Syntactic architecture and its consequences II

Between syntax and morphology

Edited by

András Bárány

Theresa Biberauer

Jamie Douglas

Sten Vikner

Open Generative Syntax 10

language science press

Open Generative Syntax

Editors: Elena Anagnostopoulou, Mark Baker, Roberta D'Alessandro, David Pesetsky, Susi Wurmbrand

In this series:

1. Bailey, Laura R. & Michelle Sheehan (eds.). Order and structure in syntax I: Word order and syntactic structure.
2. Sheehan, Michelle & Laura R. Bailey (eds.). Order and structure in syntax II: Subjecthood and argument structure.
3. Bacskai-Atkari, Julia. Deletion phenomena in comparative constructions: English comparatives in a cross-linguistic perspective.
4. Franco, Ludovico, Mihaela Marchis Moreno & Matthew Reeve (eds.). Agreement, case and locality in the nominal and verbal domains.
5. Bross, Fabian. The clausal syntax of German Sign Language: A cartographic approach.
6. Smith, Peter W., Johannes Mursell & Katharina Hartmann (eds.). Agree to Agree: Agreement in the Minimalist Programme.
7. Pineda, Anna & Jaume Mateu (eds.). Dative constructions in Romance and beyond.
8. Kastner, Itamar. Voice at the interfaces: The syntax, semantics, and morphology of the Hebrew verb.
9. Bárány, András, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.). Syntactic architecture and its consequences I: Syntax inside the grammar.
10. Bárány, András, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.). Syntactic architecture and its consequences II: Between syntax and morphology.

ISSN: 2568-7336

Syntactic architecture and its consequences II

Between syntax and morphology

Edited by

András Bárány

Theresa Biberauer

Jamie Douglas

Sten Vikner

language
science
press

Bárány, András, Theresa Biberauer, Jamie Douglas & Sten Vikner (ed.). 2020.
Syntactic architecture and its consequences II: *Between syntax and morphology*
(Open Generative Syntax 10). Berlin: Language Science Press.

This title can be downloaded at:
http://langsci-press.org/catalog/book/276
© 2020, the authors
Published under the Creative Commons Attribution 4.0 Licence (CC BY 4.0):
http://creativecommons.org/licenses/by/4.0/

ISBN: 978-3-96110-288-4 (Digital)
978-3-96110-289-1 (Hardcover)

ISSN: 2568-7336
DOI: 10.5281/zenodo.4081038
Source code available from www.github.com/langsci/276
Collaborative reading: paperhive.org/documents/remote?type=langsci&id=276

Cover and concept of design: Ulrike Harbort
Typesetting: András Bárány, Jamie Douglas, Felix Kopecky
Proofreading: Amir Ghorbanpour, Amy Lam, Andreas Hölzl, Bev Erasmus, Christian Döhler, George Walkden, Ikmi Nur Oktavianti, Jeroen van de Weijer, Jessica Brown, Lachlan Mackenzie, Madeline Myers, Jean Nitzke, Radek Šimík, Teodora Mihoc, Tom Bossuyt
Fonts: Libertinus, Arimo, DejaVu Sans Mono, Source Han Serif
Typesetting software: XeLaTeX

Language Science Press
xHain
Grünberger Str. 16
10243 Berlin, Germany
langsci-press.org

Storage and cataloguing done by FU Berlin

You say you want a revolution
Well you know
We all want to change the world
You tell me that it's evolution
Well you know
We all want to change the world

Don't you know it's gonna be alright

— The Beatles, *Revolution 1*

Contents

Introduction

András Bárány
Bielefeld University

Theresa Biberauer
University of Cambridge, Stellenbosch University, University of the West Cape

Jamie Douglas
University of Cambridge

Sten Vikner
Aarhus University

The three volumes of *Syntactic architecture and its consequences* present contributions to comparative generative linguistics that "rethink" existing approaches to an extensive range of phenomena, domains, and architectural questions in linguistic theory. At the heart of the contributions is the tension between descriptive and explanatory adequacy which has long animated generative linguistics and which continues to grow thanks to the increasing amount and diversity of data available to us. As the three volumes show, such data from a large number of understudied languages as well as diatopic and diachronic varieties of well-known languages are being used to test previously stated hypotheses, develop novel ideas and expand on our understanding of linguistic theory.

The volumes feature a combination of squib- and regular-length discussions addressing research questions with foci which range from micro to macro in scale. We hope that together, they provide a valuable overview of issues that are currently being addressed in generative linguistics, broadly defined, allowing readers to make novel analogies and connections across a range of different research strands. The chapters in Volume 1, *Syntax inside the grammar*, and Volume 3, *Inside syntax*, address research topics both at the syntactic interfaces and

András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner. 2020. Introduction. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, v–vii. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280625

in syntax proper, such as language change, complexity, and variation, as well as alignment types, case, agreement, and the syntax of null elements.

The contributions to the present, second volume, *Perspectives from morphosyntax*, address research questions and developments in morphosyntax. The volume is divided into two parts, dealing with architectural (Part I) and structural issues in morphosyntax (Part II).

The chapters in Part I, *Architectural issues in morphosyntax*, take on classic issues in grammar and provide new perspectives on questions such as universality and variation (Watumull & Chomsky), language evolution and variation (Grohmann & Leivada), as well as the architectural underpinnings of recent syntactic theory. These involve the role of the structure-building operation Merge (Zeijlstra; Moro) as well as the structure-removing operation Remove (Müller), and cross-linguistic questions relating to labelling (Tsoulas), the nature of linearisation (Johnson), phases and cyclicity (Gallego), phrase structure (Lasnik & Stone), and constraints on extraction from conjuncts and adjuncts (Bošković). Myler's chapter explores how formal syntax can make predictions about surface frequencies in word order variation, while the age-old question of lexical and syntactic categories is addressed from different perspectives in the chapters by Brandner, Kenesei, and Moro.

Part II, *Structural issues in morphosyntax*, starts with chapters reconsidering properties of relative pronouns and relative clauses (Daskalaki; Douglas). The following chapters deal with second-position and third-position effects in constituent order (Mitrović; Meelen, Mourigh & Cheng). Several contributions deal with the structure of and microvariation in noun phrases, for example, with respect to demonstratives (Cinque; Ledgeway; Kinn), and the properties and syntactic representation of person splits in Romance (Manzini and Savoia), as well as microvariation in passives in varieties of Dutch (Haegeman).

Taken together, then, the contributions to this volume, many of which have clearly been influenced and inspired by Roberts (2010; 2012), Roberts & Roussou (2003), Roberts & Holmberg (2010), Biberauer & Roberts (2012; 2015), and Biberauer et al. (2014) give the reader a sense of current research into morphosyntax and morphosyntactic variation.

References

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.

Biberauer, Theresa & Ian Roberts. 2012. On the significance of what hasn't happened. Paper presented at the 14th Diachronic Generative Syntax Conference, Lisbon.

Biberauer, Theresa & Ian Roberts. 2015. Clausal hierarchies. In Ur Shlonsky (ed.), *Beyond functional sequence* (The cartography of syntactic structures 10), 295–313. Oxford: Oxford University Press.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.

Roberts, Ian. 2012. Macroparameters and minimalism: A programme for comparative research. In Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo & Juanito Avelar (eds.), *Parameter theory and linguistic change*, 320–335. Oxford: Oxford University Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.

Part I

Architectural issues in morphosyntax

Chapter 1

Rethinking universality

Jeffrey Watumull
Oceanit

Noam Chomsky
University of Arizona, Massachusetts Institute of Technology

For a discrete infinity of reasons, Ian Roberts is to be celebrated. Here we discuss how his important work has caused us to rethink what could be, arguably, the most unbelievable and extraordinary aspect of language: its *universality*. In particular, we proffer Roberts' theory of parameter hierarchies to corroborate an *economy thesis* – a thesis implying that the quiddities of language transcend *human* language, and would obtain of *any* language *anywhere* in the universe.

1 Beyond the infinite

> As far as anyone knows, spaceships have been successfully built by exactly one civilisation in the entire history of the universe: by post-1957 humans (the Space Age actually happens to coincide exactly with my lifetime, although I had nothing to do with it) (Roberts 2017: 1)

Ian Roberts may not have been amongst those to *engineer* the Space Age, but he is one of the best to have *explained* (indirectly) how it was possible, and *explanation* is the prerequisite for all progress in scientific understanding and its technological applications. Specifically, Roberts has over his career explained how *human language* – its structure, acquisition, and historical change – has propelled our species to being the paragon of animals – to go "beyond the infinite" in Kubrick's words.

Jeffrey Watumull & Noam Chomsky. 2020. Rethinking universality. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 3–24. Berlin: Language Science Press.
DOI: 10.5281/zenodo.4280627

> Chimps, who allegedly share around 98 percent of their genes with us, […] show no interplanetary ambitions […]. Our extra 2 percent makes us extremely good – by the standards of everything else in the known universe, unbelievably, extraordinarily, *cosmically* good – at generating, storing and transmitting knowledge. How do we do it? With *language*.
>
> (Roberts 2017: 1–2)

In this, the sixth decade of Roberts' cosmic existence, we celebrate him and how his work has caused us to rethink what could be, arguably, the most unbelievable and extraordinary aspect of language: its *universality*. In particular, we proffer Roberts' theory of parameter hierarchies to corroborate an *economy thesis* – a thesis implying that the quiddities of language transcend *human* language, and would obtain of *any* language *anywhere* in the universe.

2 A universal instrument

The human mind, Descartes argued, is undoubtedly in some sense a "universal instrument". We cannot know with certainty what he intended by this provocative comment, but we do know that the Cartesians would have understood language as fundamental to any nontrivial notion of "universality" because it is language that empowers humans to generate an unbounded set of hierarchically structured expressions that can enter into effectively infinitely many thoughts and actions – that is, the competence of every human, but no beast or machine, to use language in creative ways appropriate to situations but not caused by them, and to formulate and express these thoughts coherently and without bound, perhaps "incited or inclined" to speak in particular ways by internal and external circumstances but not "compelled" to do so. Of course in the pre-Turing world, the Cartesians did not know how a finite "machine" such as the brain could generate the infinity of expressions of natural language, and therefore posited a soul where we need only posit a neurobiological Turing machine (obviously idealized with unbounded memory, etc.). Nevertheless Descartes intuited the essence of Turing universality: "Only a spiritual entity could achieve the limitlessness of interactive language, putting words together in indefinitely many ways", and to do so in ways that are "free" (i.e., not compelled by internal or external conditions) and intelligible and appropriate to situations, and to do so over an unbounded range in different domains.

> Any material machine must specialize: while a machine might do very well some of the things people do, it would necessarily be unable to do others.

> Any part or organ needed a particular configuration to achieve a task, and it was impossible to have enough different parts with the requisite configurations in a single machine to make it act in all the contingencies of life in the same way that our reason makes us act. Only disembodied reason could be 'a universal instrument'. (Riskin 2017: 63)

Of course the genius of Turing was to discover that "[i]t is possible to invent a single machine which can be used to compute any computable sequence"; he called this mathematical object, appropriately, the "universal machine" (Turing 1937: 243).

Linguistic competence (and especially its creative use), in concert with other mental faculties, establishes the general intelligence necessary for the evolutionary "great leap forward" of our species (see Chomsky 2016). As Roberts (2017: 182) conjects, "there might have been a crucial mutation in human evolution which led, in almost no time from an evolutionary perspective, from [humans living in] caves to [their creating knowledge of such sophistication as to enable us to imagine and construct things as complex as, say,] spaceships. It's a plausible speculation that the mutation in question was whatever it is that makes our brains capable of computing recursive syntax, since it's the recursive syntax that really gives language – and thought – their unlimited expressive power. It's one small step from syntax to spaceships, but a great leap for humans". A great leap for humans – and *only* humans, evidently (see Berwick & Chomsky 2016). The architecture of intelligence necessitates "provisions for recursive, hierarchical use of previous results" as manifested in the "articulation" of a complex structure into descriptions of "elementary figures" and "subexpressions designating complex subfigures", with a "figure first divided into two parts; and then with each part described using the same machinery" (Minsky 1963: 16). The recursive capacity of intelligence is most manifest in natural language:

> Whatever we can express or describe, we can treat its expression or description as though it was a single component inside another description. In languages, this corresponds to using embedded phrases and clauses. That final trick – of representing prior thoughts as things – gives our minds the awesome power to use the same brain-machinery over and over again, to replace entire conceptualizations by compact symbols, and hence to build gigantic structures of ideas the way our children build great bridges and towers from simple separate blocks. It lets us build new ideas from old ones; in short, it makes it possible to think. The same is true of our [future] computers. (Minsky 1985: 124)

Thus we might expect any (super-)human-level intelligence anywhere in the universe – including any genuine artificial intelligence ("our [future] computers") we create – to be recursive in this way.

It has been assumed that the essential properties of human language are not only unique, but *logically contingent*:

> Let us define "universal grammar" (UG) as the system of principles, conditions, and rules that are elements or properties of all human languages not merely by accident but by necessity – of course, I mean biological, not logical necessity. Thus UG can be taken as expressing "the essence of human language". (Chomsky 1975: 29)

> There is no *a priori* reason to expect that human language will have such properties; Martian could be different." (Chomsky 2000: 16)

This assumption, we submit, merits rethinking in light of Roberts' work and progress in the Minimalist program more generally (Chomsky 1995). Recent work demonstrating the *simplicity* (Watumull et al. 2017) and *optimality* (Chomsky et al. 2019) of language increases the cogency of a conjecture that at one time would have been summarily dismissed as absurd: "the basic principles of language are formulated in terms of notions drawn from the domain of (virtual) conceptual necessity", the domain defined by "general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, nonredundancy, and the like" (Chomsky 1995: 171, 1) that render linguistic computation interestingly optimal. To the extent that this *strong Minimalist thesis* (SMT) is true, the essential – computational (even mathematical) – properties of language would derive from laws of nature – language- and even biology-independent principles that, once realized in the mind/brain, *do* entail particular properties as logically necessary. For instance, it is simply a fact of logic that the simplest (optimal) form of the recursive procedure generative of syntactic structures, Merge, has two and only two forms of application (i.e., external and internal). Relatedly, *given* the nature of the structures Merge generates, minimal structure distance is *necessarily* the simplest computation for the structure dependence of rules. And so on and so forth (see Berwick et al. 2011; Chomsky 2013; Watumull 2015 for additional examples).

Research in the Minimalist program starts with the optimality conjecture and proceeds to inquire whether and to what extent it can be sustained given the observed complexities and variety of natural languages. If a gap is discovered, the task is to inquire whether the data can be reinterpreted, or whether principles of simplicity and optimal computation can be reformulated, so as to solve

the puzzles within the framework of SMT, thus generating some support, in an interesting and unexpected domain, for Galileo's precept that nature is simple and it is the task of the scientist to demonstrate it.

As we discover more and more of "the essence of human language" to be defined by (virtual) conceptual necessity, the less and less absurd it is to question just how contingent a phenomenon human language really is. It may well be with language as with other phenomena studied in the natural sciences that, in the words of the sage physicist J.A. Wheeler, "[b]ehind it all is surely an idea so simple, so beautiful, that when we grasp it – in a decade, a century, or a millennium – we will all say to each other, how could it have been otherwise?" (Wheeler 1986: 386). In other words, there may well be some a priori reasons to expect human language to have the (essential) properties it does; or, to put it whimsically, the Martian language might *not* be so different from human language after all. In short, the *universality* of universal grammar needs to be rethought.

3 Simplicity itself

Our rethinking is based on a rethinking – or reminding – of *simplicity* as originally conceived in generative linguistics. "[S]implicity, economy, compactness, etc." were proffered in the first work on generative grammar as criteria the grammar of a language must satisfy: "Such considerations are in general not trivial or "merely esthetic". It has been recognized of philosophical systems, and it is, I think, no less true of grammatical systems, that the motives behind the demand for economy are in many ways the same as those behind the demand that there be a system at all" (Chomsky 1951: 1, 67). This proposition echoed that of Goodman (1943: 107): "The motives for seeking economy in the basis of a system are much the same as the motives for constructing the system itself". The idea is elementary but profound: if the theory is no more simple, economical, compact, etc. than the data it is proffered to explain, it is not a theory at all; hence the more compressed the theory, the more successful – i.e., the more explanatory – it is.

The mathematician Gregory Chaitin (2005: 64) has formalized this idea in terms of algorithmic information theory: "a scientific theory [can be thought of] as a binary computer program for calculating observations, which are also written in binary"; a generative grammar can thus be thought of as a program for generating syntactic structures. "And you have a law of nature if there is compression, if the experimental data is compressed into a computer program", equivalently a grammar, "that has a smaller number of bits than are in the data that it explains", or generates. "The greater the degree of compression, the better

the law, the more you understand the data. But if the experimental data cannot be compressed, if the smallest program for calculating it is just as large as it is [...], then the data is lawless, unstructured, patternless, not amenable to scientific study, incomprehensible. In a word, random, irreducible". In the terms of generative grammar (Chomsky & Miller 1963: 285):

> As a matter of principle, a grammar must be finite. If we permit ourselves grammars with an unspecifiable set of rules[,] we can simply adopt an infinite sentence dictionary. But that would be a completely meaningless proposal. Clearly, a grammar must have the status of a theory about those regularities that we call the syntactic structure of the language.

To have the status of a theory, the grammar must be compressed, generating – and thereby explaining – the regularities in syntactic structures.

This idea is appreciated surprisingly seldom today: many computational cognitive scientists and machine learning theorists (and hence virtually all "artificial intelligence" (AI) labs in academia and industry) have perversely redefined a successful theory or computer program to be one that merely approximates or classifies unanalyzed data. This contrasts dramatically with the Enlightenment definition in which data are selectively analyzed as evidence for/against conjectured explanations (see Popper 1963; Chomsky 2000; Deutsch 2011). The machine learning systems (e.g., deep learning neural nets, reinforcement learning techniques, etc.) so popular in the current "AI spring" are *weak AI*: brute-force systems laboriously trained to "unthinkingly" associate patterns in the input data to produce outputs that approximate those data in a process with no resemblance to human cognition (thus betraying Turing's original vision for AI). These systems will never be genuinely intelligent, and are to be contrasted with the *strong – anthronoetic – AI* Turing envisioned: a program designed to attain human-level competence with a *human-style* typified by *syntactic generativity* and *semantic fluidity* – to think *the way* a human thinks. Today such programs, based on generative grammars, are finally being built.[1]

The early discussions on simplicity were addressing the logic of theory construction by the scientist, but later (Chomsky 1965: 4) this logic was analogized to the learning of language by children: "The problem for the linguist, as well as for the child learning the language, is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer". To determine the grammar (qua "theory" *in* the mind of the learner and qua theory *of* the mind by the linguist), some procedure to evaluate candidate grammars

[1]https://www.oceanit.com/science-technology/artificial-intelligence/

is necessary. Specifically, a format-evaluation framework: "(v) specification of a function m such that $m(i)$ is an integer associated with the grammar G_i as its value (with, let us say, lower value indicated by higher number)" (Chomsky 1965: 31). Naturally, "simpler" grammars are more highly valued, but, then as now, "simplicity" is complex: "In the context of this discussion, 'simplicity' (that is, the evaluation measure m of (v)) is a notion to be defined within linguistic theory along with "grammar", "phoneme", etc. Choice of simplicity measure is rather like determination of the value of a physical constant" (Chomsky 1965: 37–38). Goodman (1943: 107–108) too was cognizant of the complexity of simplicity, observing that "the mere counting of primitives is no satisfactory measure" because "by the purely mechanical application of certain logical devices, we can readily reduce all the primitives of any system to one". Thus while Goodman searched for a general notion of simplicity applicable to all systems, a specific notion applicable to language was sought in generative linguistics, and both ultimately "failed" (i.e., superseded by better notions – characteristic of a healthy science): the former for technical reasons, the latter because of the success of the principles-and-parameters (P&P) framework (Chomsky 1981), which obviated the need for any simplicity measure of the type envisioned for the format-evaluation framework.

4 The principles-and-parameters mission

In P&P, language acquisition is the process of setting the values for the finitely many universal parameters of the initial state of the language faculty (UG). The apparent complexity and diversity of linguistic phenomena is illusory and epiphenomenal, emerging from the interaction of invariant principles under varying conditions. This was a radical shift from the early work in generative linguistics, which sought only an evaluation measure that would select among alternative theories of a language (grammars) – the simplest congruent with the format encoded in UG and consistent with the primary linguistic data. But with the P&P shift in perspective, simplicity can be rethought, though this was not initially appreciated. As discussed in the earliest work in generative linguistics, notions of simplicity assume two distinct forms: the imprecise but profound notion of simplicity that enters into rational inquiry generally, and the theory-internal measure of simplicity that selects among I-languages. The former notion of simplicity is language-independent, but the theory-internal notion is a component of UG, a subcomponent of the procedure for determining the relation between experience and I-language (again, something like a physical constant). In early work, the internal notion was implemented in the form of the evaluation procedure to select among proposed grammars/I-languages consistent with the UG format

for rule systems. But, as Ian Roberts (2012) and others (e.g., Sheehan et al. 2017) discovered, the P&P approach transcends that limited, parochial conception of simplicity: with no evaluation procedure, there is no internal notion of simplicity in the earlier sense. There remains only the universal notion of simplicity.

In P&P, grammars – I-languages – are simple, but, as evidenced in Roberts' work (e.g., Roberts & Holmberg 2010), they are so by virtue of third-factor principles of computational efficiency (Chomsky 2005), not by analogy to theory-construction or by stipulation in UG. In fact, rather than "simple", we propose to define P&P-style acquisition as "economical", which, in the Leibnizian spirit, we understand to subsume simplicity:

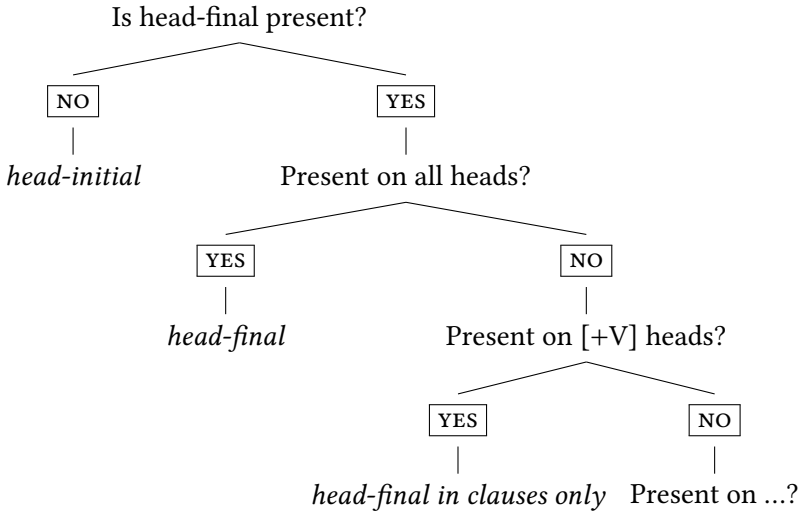
> The most economical idea, like the most economical engine, is the one that accomplishes most by using least. Simplicity – or fuel consumption – is a different factor from power [i.e., generative capacity, empirical coverage, etc.] but has to be taken equally into consideration […]. The economy of a basis may be said to be the ratio of its *strength* to its simplicity. But superfluous power is also a waste. Adequacy for a given system is the only relevant factor in the power of a basis; and where we are comparing several alternative bases for some one system, as is normally the case, that factor is a constant. Thus in practice the simplest basis is the most economical.
>
> (Goodman 1943: 111)

Economy, in other words, is a *minimax* notion. In Leibniz's words (see Roberts & Watumull 2015): "the simplicity of the means counterbalances the richness of the effects" so that in nature "the maximum effect [is] produced by the simplest means". This notion is enshrined in the Galilean ideal (see Chomsky 2002).

One economical form of P&P-style learning explicable in terms of third-factors is the traversal of a parameter hierarchy (see Roberts 2012; Biberauer 2016) – parameter specification. In such a system, the child is not unthinkingly enumerating and evaluating grammars.[2] Instead, the I-language matures to a steady state in a relatively deterministic process of "answering questions" that *emerge naturally and necessarily* in the sense that there exist "choices" in acquisition that logically must be "made" for the system to function at all; none of the parameters need be encoded in the genetic endowment (see Obata et al. 2015 for similar ideas). This is the ideal, of course. Like SMT generally, how closely it can be approximated is an empirical matter, and there remain many challenges.

[2]Such an inefficient and unintelligent technique is the modus operandi of many machine learning (weak AI) systems.

Parameter specification – i.e., the P&P-conception of "learning" as the specification of values for the variables in I-language – can be schematized as a decision tree (parameter hierarchy) which, as Roberts has shown, is governed by minimax economy: minimizing formal features (feature-economy) coupled with maximizing accessible features (input-generalization). Traversal of a hierarchy – a conditional-branching Turing machine program – is inevitably economical in that the shortest (in binary) and most general parameter settings are necessarily "preferred" in the sense that the faster the computation halts, the shorter the parameter settings. For instance, to specify word-order, a series of binary queries with answers of increasing length and decreasing generality (microparameters) is structured thus:



For compatibility with computability theory and Boolean logic, the parameter hierarchy can be translated as follows:

(1) Hierarchy: H
State T: Decision problem
Yes: 0/1 (0 = transition to state T+1) (1 = halt and output parameter specification for H)
No: 0/1 (0 = transition to state T+1) (1 = halt and output parameter specification for H)

(2) Hierarchy: Word order
State 1: Is head-final present?
Yes: Output 0 (transition to State 2)
No: Output 1 (halt and output "head-initial")
State 2: Present on all heads?
Yes: Output 1 (halt and output "head-final")
No: Output 0 (transition to State 3)
State 3: Present on [+V] heads?
Yes: Output 1 (halt and output "head-final in clause only")
No: Output 0 (transition to State 4)
…

So in P&P, the logic is not "enumerate and evaluate" with stipulative (theory-internal) simplicity measures: it is "compute all and only what is necessary", which implies the language-independent reality of economy in that, as with the parameter hierarchies, the process answers all and only the questions it needs to. It is not that there is any explicit instruction in the genetic endowment to prefer simple answers: it is simply otiose and meaningless to answer unasked questions (i.e., once the parameters are set, the computation halts).[3]

Moreover the "answers" to "questions" can be represented in binary. Indeed binary is a *notation-independent* notion necessary and sufficient to *maximize* computation with *minimal* complexity: functions of arbitrarily many arguments can be realized by the composition of binary (but not unary) functions – a truth of minimax logic with "far-reaching significance for our understanding of the functional architecture of the brain" (Gallistel & King 2010: x). The mathematical and computational import of binary was rendered explicit in the theories of Turing (1937) and Shannon (1948), the former demonstrating the necessarily digital – hence ultimately binary – nature of *universal computation* (a universal Turing machine being the most general mathematical characterization of computation); the latter formalizing *information* in terms of *bits* (binary digits). The consilience of these ideas is our economy thesis: human language is based on simple representations (i.e., bits) and strong computations (i.e., the binary functions of Turing machines) – and "economy of a basis may be said to be the ratio of its *strength* to its simplicity" (Goodman 1943: 111).

[3] In this way it is trivial to derive Ockham's razor from virtual conceptual necessity. If the law of parsimony is not to multiple entities beyond necessity, and language conforms to conceptual necessity, then ergo it is maximally parsimonious. As Wittgenstein (1922) observed: "Ockham's maxim is, of course, not an arbitrary rule, nor one that is justified by its success in practice: its point is that unnecessary units in a sign-language mean nothing" (5.47321); "If a sign is *useless*, it is meaningless. That is the point of Ockham's maxim" (3.328).

5 Universal economy

As one of the "general considerations of conceptual naturalness that have some independent plausibility", economy would be a factor that obtains of any optimally "designed" (natural or artificial) computational system. So, rethinking universality, if the Martian language were optimal in the sense of conforming to virtual conceptual necessity, then it might be surprisingly similar to human language. In point of fact, we ought not to be too surprised. It is now well established by biologists that *convergence* is a common theme in any evolutionary process:

> the number of evolutionary end-points is limited: by no means is everything possible. [Because of evolutionary convergence,] what is possible usually has been arrived at multiple times, meaning that the emergence of the various biological properties is effectively inevitable.
>
> (Conway Morris 2013: xii–xiii)

Indeed, the paleontologist Simon Conway Morris argues that human-style intelligence was effectively inevitable given the initial conditions of evolution on Earth. And there is no reason a priori to assume that the principle of evolutionary convergence is unique to the biology of a particular planet. Quite the contrary, if we accept the rational form of inquiry in which the principle is understood abstractly in a computational framework. The idea is that *any* computational system *anywhere* made of *anything* is governed by *laws* of computation. As the cognitive scientist C.R. Gallistel and computer scientist Adam King argue persuasively (Gallistel & King 2010: 167):

> The functional structure of modern computers is sometimes discussed by neuroscientists as if it were an accidental consequence of the fact that computing circuits are constructed on a silicon substrate and communicate by means of pulses of electrical current sent over wires. Brains are not computers, it is argued, because computers are made of silicon and wire, while brains are made of neurons. We argue that, on the contrary, several of the most fundamental aspects of the functional structure of a computer are dictated by the logic of computation itself and that, therefore, they will be observed in any powerful computational device, no matter what stuff it is made of. In common with most contemporary neuroscientists, we believe that brains are powerful computational devices. We argue, therefore, that those aspects of the functional structure of a modern computer that are dictated by the logic of computation must be critical parts of the functional structure of brains. (Gallistel & King 2010: 167)

This argument simply reiterates Turing's (1950: 446) thesis that "[i]f we wish to find such similarities [as may exist between minds and machines] we should look [not at their substrates, but] rather for mathematical analogies of function". And given this universality of the functional, mathematical architecture of computation, it is possible that we may need to rethink how uniquely human or even uniquely biological our modes of mental computation really are. One interesting implication is that we must rethink any presumptions that extraterrestrial intelligence or artificial intelligence would really be all that different from human intelligence.

So we assume that human language is a computational process that can be characterized by a Turing machine (see Watumull 2015). It is possible to explore the space of all possible Turing machines (i.e., the space of all possible computer programs), not exhaustively of course, but with sufficient breadth and depth to make some profound discoveries. The late Marvin Minsky, founder of the artificial intelligence laboratory at MIT, and his student Daniel Bobrow, once enumerated and ran some thousands of the simplest Turing machines (computer programs with minimal numbers of rules). Intriguingly, out of the infinity of possible behaviors, only a surprisingly small subset emerged. These divided into the trivial and the nontrivial. The boring programs either halted immediately or erased the input data or looped indefinitely or engaged in some similar silliness. The remainder, however, were singularly interesting: *all* of these programs executed an effectively *identical* counting function – a primitive of elementary arithmetic. In fact, this operation reduces to a form of Merge (see Chomsky 2008). More generally, these "A-machines" (*A* for *arithmetic*) prove a point:

> [I]t seems inevitable that, somewhere, in a growing mind some A-machines must come to be. Now, possibly, there are other, really different ways to count. So there may appear, much, much later, some of what we represent as 'B-machines' – which are processes that act in ways which are similar, but not identical to, how the A-machines behave. But, our experiment hints that even the very simplest possible B-machine will be so much more complicated that it is unlikely that any brain would discover one before it first found many A-machines. (Minsky 1985: 121)

> Let us think of this exploration as exposing parts of some infinite 'universe of possible computational structures'. Then this tiny fragment of evidence suggests that such a universe may look something like [Figure 1.1]. (Minsky 1985: 120)

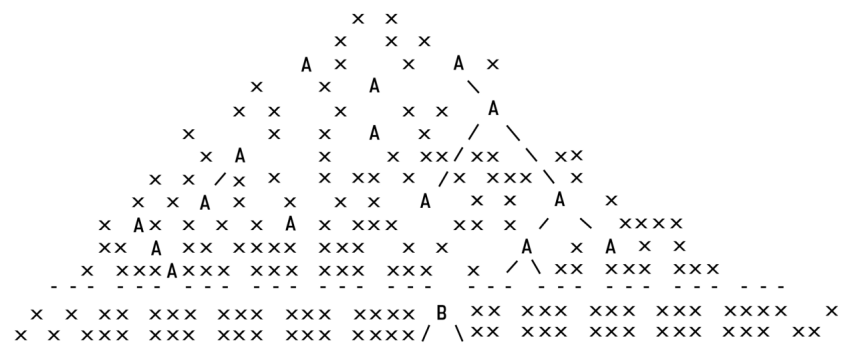


Figure 1.1: Representation of a universe with "A" and "B-machines" (Minsky 1985: 120)

This is evidence that arithmetic – the foundation of any mathematical/computational system – as represented in an A-machine – reducible to Merge – is technically an *attractor* in the *phase space* of possible mathematical structures:

> any entity who searches through the simplest processes will soon find fragments which do not merely resemble arithmetic but *are* arithmetic. It is not a matter of inventiveness or imagination, only a fact about the geography of the universe of computation. (Minsky 1985: 122)

Curiously, some physicists have argued that human mathematics is contingent: "the next batch of aliens might turn out to be different" (Alford 2006: 774), with no recognizable rules or systems. This objection echoes once regnant dogma in linguistics that "[human] languages could differ from each other without limit and in unpredictable ways" such that linguists ought to proceed "without any preexistent scheme of what a language must be" (Joos 1957: 96, v), implying that any two human languages could be as different from each other as any one could be from an alien language. But this dogma could not withstand critical scrutiny, and was dispelled with the advent of generative linguistics and its formulation of universal grammar – the theory of the abstract grammatical system encoded genetically in *Homo sapiens sapiens* – and crucially by the deeper empirical inquiries into the languages of the world undertaken within the framework of generative grammar (e.g., the spectacular demonstration that Warlpiri, contrary to all appearances, has the standard hierarchical structures universal to natural languages (see Hale 1976; Legate 2001). To the extent that SMT is true, general properties derivative of this formal system define the properties universal to particular languages. Therefore we should indeed study these particular languages

with a "preexistent scheme of what a language must be" because UG and general principles of computation constrain the space of possible linguistic properties. And thus languages could not "differ from each other without limit", but only in "[predictable] ways".

The thesis that arithmetic is an *attractor* in the *phase space* of possible mathematical structures obviously generalizes beyond arithmetic to all simple computations (see Wolfram 2002 for countless examples). "Because of this, we can expect certain 'a priori' structures to appear, almost always, whenever a computational system evolves by selection from a universe of possible processes" (Minsky 1985: 119). Analogously, we submit that it is not implausible that an evolutionary search through the simplest computations will soon find something like Merge. Merge is an operation so elementary as to be subsumed somehow in every more complex computational procedure: take two objects X and Y already constructed and form the object Z without either modifying X or Y, or imposing any additional structure on them: thus Merge(X, Y) = {X, Y}.[4] This simple assumption suffices to derive in a principled (necessary) way a complex array of otherwise arbitrary (contingent) phenomena such as the asymmetry of the conceptual-intentional and sensory-motor interfaces (entailing the locus of surface complexity and variety), the ubiquity of dislocation, structure-dependence, minimal structural distance for anaphoric and other construals, the difference between what reaches the mind for semantic interpretation and what reaches the apparatus of articulation and perception (see Chomsky 2017).

6 The dawn of language

As we discussed in terms of our economy thesis, simplicity can be defined in algorithmic information theory (or the theory of program-size complexity): the complexity of a program is measured by its maximally compressed length in bits so that the simplest program is that with the shortest description. A search of the phase space of possible programs, whether conducted consciously (e.g., by us, extraterrestrials, etc.) or unconsciously (e.g., by modern computers, evolution, etc.) automatically proceeds in size order from the shortest and increasing to programs no shorter than their outputs (these incompressible programs are effectively lists); many complex programs would subsume simpler programs as the real numbers subsume the natural numbers. And, as demonstrated logically and empirically, "any evolutionary process must first consider relatively simple

[4] This formulation of Merge requires some rethinking in ways that we can put aside here (see Watumull et al. in press for discussion).

systems, and thus discover the same, isolated, islands of efficiency" (Minsky 1985: 122). Why are the simple systems (e.g., Merge) so sparsely distributed in the phase space of possible processes? (Why are they "islands" in the computational universe?) Why are there no "similar" processes in the neighborhood? (There is not something "like" arithmetic out there: there is just arithmetic, "cold and austere, […] yet sublimely pure, and capable of a stern perfection such as only the greatest art can show" in Bertrand Russell's words.) The answer must be that small sets of rules (e.g., Merge) can generate unbounded complexity, but the converse is not in general true: it is simply a mathematical fact (a tautology) that there is only a small set of small sets of rules, and thus not all complex phenomena can be generated by small sets of rules (there is simply not a sufficient number of small sets of rules "to go around"). This explains why, for instance, one cannot fiddle with arithmetic: one cannot posit its simple rules, generate a universe of consequences, and then make changes to that universe and expect the simple rules to cover the "revised" universe (e.g., one cannot remove a number or change a sum, product, etc.). Analogously, having posited Merge and executed it to generate the discrete infinity of syntactic structures, one cannot modify the logic (e.g., structure dependence) that obtains of those structures by dint of their having been generated by Merge and still expect Merge to generate new structures that conform to the modified logic, for the modified system is now "miraculous" in the technical sense of possessing properties that did not emerge from the rules themselves (or nonarbitrary third factors, i.e., laws of nature). And there cannot be infinitely many sets of small rules in the neighborhood of Merge to produce the effect of continuity. Thus there can only be *islands* of computation, not *continents*.

Thus it may well be that, given the universal and invariant laws of evolution, convergence on systems – Turing machines – virtually identical to those "discovered" in our evolutionary history is inevitable.[5] Hence our rethinking the proposition "Martian could be different".

The fact that simple computations are attractors in the phase space of possible computations goes some way to explaining why language should be optimally designed (insofar as SMT holds) in that an evolutionary search is likely to converge on it, which leads us to consideration of the origin of language. Convergence is a consequence of constraints. As with intelligence, evolution and development are possible only by coupling scope with constraints. Stated generally: the scope

[5]Indeed we might speculate that were we to "wind the tape of life back" and play it again, in Stephen Jay Gould's phrasing, not only would something like Merge reemerge, but something like humans could well be "inevitable", as some biologists have suggested (see Conway Morris 2013).

of any creative process is a function of its operating within limits. In the context of evolution, for instance, Stuart Kauffman (1993: 118) observes,

> Adaptive evolution is a search process – driven by mutation, recombination, and selection – on fixed or deforming fitness landscapes. An adapting population flows over the landscape under these forces. The structure of such landscapes, smooth or rugged, governs both the evolvability of populations and the sustained fitness of their members. The structure of fitness landscapes inevitably imposes limitations on adaptive search.

The analogy to mind is deeply nontrivial, for "intellectual activity consists mainly of various kinds of search" (Turing 1948: 431).

The evolution of language is mysterious (see Hauser et al. 2014), but SMT is consistent with the limited archeological evidence that does exist on the emergence of language, evidently quite recently and suddenly in the evolutionary time frame (see Tattersall 2012).[6] Furthermore there is compelling evidence for SMT in the design of language itself. For instance, it is a universal truth of natural language that the rules of syntax-semantics are structure-dependent (see Berwick et al. 2011): hierarchy, not linearity, is determinative in the application of rules and interpretation of expressions. This implies a far-reaching thesis with many consequences: linear order is a peripheral property of language, emerging only in externalization at the sensory-motor interface (where serial ordering is necessary). If this thesis holds generally, then Aristotle's dictum that language is "sound with meaning" should be revised: language is not sound with meaning, but rather meaning with sound (or some other modality of externalization), a very different concept, reflecting a different traditional idea: that language is fundamentally an instrument of thought – "audible thinking", "the spoken instrumentality of thought", as William Dwight Whitney expressed the traditional conception (see Chomsky 2013), consistent with the Cartesian idea that language is a central component of our mind as a "universal instrument", endowing us with general intelligence. As François Jacob suggested (see Berwick & Chomsky 2011), plausibly, "the role of language as a communication system between individuals would have come about only secondarily" to the emergence of generative syntax (Merge, we would now say) and its mapping of structures to the conceptual-intentional system for semantic interpretation. "The quality of language that makes it unique does not seem to be so much its role in communicating directives for action" or other typical features of animal communication, but rather

[6] There is quite compelling evidence that since the trek of our ancestors from Africa some 50,000 years ago, the language faculty has undergone no significant change, and not very long before (in evolutionary time) there is no evidence that it existed at all.

"its role in symbolizing, in evoking cognitive images", in molding our notion of reality and yielding our capacity for thought and planning, through its unique property of allowing "infinite combinations of symbols" and therefore "mental creation of possible worlds". Thus the most reasonable speculation today – and one that opens productive lines of research – is that from some simple rewiring of the brain, Merge emerged, naturally in its simplest form, providing the basis for unbounded and creative thought – the "great leap forward" evidenced in the archeological record and in the remarkable differences distinguishing modern humans from their predecessors and the rest of the animal kingdom (see Huybregts 2017; Berwick & Chomsky 2016 for in-depth discussion of these topics).

If this conjecture can be sustained, we could answer the question why language should be optimally designed: optimality would be expected under the postulated conditions, with no selectional or other pressures operating; the emerging system should just follow the laws of nature such as minimal computation and more "general considerations of conceptual naturalness that have some independent plausibility, namely, simplicity, economy, symmetry, nonredundancy, and the like" – rather the way a snowflake forms. If this is correct, then, contrary to what was once presumed, there *would* be a priori reasons to expect any language anywhere in the universe would resemble human language; the "principles, conditions, and rules that are elements or properties of all human languages" *would* be *logically* necessary, deriving from laws of nature. And so, just as physicists seek "an idea so simple, so beautiful, that [...] we will all say to each other, how could it have been otherwise?", in the study of language we search for – and are discovering – objects of great beauty and simplicity.

7 The wonders of language

> It is [...] quite possible that we, as a species, have crossed a cognitive threshold. Our capacity to express anything, through the recursive syntax and compositional semantics of natural language, might have taken us into a cognitive realm where anything, everything, is possible. Effectively, having language has made us the equal of any extraterrestrial.
>
> (Roberts 2017: 181–182)

Notwithstanding the universal logic of computation, it is obviously necessary that there exist *constraints* on the mind if it is to have any *scope* at all, and these constraints may very well be uniquely human. Taking the extreme case, suppose that the human mind is a universal Turing machine (see Watumull 2015).

Such a mind could be a *universal explainer*. The argument is simple: a universal Turing machine can emulate any other Turing machine (i.e., a universal computer can run any program); a program is a kind of theory (written to be readable/executable by a computer); thus a universal Turing machine can compute any theory; and thus, assuming that everything in the universe could in principle be explained by and understood within some theory or other (in other words, assuming no magic, miracles, etc.), a universal Turing machine – a Turing-universal mind – could explain and understand everything. It is an intriguing conclusion, and not obviously false, but numerous objections could be posed. For instance,

> an arbitrary Turing machine, or an unrestricted rewriting system, is too unstructured to serve as a grammar […]. Obviously, a computer program that succeeded in generating sentences of a language would be, in itself, of no scientific interest unless it also shed some light on the kinds of structural features that distinguish languages from arbitrary, recursively enumerable sets. (Chomsky 1963: 360)

Beyond language, if a Turing-universal mind is to be a universal explainer, it should not generate all possible explanations, true and false, because that would be merely to restate the problem of explaining nature: deciding which in an infinite set of explanations are the true (or best) explanations is as difficult as constructing the best explanations in the first place. There must be "limits on admissible hypotheses", in the words of Charles Sanders Peirce (see Chomsky 2006). This interdependence of scope and limits has been expounded by many creative thinkers and analyzed by (creative) philosophers of esthetics: the beauty of jazz emerges not by "playing anything", but only when the improvisation is structured, canalized; the beauty of a poem is a function of its having to satisfy the constraints of its form, as the mathematician Stanislaw Ulam (1976: 180) observed,

> When I was a boy I felt that the role of rhyme in poetry was to compel one to find the unobvious because of the necessity of finding a word which rhymes. This forces novel associations and almost guarantees deviations from routine chains or trains of thought. It becomes paradoxically a sort of automatic mechanism of originality.

Thus from science to art, we see that the (hypothesized) infinite creativity of the Turing-universal human mind is non-vacuous and useful – and beautiful – only if it operates within constraints – constraints that appear to be uniquely human.

> So understanding language means understanding a very big part of what it is to be human, what it is to be you. And that is perhaps the greatest wonder of language of all. (Roberts 2017: 182)

The wonders of language Ian Roberts has illuminated are beyond counting; we have surveyed but a twinkling here. Indeed, of his work we might say, in closing, "my God! – *it's full of stars*!" (Clarke 1968: 202).

Abbreviations

SMT	strong Minimalist thesis	UG	Universal Grammar

References

Alford, Mark. 2006. On math, matter, and mind. *Foundations of Physics* 36(6). 765–794. DOI: 10.1007/s10701-006-9048-x.

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its evolution. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Berwick, Robert C. & Noam Chomsky. 2016. *Why only us: Language and evolution*. Cambridge, MA: MIT Press.

Berwick, Robert C., Paul Pietroski, Beracah Yankama & Noam Chomsky. 2011. Poverty of the stimulus revisited. *Cognitive Science* 35(7). 1207–1242. DOI: 10.1111/j.1551-6709.2011.01189.x.

Biberauer, Theresa. 2016. Going beyond the input (and UG): An emergentist generative perspective on syntactic variation, stability and change. Paper presented at The Emergence of Modern Hebrew Workshop, Hebrew University of Jerusalem, 6 July 2016.

Chaitin, Gregory. 2005. *Meta math: The quest for omega*. New York: Vintage Books.

Chomsky, Noam. 1951. *The morphophonemics of modern Hebrew*. University of Pennsylvania. (MA thesis).

Chomsky, Noam. 1963. Formal properties of grammars. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, vol. II, 323–418. New York: Wiley & Sons.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1975. *Reflections on language*. New York: Pantheon.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2002. *On nature and language*. Cambridge: Cambridge University Press.
Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.
Chomsky, Noam. 2006. *Language and mind*. Cambridge: Cambridge University Press.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.
Chomsky, Noam. 2016. *What kind of creatures are we?* New York: Columbia University Press.
Chomsky, Noam. 2017. The language capacity: Architecture and evolution. *Psychonomic Bulletin & Review* 24(1). 200–203. DOI: 10.3758/s13423-016-1078-6.
Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions and challenges. *Catalan Journal of Linguistics*. DOI: 10.5565/rev/catjl.288.
Chomsky, Noam & George Miller. 1963. Introduction to the formal analysis of natural languages. In R. Duncan Luce, Robert R. Bush & Eugene Galanter (eds.), *Handbook of mathematical psychology*, vol. II, 269–321. New York: Wiley & Sons.
Clarke, Arthur C. 1968. *2001: A space odyssey*. New York: Penguin.
Conway Morris, Simon. 2013. *Inevitable humans in a lonely universe*. Cambridge: Cambridge University Press.
Deutsch, David. 2011. *The beginning of infinity: Explanations that transform the world*. New York: Viking.
Gallistel, Charles Randy & Adam Philip King. 2010. *Memory and the computational brain: Why cognitive science will revolutionize neuroscience*. New York: Wiley–Blackwell.
Goodman, Nelson. 1943. On the simplicity of ideas. *Journal of Symbolic Logic* 8(4). 107–121. DOI: 10.2307/2271052.

Hale, Kenneth L. 1976. The adjoined relative clause in Australia. In R.M.W. Dixon (ed.), *Grammatical categories in Australian languages*, 78–105. Canberra: Australian Institute of Aboriginal Studies.

Hauser, Marc D., Charles D. Yang, Robert C. Berwick, Ian Tattersall, Michael J. Ryan, Jeffrey Watumull, Noam Chomsky & Richard C. Lewontin. 2014. The mystery of language evolution. *Frontiers in Psychology* 5. DOI: 10.3389/fpsyg.2014.00401.

Huybregts, Riny. 2017. Phonemic clicks and the mapping asymmetry: How language emerged and speech developed. *Neuroscience and Behavioral Reviews* 81(B). 279–294. DOI: 10.1016/j.neubiorev.2017.01.041.

Joos, Martin. 1957. *Readings in linguistics: The development of descriptive linguistics in America since 1925*. Washington, DC: American Council of Learned Societies.

Kauffman, Stuart. 1993. *Self-organization and selection in evolution*. Oxford: Oxford University Press.

Legate, Julie Anne. 2001. The configurational structure of a nonfigurational language. *Linguistic Variation Yearbook* 1(1). 61–104. DOI: 10.1075/livy.1.05leg.

Minsky, Marvin. 1963. Steps towards artificial intelligence. In Edward A. Feigenbaum & Julian Feldman (eds.), *Computers and thought*, 406–450. New York: McGraw–Hill.

Minsky, Marvin. 1985. Why intelligent aliens will be intelligible. In Edward Regis (ed.), *Extraterrestrials: Science and alien intelligence*, 117–128. Cambridge, MA: MIT Press.

Obata, Miki, Samuel Epstein & Marlyse Baptista. 2015. Can crosslinguistically variant grammars be formally identical? Third factor underspecification and the possible elimination of parameters of UG. *Lingua* 156. 1–16. DOI: 10.1016/j.lingua.2014.12.003.

Popper, Karl. 1963. *Conjectures and refutations*. New York: Routledge.

Riskin, Jessica. 2017. *The restless clock: A history of the centuries-long argument over what makes living things tick*. Chicago: University of Chicago Press.

Roberts, Ian. 2012. Macroparameters and minimalism: A programme for comparative research. In Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo & Juanito Avelar (eds.), *Parameter theory and linguistic change*, 320–335. Oxford: Oxford University Press.

Roberts, Ian. 2017. *The wonders of language or how to make noises and influence people*. Cambridge: Cambridge University Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Shee-

han (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.
Roberts, Ian & Jeffrey Watumull. 2015. Leibnizian linguistics. In Ángel J. Gallego & Dennis Ott (eds.), *50 years later: Reflections on Chomsky's* Aspects, 211–222. Cambridge, MA: MIT Press.
Shannon, C.E. 1948. A mathematical theory of communication. *Bell System Technical Journal* 27(3). 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x.
Sheehan, Michelle, Theresa Biberauer, Ian Roberts & Anders Holmberg. 2017. *The Final-over-Final Condition: A syntactic universal*. Cambridge, MA: MIT Press.
Tattersall, Ian. 2012. *Masters of the planet: The search for our human origins*. New York: Palgrave Macmillan.
Turing, Alan M. 1937. On computable numbers, with an application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society* s2-42(1). 230–265. DOI: 10.1112/plms/s2-42.1.230.
Turing, Alan M. 1950. Computing machinery and intelligence. In Jack B. Copeland (ed.), *The essential Turing*, 433–464. Oxford: Clarendon Press.
Turing, Alan M. 1948. Intelligent machinery. In Jack B. Copeland (ed.), *The essential Turing*, 395–432. Oxford: Clarendon Press.
Ulam, Stanislaw. 1976. *Adventures of a mathematician*. New York: Scribner's.
Watumull, Jeffrey. 2015. *The linguistic Turing Machine*. University of Cambridge. (Doctoral dissertation).
Watumull, Jeffrey, Noam Chomsky & Ian Roberts. 2017. Economy in grammar learning. Paper presented at the workshop *Simplicity in Grammar Learning*, 23 September 2017.
Watumull, Jeffrey, Ian G. Roberts & Noam Chomsky. In press. Universal Grammar. In Douglas Vakoch (ed.), *Xenolinguistics: Toward a science of extraterrestrial language*. Oxford: Oxford University Press.
Wheeler, John Archibald. 1986. How come the quantum? *Annals of the New York Academy of Sciences* 480(1). 304–316. DOI: 10.1111/j.1749-6632.1986.tb12434.x.
Wittgenstein, Ludwig. 1922. *Tractatus logico-philosophicus*. London: Routledge.
Wolfram, Stephen. 2002. *A new kind of science*. Champagne, IL: Wolfram Media.

Chapter 2

Reconciling linguistic theories on comparative variation with an evolutionarily plausible language faculty

Kleanthes K. Grohmann
University of Cyprus, Cyprus Acquisition Team

Evelina Leivada
UiT The Arctic University of Norway, Cyprus Acquisition Team

This work aims to reconcile the atomic objects of study typically assumed within comparative variation studies with an evolutionarily plausible faculty of language. In the process, we formulate and address the *incompatibility problem*, the observation that studying comparative (micro)variation has progressively led to an evolutionarily implausible Universal Grammar. We identify a solution to this problem through arguing in favour of a so-called emergentist approach to some linguistic primitives. We then address the *granularity mismatch problem* and argue on the basis of this emergentist approach firstly, that linguistic and neurocognitive studies of language may be brought to the same level of granularity, and secondly, that specific insights from comparative variation can inform an evolutionarily plausible approach to human language.

1 Introduction

The topic of language variation and how it informs our study of the faculty of language (FL) together with its initial state are currently at the forefront of linguistic research (for latest overviews, see e.g. Hinzen 2014; Trettenbrein 2015;

Kleanthes K. Grohmann & Evelina Leivada. 2020. Reconciling linguistic theories on comparative variation with an evolutionarily plausible language faculty. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 25–42. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280629

Berwick & Chomsky 2016). As a matter of fact, the exploration of variation from a comparative, cross-linguistic perspective can be considered one of the very few topics which both linguists and cognitive neuroscientists agree merits further attention.

A representative perspective of the first area of research is that of generative linguist Noam Chomsky. When asked in a recent interview what the main advantages and/or reasons to study linguistic variation are, he reiterated a view that has been repeatedly explored in his work: In order to determine the capacity to use and understand language, we need to know "what options it permits" (Chomsky 2015). Put differently, if we want to understand FL and its initial state, Universal Grammar (UG), we must determine what structures UG is capable of generating. In the same vein, we should also determine what structures UG is *not* capable of generating as striking typological gaps across phylogenetically diverse languages call for explanations that can enrich our theory of language (see Biberauer, Holmberg & Roberts 2014 for a concrete example). From a linguistic perspective, we will call this the "insider" view.

To pursue the analogy, the perspective of cognitive neuroscientist Peter Hagoort can be described as the "outsider" view. Hagoort devoted part of his plenary talk at the 47th annual meeting of the Linguistic Society of Europe to how linguistics, once seen as a key player in the field of cognitive science, has seen its influence fade over the years (Hagoort 2014). This alienation directly relates to how linguists have presented their discoveries in the study of language variation. Often linguists have captured aspects of comparative variation through postulating primitives that they did not grow or derive in any sense, typically by assuming that a UG-encoded feature drives the relevant linguistic representation. Such postulations cannot be informative in the long run. Perhaps they can be successfully employed when one deals with some language A or B, but when the aim is broader (e.g., to approach our language-readiness and UG as its initial state), then such postulations are rather impeding progress.

In this context, the two most important questions to be addressed are (i) why this alienation across disciplines is happening and (ii) whether there is a remedy for this situation. The second question is the topic of §2. With respect to the first question, it seems that the reason is in part the way the topic of language variation has been approached over the last few years. More specifically, discussing comparative syntax and the way parametric models capture variation (see, for example, the recent collection of papers in Fábregas et al. 2015), Biberauer, Holmberg, Roberts & Sheehan (2014) argue that linguistic descriptions that have emerged since Chomsky (1981) have achieved an increasingly high level of

descriptive adequacy, but sacrificed explanatory adequacy due to the postulation of more and more entities in UG. In their words:

> Arguably, the direction that [principles & parameters] (P&P) theory has taken reflects the familiar tension between the exigencies of empirical description, which lead us to postulate ever more entities, and the need for explanation, which requires us to eliminate as many entities as possible. In other words, parametric descriptions as they have emerged in much recent work tend to sacrifice the explanatory power of parameters of Universal Grammar in order to achieve a high level of descriptive adequacy. (Biberauer, Holmberg, Roberts & Sheehan 2014: 104)

Describing linguistic data and formulating observations or generalisations over these data may then offer observational adequacy, possibly even descriptive adequacy, but not explanatory adequacy.

Although Biberauer, Holmberg, Roberts & Sheehan's point is well-taken, it is only a part of the issue at hand. Another part is presented by Yang (2004) when he writes that

> adult speakers, at the terminal state of language acquisition, *may retain multiple grammars, or more precisely, alternate parameter values*; these facts are fundamentally incompatible with the triggering model of acquisition […] *It is often suggested that the individual variation is incompatible with the Chomskyan generative program.* (Yang 2004: 50–51)

We can thus phrase the full problem as follows:

(1) *The incompatibility problem*: Studying microvariation has led to a model entailing an evolutionarily implausible UG/FL.

Put differently, we have managed to describe many linguistic structures across different languages, but now we have trouble explaining the ontology of the biological "structure" underlying their existence: UG. Given the short time scale typically assumed for evolution, the higher the degree of linguistic specificity encoded in UG, the more difficult the task of accounting for it in evolutionary terms.

Reconciling a bottom-up approach to UG and a resulting evolutionarily plausible FL with the findings from the literature on language variation has the potential to solve not only the incompatibility problem but also *Poeppel's problem*. More specifically, this reconciliation can overcome the granularity mismatch

considerations according to which linguistic and neuroscientific studies of language operate with objects of different granularity in a way that makes the construction of interdisciplinary bridges particularly difficult (cf. the granularity mismatch problem in Poeppel & Embick 2005). A bottom-up approach to UG entails a non-overarticulated UG which consists of a few computational principles (as Di Sciullo et al. 2010 have argued) only, leaving outside of this component many of the linguistic primitives that have been ascribed to it within comparative variation studies.

In this context, the next section discusses the importance of studying variation from a comparative, cross-linguistic perspective while at the same time maintaining a bottom-up approach to UG (i.e. an approach to UG from below that seeks to ascribe to it as little as possible, while maximizing the role of the other two factors in language design; Chomsky 2007). Pursuing a bottom-up vs. a top-down approach matters because depending on how much one ascribes to UG, the plausibility of the latter from an evolutionary perspective changes significantly. Our main aim is to offer the following solution to the incompatibility problem: An emergentist approach to some UG primitives can reconcile the Chomskyan generative program and the individual variation attested in reality. §3 then aims to offer a concrete demonstration of how relevant findings and primitives from the field of language variation can inform a biological approach to human language. §4 concludes and presents some suggestions for future work on this topic.

2 An emergentist approach to UG primitives

The second question that arose in the context of Hagoort's view on the interaction of linguistics with the larger field of cognitive science is whether there is a remedy for the observed decreased influence of linguistics. Hagoort (2014) offers five different directions for rectifying this issue. We apply some of these directions through pursuing an approach to UG primitives from below (Chomsky 2007), while at the same time retaining in our theory of FL some of the theoretical notions that pertain to the comparative variation literature. This combination has the potential of killing two birds with one stone, solving not only the incompatibility problem but also doing justice to the patterns of (micro)variation that are attested across languages in the following, two-step way:

I. Disentangling variation by teasing apart the different contributing factors which are responsible for deriving it in a way that does justice to sociolinguistic and psycho-/neurolinguistic aspects of language use, such as mono- vs. bilingual acquisition trajectories, the sociolinguistic status of the linguistic input, and the non-linguistic part of the environment.

II. Keeping UG primitives to a minimum in order to effectively comply with both minimalist principles and evolutionary constraints.

Point (I) has a second part that will not be addressed in this paper but that should be kept in mind nevertheless if the goal is to construct interdisciplinary bridges and overcome the granularity mismatch problem: Embedding the theory of language variation that emerges from step (I) into a "shared context of justification" (Hagoort 2014) by obtaining reliable data from different language groups, each of which may contribute its own characteristics towards deriving variation.[1] In practice, this would mean that careful elicitation of data should be followed by an attempt to interpret the data through *deriving* their properties rather than assuming that they are driven by a new, ad hoc postulated feature. If the aim is to understand FL rather than describe structure A in language B, then this process of interpretation should also be cautious to not rely on assumptions that are hard to sustain in the long run and quickly decompose under the light of interdisciplinary examination.

Talking about different contributing factors in (I) boils down to realising that variation across developmental paths of individuals that speak the same language can be the outcome of different modalities, environmental factors, non-linguistic features that affect linguistic development, and so on. For instance, research has shown that non-standard varieties allow for greater grammatical fluidity in a way that blurs the boundaries across different varieties. This, in turn, affects speakers' perceptions of whether a specific variant belongs to their linguistic repertoire or not (Cheshire & Stein 1997; Henry 2005). Another contributing factor is the trajectory of language acquisition and subsequent development, and the circumstances in which it takes place. For example, non-heritage speakers of a language may differ from heritage speakers of the *same* language with respect to the amount of variation attested in their repertoire (Montrul 2002; 2008; Lohndal & Westergaard 2016). The sociolinguistic status of the language(s) one is exposed to (the mono- vs. bilingual trajectory is in and of itself another factor that leads to variation) is yet another potential source of variation: In the case of non-standard varieties, speakers' perceptions about their native grammatical

[1] Hagoort (2014) argues that running sentences in one's head and consulting a colleague is fine for discovering interesting phenomena and possible explanations (the "context of discovery"), but it does not suffice as "the context of justification", due to innate confirmation biases and the fallibility of introspection. Thus, "to justify one's theory, empirical data have to be acquired and analysed according to the quantitative standards of the other fields of cognitive science". In the context of addressing the incompatibility problem, Hagoort's perspective is relevant because it shows how findings that may target points of grammatical (micro)variation should be analysed and interpreted.

variants are likely to be affected by their knowledge that many of their dialectal structures are considered unacceptable or "incorrect" by speakers of the standard variety (Henry 2005 for Belfast English; Leivada, Papadopoulou, Kambanaros, et al. 2017 for Cypriot Greek) in a way that enhances grammatical fluidity. Also, in those cases in which a standard variety co-exists with a structurally proximal, non-standard variety, the discreteness across grammatical variants at times fades away by the emergence of intermediate (Cornips 2006) or "diaglossic" speech repertoires (Auer 2005), resulting once more in a greater degree of variation (see also Rowe & Grohmann 2014 and relevant references cited for Cypriot Greek).

Understanding the multitude of faces that variation can acquire (for a more extensive overview, see Leivada 2015a) is of key importance when it comes to approaching UG primitives from an emergentist perspective. The reason is that cross-linguistic variation has long been described as part of UG, that is, deriving from UG parameters. Showing that patterns of variation are not as stabilised or uniform as the traditional UG parameters-account predicts opens the way for an emergentist approach to linguistic primitives that were traditionally viewed as part of UG. Understanding what terms like "stabilised" or "uniform" refer to in the present context requires shifting our attention to how variation *within* linguistic communities has been approached.

A crucial challenge for any approach to variation derives from the mainstream conception of the notion of "surface variation" (i.e. grammatical variation among speakers of the same language that is not the result of any acquired or developmental pathology) *within* a linguistic community. For example, Chomsky's idealised picture of a "completely homogeneous speech community" and an "ideal speaker-listener [...] who knows its language perfectly" (Chomsky 1965: 3) is often assumed together with the assumption that the so-called "linguistic genotype" is uniform across the species in the absence of severe and specific pathology (Anderson & Lightfoot 2000). Another related idea is that attained adult performance is "essentially homogeneous with that of the surrounding community", unless again a pathology is present (Anderson & Lightfoot 2000: 698). When translated into empirical terms, idealisations like these, although theoretically well-argued in their original context, paint a picture directly related to both Hagoort's and Poeppel's considerations. More specifically, by not doing justice to the patterns of surface variation that are attested in reality, theoretical linguistics may *lose* a significant part of its potential for interactions with fields that deal with recent sign language emergence, evolutionary linguistics, or sociolinguistics. Despite what the idealised picture suggests, variation can be found even in the absence of any pathology, even among speakers of the same language, and even within a native speaker who has passed the L1 acquisition period. The core

of this idea can be analysed across two dimensions, the linguistic dimension and the developmental one.

The developmental dimension refers to the fact that the presence of a severe and specific pathology is not a necessary condition for obtaining variation, even among neurotypical speakers of the same language. Individuals that share a diagnosis of cognitive disorder (or the absence of one) are not necessarily uniform in terms of their innate endowment: Individuals with a pathogenic variant of a gene can be impaired in a non-uniform fashion (variable expressivity), which may result in different cognitive phenotypes at times not reaching a cut-off point where the diagnosis of a specific pathology is possible. To demonstrate this with two examples, Fowler (1995) observes that there is tremendous variability with regard to language function in individuals with Down syndrome (variable expressivity). And it has also been observed that the existence of subsyndromal schizotypal traits in the general population is higher than average in first-degree relatives of patients with schizophrenia (Calkins et al. 2004). This led to the realisation that

> schizophrenia is not, despite its clinically important and reliable categorical diagnosis […], a binary phenotype (present, absent) with sudden disease onset. (Ettinger et al. 2014: 1)

In other words, some pathological characteristics might be present even if the cut-off point for reaching a diagnosis is not met – and, on the other hand, a diagnosis of schizophrenia might be reached, even if the pathological characteristics manifested among individuals with the same diagnosis are far from uniform. Together, these two examples suggest that it is equally plausible to expect that attained adult performance is not uniform among members of the same linguistic community in the absence of a pathology or in the presence of the same pathology.

With respect to the linguistic dimension, this is where factors related to non-standard varieties and inherent grammatical fluidity enter the picture. Evidently, not all linguistic communities are homogeneous, and in many cases this variation goes well beyond bi- or multilingualism. Similarly, in the case of recent language emergence de novo, as in the case of Al-Sayyid Bedouin Sign Language (ABSL) and other sign languages, fieldwork has shown that not only is the development of grammatical markers subject to environmental factors (e.g., time, distribution of speakers/signers, etc.), but also that great grammatical fluidity is attested at the various stages in the development of a language. In these recently emerged languages, points of variation ("parameters" in generative terms) are *not* fixed in

terms of their values, resulting in the realisation of alternate settings both within and across speakers (Washabaugh 1986; Sandler et al. 2011).

To mention a concrete example, consider the head-directionality parameter. S(ubject) O(bject) V(erb) is the prevalent word order among ABSL signers; this was, however, established as the prevalent order from the second generation of signers onwards only (Sandler et al. 2005), meaning that for some time the manifestations of this "parameter" were more fluid than what a stabilised parameter value would permit. Even more important is the fact that variation exists past the "stabilisation" point: Sandler et al. (2005: 2663) report the existence of some (S)VO patterns. As Leivada (2015a) argues in her discussion of ABSL, the fact that SOV patterns became robust in the second generation of speakers illustrates that variation is present when certain grammatical properties are still emerging. Fluctuating parameter values within a syntactic environment are incompatible with the idea that a parameter value is fixed past the terminal state of acquisition. Observing that this fluctuation exists in various cases, be it non-standard varieties or recently emerged grammars, is an indication that the head-directionality parameter "should indeed be better viewed as a surfacey decision that allows for varying realizations, rather than a fixed, deeply rooted syntactic parameter" (Leivada 2015a: 48). This does not mean that points of variation are unfixed and eventually culminate in an "anything goes" grammar, but it does mean that this surface decision is not (i) syntactic (i.e. Chomsky in recent work has explicitly recognized that variation between grammars is a matter of variable externalization; see Berwick & Chomsky 2011: 41), (ii) UG-encoded, or (iii) binary, as the classical parametric approach would suggest. Non-binarity is particularly evident in case of bidialectal speakers; their linguistic repertoire may include functionally equivalent variants (Kroch 1994) with *different* values that are alternatively realized in the *same* syntactic environment (Leivada, Papadopoulou & Pavlou 2017).

An emergentist approach to some linguistic primitives that were previously thought to be parts of UG will be able to reconcile the Chomskyan generative program (and especially UG, as one of its main pillars) with the patterns of variation that are attested in reality (see Yang's 2004 point mentioned earlier). Moreover, an emergentist approach will solve the incompatibility problem, as the number of linguistic primitives allocated to UG will be reduced. The notion of *emergent parameters* (Roberts & Holmberg 2010; Roberts 2012; Biberauer, Roberts & Sheehan 2014; Biberauer & Roberts 2017) is an important step in this direction. The central idea behind emergent parameters is that instead of postulating a richly specified parametric endowment as part of the initial state of our FL (UG; Chomsky 1981), parameters are derived (i.e. emergent) properties falling out of the interaction of

Chomsky's (2005) three factors in language design (Biberauer, Holmberg, Roberts & Sheehan 2014). In the context of emergent parameters in which UG does not provide a pre-specified "menu" of parametric choices, Biberauer, Roberts & Sheehan (2014) note that it is very important to provide independent motivation for the plausibility of the parameters that acquirers will postulate as well as for the sequence in which each point of variation should be considered. Here lies the solution to the incompatibility problem and a first step towards approaching the granularity mismatch problem.

With respect to the incompatibility problem, if the points of variation that are meaningful from a comparative (micro)variation perspective are treated as emergent properties, they are no longer translated as innately specified options. The consequence of this move is that UG would be considerably deflated and much easier to discuss from an evolutionary perspective. As Chomsky (2007) has very convincingly argued, for any given component or structure, the less attributed to structure-specific factors for determining the development of an organism, the more feasible the study of its evolution, hence the need for a bottom-up approach to UG.

In relation to the granularity mismatch problem, the important component of the "emergent parameters"-account lies in the element of *interaction*. As Biberauer, Roberts & Sheehan (2014) explicitly claim, it is the interaction of the second factor (linguistic input) and the third factor (non-language-specific principles of cognition) plus the language-readiness (provided by the first factor, UG); that delivers emergent parameters. To illustrate this with an example, let's return to the head-directionality parameter, which makes reference to the position of a head in relation to its dependents. Traditional accounts of grammar would describe Japanese as a head-final and English as a head-initial language, with the difference between the two explained in terms of the different value to which the head-directionality parameter is set. The typological preference given to harmonic orders (i.e. *consistent* head-initial or head-final patterns within a language; see Hawkins 2010) might also be taken to suggest that a UG-based head-directionality parameter is indeed operative and, once set, its effects are diffused across different syntactic environments.[2] Alternatively, one could argue that the realisation of the head in relation to its dependents does not boil down to setting a UG-based parameter. This latter approach should be preferred because it is compatible with the fact that variation *can* be attested past the "setting" state in the repertoire of a neurotypical, adult speaker who has fully acquired her language

[2] A reviewer points out that this is not assumed within the emergentist approach just outlined. Indeed, it is not and we do not embrace this explanation either; we only point out that it is an alternative explanation, which, however, should not be preferred, since it does not accommodate the patterns of variation that are attested.

(as suggested in the case of ABSL). If one chooses to approach this parameter as an emergent parameter, the interaction of this grammatical choice with principles of general cognitive architecture becomes meaningful. For example, why are harmonic orders preferred if they are not *imposed* by the setting of a predetermined parameter? Of course, an emergent parameter would also need to be "set" in order to reflect the options that are permitted in the adult grammar, but crucially by not being encoded in UG, its variable realizations within and across speakers of the same language (e.g., in the form of functionally equivalent variants; Leivada, Papadopoulou & Pavlou 2017) would not be a problem for our theory of UG and/or FL.

Roberts (2016b) suggests that these generalisation effects are related to the computational conservatism of the learning device. This is formally captured by his *input generalisation*: "There is a preference for a given feature of a functional head F to generalise to other functional heads G, H …" (cf. Roberts 2007: 275) – that is, to "maximise available features" (Biberauer & Roberts 2016; Roberts 2016b). This computational conservativism is a third factor principle. If so, preference for harmonic orders no longer amounts to a UG-wired principle or parameter, but to the way human memory or even learning more broadly works. It has been shown that sequence edges are particularly salient positions and facilitate learning in a way that gives rise to *either* word-initial *or* word-final processes much more often than otherwise (see, for example, Endress et al. 2009 on the prevalence of prefixing and suffixing across languages in comparison to the rarity of infixing). At the syntactic level, Dryer (1992) observes the following correlation with respect to generalisation effects in relation to the position of the Head on the basis of 434 languages: OV languages are mostly postpositional and VO languages are mostly prepositional. From Dryer's dataset, Hawkins (2010) calculated that the vast majority of languages (93%) are consistently OV-postpositional or VO-prepositional. Hawkins (2010) approaches harmonic word-orders in terms of third factor demands, and, more specifically, a processing preference that favours shorter processing domains. Evidently, the workings of comparative (micro)variation which deal with headedness patterns across typologically different languages can now be revisited and explained from a different perspective. This perspective involves the *interaction* of linguistic patterns with the driving forces of general cognition in a way that addresses Hagoort's considerations. With respect to the "messy" patterns of variation that just do not fit in the classical notion of a binary parameter, but that are just as uncontroversially there, an emergentist approach has the potential to cover these too. If parameters are emergent and allow for non-binary realizations, then the incompatibility that Yang (2004) correctly observes between these "messy" patterns and UG disappears.

Despite its theoretical and empirical benefits, this interaction may not solve the *granularity mismatch problem*. It may contribute to the construction of interdisciplinary bridges in some respects, but still a good portion of primitives may be left unmapped across disciplines. Put differently, even if parameters or other linguistic primitives are explained through an emergentist approach, this would not entail that the granularity mismatch problem has been solved. This could be due to the complicated nature of the task at hand; as Hornstein (2009: 156–157) argues, "the right theory of grammar will be one that has (roughly) the empirical coverage of [government-and-binding theory], *and* that 'solves' Plato's problem, Darwin's problem, *and* the granularity mismatch problem" (emphasis added).[3] In other words, given how polylithic both the problem and its solutions are, there can be no a priori guarantee of success. Despite recognising this possibility, the next section will follow Hagoort's (2014) suggestion to maximise the interdisciplinary contributions of linguistics within a larger cognitive (neuro)science environment. We endeavour to approach a constraint, which in the linguistics literature has been called "linguistic" or "syntactic" more often than not, in neurocognitive terms.

3 Levels of granularity: Anti-identity as a case study

Anti-identity has received many distinct names in the linguistics literature; consider, for example, the *obligatory contour principle* in phonology (Odden 1986), *identity avoidance* (van Riemsdijk 2008), *distinctness* (Richards 2010), *X-within-X recursion* (Arsenijević & Hinzen 2012). This is also the basis for *anti-locality* relations in syntax (Grohmann 2003, recently surveyed with additional references in Grohmann 2011). Regardless of the level of linguistic analysis at stake, anti-identity in general describes the absence of adjacent elements of the same category (e.g., [*XX] in syntax).

There are different ways to approach this phenomenon. In the linguistics literature, it has been approached in terms of a UG-imposed well-formedness ban that precludes the adjacency of same-category elements (see Richards 2010 for a more detailed discussion). This position would place the ban in UG, together with the configurations of categorial features that the ban is sensitive to. Alternatively, one could aim to keep UG at a minimum and see whether [*XX] can be shown to boil down to a general, cognitive principle. A first step in this direction

[3] According to Hornstein (2009), Darwin's problem refers to "the logical problem of language evolution", how language emerged in the species (see also Boeckx & Grohmann 2007 on the relation between Plato's problem and Darwin's problem).

is made by van Riemsdijk (2008) when he briefly argues that identity avoidance might be "a general principle of biological organization" (p. 242). If so, one expects to find its manifestations not only in language, but also in other domains of cognition.

Taking one step back, if this comparison across cognitive domains is fruitful, one would have successfully mapped an element that appears in the "parts list" (i.e. a list that enumerates concepts canonically used in the fields of study it represents; see Poeppel & Embick 2005) of two different disciplines. In more recent work, Poeppel (2012) talks about the *mapping problem*. In his words, the mapping problem "addresses the relation between the primitives of cognition (here speech, language) and neurobiology. Dealing with this mapping problem invites the development of linking hypotheses between the domains" (Poeppel 2012: 34). Developing these linking hypotheses is the only route to potentially solving the granularity mismatch problem. Returning now to the case at hand, linking hypotheses *can* be constructed for [*XX].

It seems to be true that humans do not like repetitions in general and that anti-identity in language is not the result of a linguistic ban but of a bias that finds application in other domains of human cognition too. Walter's (2007) biomechanical repetition avoidance hypothesis proposes a *physiological* motivation for this dislike: Repetition of articulatory gestures is relatively difficult, and this difficulty results in phonetic variation; that is, in [XX] it is likely that the two elements are not spelled out identically. We propose the term "novel information bias", which has a *cognitive* motivation: It refers to the well-demonstrated fact that subjects are unable to tokenise multiple adjacent instances of the same type (Treisman & Kanwisher 1998, Walter 2007) because of a general bias in the perceptual system to be more attentive to novel sensory information than to repeated information (Leivada 2017).

In the body of research by Kanwisher (1987 et seq.), *repetition blindness* has been described as the result of difficulties in detecting repeated tokens in rapid serial visual presentations of words. Another illustration is the *apparent motion illusion*: Identical stimuli flashed in different locations are largely perceived as a single moving stimulus; in other words, subjects show a clear preference for a representation of different tokens as one moving token (Vetter et al. 2012). What this means in the context of [*XX] is that talking about a general cognitive bias on anti-identity instead of a UG-wired linguistic constraint that bans [*XX] explains why a limited number of [XX] patterns do surface cross-linguistically (as shown in Leivada 2015b). In sum, the strong preference for anti-identity in language has to do with the way our brain computes types and tokens, and not with a syntactic ban on same-category embedding.

Overall, this approach to anti-identity can be extended to other UG primitives such as parameters or categorial features. In line with Poeppel & Embick's (2005) suggestion to "tak[e] linguistic categories seriously and us[e] them to investigate how the brain computes with such abstract categorical representations" (p. 107), this approach can lead to an evolutionarily plausible UG, while at the same time describing and accounting for the patterns of variation that one has to deal with in the field of comparative variation.

4 Outlook

The approach to UG primitives advocated in this work is still in its earliest stages. An important thing to keep in mind for future work is that deflating UG does not equal arguing against its existence. In other words, there can be a noticeable change in the way we treat UG primitives, without denying the existence of UG (for further discussion, see Roberts 2016a and many of the contributions to that volume). The second important note is that achieving the right levels of abstraction and representation in this effort is crucial: The more linguists abstain from postulating UG-encoded primitives that are very language-specific in nature, the more progress will be made in embedding findings from linguistics in a productively shared context of justification. Last, a third part of this type of approach that is worth mentioning is the conclusion reached in Biberauer, Roberts & Sheehan (2014): What were previously thought to be hard-wired properties of FL could actually reduce to emergent properties that feature the element of interaction among the different factors in language design.

Abbreviations

ABSL	Al-Sayyid Bedouin Sign Language	FL	faculty of language
		UG	Universal Grammar

Acknowledgements

We thank two anonymous reviewers for their helpful comments. KKG's contribution was partially supported by Leventis project 3411-61041 (University of Cyprus). EL acknowledges support from European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement no. 746652.

References

Anderson, Stephen R. & David W. Lightfoot. 2000. The human language faculty as an organ. *Annual Review of Physiology* 62. 697–722. DOI: 10.1146/annurev.physiol.62.1.697.

Arsenijević, Boban & Wolfram Hinzen. 2012. On the absence of X-within-X recursion in human grammar. *Linguistic Inquiry* 43(3). 423–440. DOI: 10.1162/LING_a_00095.

Auer, Peter. 2005. Europe's sociolinguistic unity, or: A typology of European dialect/standard constellations. In Nicole Delbecque, Johan van der Auwera & Dirk Geeraerts (eds.), *Perspectives on variation: Sociolinguistic, historical, comparative*, 7–42. Berlin: De Gruyter.

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its evolution. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Berwick, Robert C. & Noam Chomsky. 2016. *Why only us: Language and evolution.* Cambridge, MA: MIT Press.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Biberauer, Theresa & Ian Roberts. 2016. Emergent parameters and syntactic complexity: New perspectives. Paper presented at 'Notions of 'feature' in linguistic theory, Cross-theoretical and cross-linguistic perspectives' workshop at the 49th Annual Meeting of the SLE 2016, Naples.

Biberauer, Theresa & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax*, 134–161. Cambridge: Cambridge University Press.

Biberauer, Theresa, Ian Roberts & Michelle Sheehan. 2014. No-choice parameters and the limits of syntactic variation. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 46–55. Somerville, MA: Cascadilla Proceedings Project.

Boeckx, Cedric & Kleanthes K. Grohmann. 2007. The *Biolinguistics* manifesto. *Biolinguistics* 1. 1–8.

Calkins, Monica E., Clayton E. Curtis, William M. Grove & William G. Iacono. 2004. Multiple dimensions of schizotypy in first degree biological relatives of schizophrenia patients. *Schizophrenia Bulletin* 30(2). 317–325. DOI: 10.1093/oxfordjournals.schbul.a007081.

Cheshire, Jenny & Dieter Stein. 1997. *Taming the vernacular: From dialect to written language*. Harlow: Longman.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax–semantics*, 1–30. Berlin: De Gruyter.

Chomsky, Noam. 2015. An interview on linguistic variation with Noam Chomsky. *Isogloss* 1. 143–145. DOI: 10.5565/rev/isogloss.11.

Cornips, Leonie. 2006. Intermediate syntactic variants in a dialect-standard speech repertoire and relative acceptability. In Gisbert Fanselow, Caroline Féry, Matthias Schlesewsky & Ralf Vogel (eds.), *Gradience in grammar: Generative perspectives*, 85–105. Oxford: Oxford University Press.

Di Sciullo, Anna Maria, Massimo Piattelli-Palmarini, Kenneth Wexler, Robert C. Berwick, Cedric Boeckx, Lyle Jenkins, Juan Uriagereka, Karin Stromswold, Lisa Lai-Shen Cheng, Heidi Harley, Andrew Wedel, James McGilvray, Elly van Gelderen & Thomas G. Bever. 2010. The biological nature of human language. *Biolinguistics* 4(1). 4–34.

Dryer, Matthew S. 1992. The Greenbergian word order correlations. *Language* 68(1). 81–138. DOI: 10.2307/416370.

Endress, Ansgar D., Marina Nespor & Jacques Mehler. 2009. Perceptual and memory constraints on language acquisition. *Trends in Cognitive Sciences* 13(8). 348–353. DOI: 10.1016/j.tics.2009.05.005.

Ettinger, Urlich, Inga Meyhöfer, Maria Steffens, Michael Wagner & Nikolaos Koutsouleris. 2014. Genetics, cognition, and neurobiology of schizotypal personality: A review of the overlap with schizophrenia. *Frontiers in Psychiatry* 5(18). DOI: 10.3389/fpsyt.2014.00018.

Fábregas, Antonio, Jaume Mateu & Michael T. Putnam (eds.). 2015. *Contemporary linguistic parameters*. London: Bloomsbury Academic.

Fowler, Anne E. 1995. Linguistic variability in persons with Down syndrome: Research and implications. In Lynn Nadel & Donna Rosenthal (eds.), *Down syndrome: Living and learning in the community*, 121–131. New York: John Wiley.

Grohmann, Kleanthes K. 2003. *Prolific domains: On the anti-locality of movement dependencies*. Amsterdam: John Benjamins. DOI: 10.1075/la.66.

Grohmann, Kleanthes K. 2011. Anti-locality: Too-close relations in grammar. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 260–290. Oxford: Oxford University Press.

Hagoort, Peter. 2014. Linguistics quo vadis? An outsider perspective. Paper presented at the 47th Annual Meeting of the Societas Linguistica Europaea, Adam Mickiewicz University, Poznań.

Hawkins, John A. 2010. The FOFC from a processing typology perspective. Paper presented at the Workshop on the Final-Over-Final Constraint: Processing, Typology and Acquisition, University of Cambridge.

Henry, Alison. 2005. Non-standard dialects and linguistic data. *Lingua* 115(11). 1599–1617. DOI: 10.1016/j.lingua.2004.07.006.

Hinzen, Wolfram. 2014. What is un-Cartesian linguistics? *Biolinguistics* 8. 226–257.

Hornstein, Norbert. 2009. *A theory of syntax: Minimal operations and Universal Grammar*. Cambridge: Cambridge University Press.

Kanwisher, Nancy G. 1987. Repetition blindness: Type recognition without token individuation. *Cognition* 27(2). 117–143. DOI: 10.1016/0010-0277(87)90016-3.

Kroch, Anthony. 1994. Morphosyntactic variation. In Katherine Beals (ed.), *Papers from the 30th regional meeting of the Chicago Linguistic Society: Parasession on variation and linguistic theory*, 180–201. Chicago: Chicago Linguistic Society.

Leivada, Evelina. 2015a. *The nature and limits of variation across languages and pathologies*. Universitat de Barcelona. (Doctoral dissertation).

Leivada, Evelina. 2015b. X-within-X structures and the nature of categories. *Biolinguistics* 9. 50–73.

Leivada, Evelina. 2017. What's in (a) label? Neural origins and behavioral manifestations of identity avoidance in language and cognition. *Biolinguistics* 11. 1–30.

Leivada, Evelina, Elena Papadopoulou, Maria Kambanaros & Kleanthes K. Grohmann. 2017. The influence of bilectalism and non-standardization on the perception of native grammatical variants. *Frontiers in Psychology*: Language Sciences 8, 205. DOI: 10.3389/fpsyg.2017.00205.

Leivada, Evelina, Elena Papadopoulou & Natalia Pavlou. 2017. Functionally equivalent variants in a non-standard variety and their implications for Universal Grammar: A spontaneous speech corpus. *Frontiers in Psychology*: Language Sciences 8, 1260. DOI: 10.3389/fpsyg.2017.01260.

Lohndal, Terje & Marit Westergaard. 2016. *Grammatical gender in american norwegian heritage language: Stability or attrition?* DOI: 10.3389/fpsyg.2016.00344.

Montrul, Silvina. 2002. Incomplete acquisition and attrition of Spanish tense/aspect distinctions in adult bilinguals. *Bilingualism: Language and Cognition* 5(1). 39–68. DOI: 10.1017/S1366728902000135.

Montrul, Silvina. 2008. *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins. DOI: 10.1075/sibil.39.

Odden, David. 1986. On the role of the Obligatory Contour Principle in phonological theory. *Language* 62(2). 353–383. DOI: 10.2307/414677.

Poeppel, David. 2012. The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology* 29(1-2). 34–55. DOI: 10.1080/02643294.2012.710600.

Poeppel, David & David Embick. 2005. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-first century psycholinguistics: Four cornerstones*, 103–118. Mahwah, NJ: Lawrence Erlbaum.

Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Roberts, Ian. 2012. Macroparameters and minimalism: A programme for comparative research. In Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo & Juanito Avelar (eds.), *Parameter theory and linguistic change*, 320–335. Oxford: Oxford University Press.

Roberts, Ian. 2016a. Introduction. In Ian Roberts (ed.), *The Oxford handbook of Universal Grammar*, 1–34. Oxford: Oxford University Press.

Roberts, Ian. 2016b. Linearization, labelling and word-order parameters. Presentation at Workshop on disharmonic word order, Chinese University of Hong Kong, October 14th.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Rowe, Charley & Kleanthes K. Grohmann. 2014. Canaries in a coal mine: Native speakerhood and other factors as predictors of moribundity, death, and diglossic shift in Cypriot Greek. *Mediterranean Language Review* 21. 121–142.

Sandler, Wendy, Irit Meir, Svetlana Dachkovsky, Carol A. Padden & Mark Aronoff. 2011. The emergence of complexity in prosody and syntax. *Lingua* 121(113). 2014–2033. DOI: 10.1016/j.lingua.2011.05.007.

Sandler, Wendy, Irit Meir, Carol A. Padden & Mark Aronoff. 2005. The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America* 102. 2661–2665. DOI: 10.1073/pnas.0405448102.

Treisman, Anne M. & Nancy G. Kanwisher. 1998. Perceiving visually-presented objects: Recognition, awareness, and modularity. *Current Opinion in Neurobiology* 8(2). 218–226. DOI: 10.1016/S0959-4388(98)80143-8.

Trettenbrein, Patrick C. 2015. The "grammar" in Universal Grammar: A biolinguistic clarification. *Questions and Answers in Linguistics* 2(1). 1–10. DOI: 10.1515/qal-2015-0005.

van Riemsdijk, Henk C. 2008. Identity avoidance: OCP effects in Swiss relatives. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 227–250. Cambridge, MA: MIT Press.

Vetter, Peter, Grace Edwards & Lars Muckli. 2012. *Transfer of predictive signals across saccades*. DOI: 10.3389/fpsyg.2012.00176.

Walter, Mary Ann. 2007. *Repetition avoidance in human language*. Cambridge, MA: MIT. (Doctoral dissertation).

Washabaugh, William. 1986. *Five fingers for survival: Deaf sign language in the Caribbean*. Ann Arbor, MI: Karoma Press.

Yang, Charles D. 2004. Toward a theory of language growth. In Lyle Jenkins (ed.), *Variation and universals in biolinguistics*, 37–56. Amsterdam: Elsevier.

Chapter 3

Rethinking remerge: Merge, movement and music

Hedde Zeijlstra
Georg-August-Universität Göttingen

In an influential paper, Katz & Pesetsky (2011) present the identity thesis for language and music, stating that "[a]ll formal differences between language and music are a consequence of differences in their fundamental building blocks (arbitrary pairings of sound and meaning in the case of language; pitch classes and pitch-class combinations in the case of music). In all other respects, language and music are identical." Katz & Pesetsky argue that, just like syntactic structures, musical structures are generated by (binary) Merge, for which they provide a number of arguments: for instance, musical structures are endocentric (each instance of Merge in music, just like in language, has a labelling head). They also argue that movement phenomena (i.e., the application of Internal Merge) can be attested in both language and music. While fully endorsing the view that musical structures are the result of multiple applications of External (binary) Merge, this paper argues that the arguments in favour of the presence of Internal Merge in music are at best inconclusive and arguably incorrect. This is, however, not taken as an argument against the identity thesis for language and music; rather, I take it to follow from it: the identity thesis for language and music reduces all differences between language and music to its basic building blocks. If the application of Internal Merge in natural language is driven by uninterpretable features (cf. Chomsky 1995; 2001; Bošković 2007; Zeijlstra 2012) that are language-specific and not applicable to music (the reason being that only building blocks that are pairings of sound and meaning can be made up of interpretable and uninterpretable features), the direct consequence is that Internal Merge cannot be triggered in music either.

Hedde Zeijlstra. 2020. Rethinking remerge: Merge, movement and music. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 43–66. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280631

1 Introduction: External and Internal Merge in language and music

Since Chomsky (1995), the operation Merge has been taken to be the primary structure-building operation in natural language. In current minimalism, syntactic movement is, moreover, considered a special instance of Merge (Internal Merge), which applies to a particular syntactic object and a part thereof (cf., inter alia, Chomsky 2005). In this sense, Internal Merge is different from External Merge, where the two input objects do not stand in an inclusion relation.

However, natural language is not the only cognitive domain where Merge is said to be a structure-building operation. As has been claimed in Lerdahl & Jackendoff (1983) and, more recently, in Katz & Pesetsky (2011), music is also a cognitive domain where structures can be taken to be generated by means of an operation like Merge. If musical structures are indeed generated by means of Merge and if movement is a special instance of Merge, the question arises whether music exhibits movement effects as well. After all, why could Internal Merge not apply in music if it can apply in natural language?

In order to account for the differences and similarities between language and music, Katz & Pesetsky (2011) entertain their so-called *identity thesis for language and music*, which states that:

> [a]ll formal differences between language and music are a consequence of differences in their fundamental building blocks (arbitrary pairings of sound and meaning in the case of language; pitch-classes and pitch class combinations in the case of music). In all other respects, language and music are identical. (Katz & Pesetsky 2011: 3)

For Katz & Pesetsky, this means that Merge should be equally effective in natural language and music and that therefore music is indeed expected to exhibit both External and Internal Merge effects. In their paper, they identify particular musical patterns that they take to reflect movement in music.

However, one may wonder whether it is correct to assume that identity thesis for language and music entails that both External and Internal Merge should apply in music. As I will argue in this paper, it all depends on what triggers Internal Merge in the first place. Internal Merge differs from External Merge in the sense that Internal Merge does not have to take elements from the numeration into the syntactic structure. If every element in the numeration needs to end up in the syntactic structure, it follows immediately that every element present in the numeration needs to undergo External Merge. But why would particular elements be required to undergo Internal Merge as well?

Following a longstanding tradition in syntactic theory, I assume that Internal Merge is triggered by so-called uninterpretable formal features – formal features that need to stand in a particular configuration with their interpretable counterparts. If that is the case, the question arises as to whether such movement-triggering features can also be attested in music. I argue they do not.

According to the identity thesis for language and music, all differences between music and language should reduce to differences in their building blocks: for Katz & Pesetsky, arbitrary pairings of sound and meaning in the case of language, and pitch classes and pitch-class combinations in the case of music. Let's focus in more detail on each type of building blocks.

Lexical items are generally thought to consist of three types of features: phonological features, syntactic or formal features, and semantic features. Phonological features are only interpretable or legible for the sensori-motor system; semantic features are only interpretable or legible for the conceptual-intentional systems; and syntactic or formal features are interpretable or legible for neither of them. In that sense, linguistic building blocks can be said to be multi-modular, not mono-modular.

Things are different when it comes to musical building blocks. One dimension in which the architecture of music is much different from that of natural language is that musical structures are not subject to compositional semantic interpretation in the sense that the meaning of a musical structure – to the extent it has any (see, for instance, Schlenker 2016 and references therein for discussion) – follows compositionally from the meaning of the parts it consists of and the way these parts are structured. While linguistic objects are built of elements that form sound-meaning pairs, the musical objects are not. Musical building blocks are mono-modular building blocks. Mono-modular building blocks are building blocks that are all interpretable or legible for the same module, in this case the sound side of music. And even if it turns out that pitch classes and pitch-class combinations are not the only available building blocks in music (and other building blocks are available as well, either inside or outside Western tonal music), those building blocks will still belong to the same sound module.

Mono- vs. multi-modularity is then a main characteristic of the differences between musical and linguistic building blocks. Now, under the view that the application of Internal Merge is indeed driven by the need of so-called uninterpretable features to be checked by their interpretable counterparts, it follows immediately that Internal Merge can only be triggered by features present on linguistic building blocks, not on musical building blocks. The reason is that uninterpretable features are defined as elements that are not part of the set of semantic features, but require a particular checking (or valuation) relation with a feature that does

belong to this set. As a consequence, no uninterpretable feature can be acquired without the presence of a semantic counterpart (see Brody 1997; Svenonius 2007; Zeijlstra 2008; 2012). But if that is correct, uninterpretable features, by definition, can only be part of building blocks that are not mono-modular. In fact, in any cognitive system whose output is not defined in terms of pairs of elements belonging to different cognitive modules (in the way that linguistic output is defined in terms of sound-meaning pairs), features that denote dependencies on elements belonging to different modules cannot exist.

If that is the case, the identity thesis for language and music should actually predict that, to the extent that Internal Merge can only be triggered by uninterpretable formal features, it can never apply to pieces of musical structure and that therefore instances of movement are expected to be absent in music.

In this article, I first further elaborate the claim that (properties of) uninterpretable features are the trigger for syntactic movement (§2). Then, in §3, I discuss Katz & Pesetsky's claim that music does not only exhibit External Merge, but also Internal Merge. In §4, I spell out some problems for the claim that music exhibits movement effects, and I provide an alternative analysis for the phenomena discussed by Katz & Pesetsky that does not allude to movement. I argue that this alternative account can equally well, if not better, explain the special behaviour of full cadences than the movement account does. §5 concludes.

2 Internal and External Merge in natural language

One of the highlights of the twenty-first-century developments in minimalism has been the operational unification of syntactic structure building and movement. While previous versions of minimalism (and its generative predecessors) took movement to involve a separate syntactic operation alongside Merge (or any other structure-building operation), Chomsky (2005) argued that nothing a priori forbids Merge to apply to previously created parts of the syntactic structure, and to remerge, or internally merge, these with the top node of the derivation (see also Starke 2001). Under this conception of Internal Merge, the question as to why natural language would display displacement operations no longer seemed to be in need of an explanation. If Merge is not restricted to External Merge, it would rather require additional explanation if language did not display movement effects.

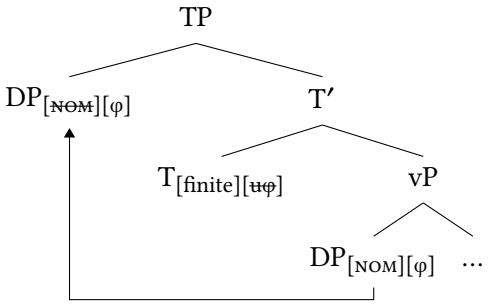
At the same time, questions still arise with respect to when Internal Merge should take place. Internal Merge differs from External Merge in the sense that Internal Merge does not have to take elements from the numeration into the

syntactic structure. If every element in the numeration needs to end up in the syntactic structure, it follows immediately that every element present in the numeration needs to undergo External Merge. But why would particular elements be required to undergo Internal Merge as well? From this perspective, there is no (external) reason that would force Internal Merge to take place.

The most straightforward solution would be to assume that Internal Merge only takes place if not applying it would render the sentence ungrammatical. Under that view, Internal Merge is a costly operation that only applies when necessary. This means that it is an operation for which a trigger is needed; and therefore, the question immediately arises as to what triggers Internal Merge.

Originally, it has been proposed by Chomsky (1995) that so-called uninterpretable features trigger movement. In a structure like (1), it is the uninterpretable [uφ] feature on T that triggers movement of the lower DP into the specifier position of the T-head, so that this feature, as well as the nominative feature on the DP, can be checked. The central conceptual motivation behind uninterpretable features as triggers for movement was that this would reduce two not well understood phenomena – the existence of semantically vacuous elements and the existence of displacement effects – to one not well understood notion: the need to remove uninterpretable features (where removal of uninterpretable features was said to take place under spec-head configuration).

(1)



This view, however, was later on rejected, primarily since it turned out that uninterpretable features could be checked at a distance (the uninterpretable feature probing down in its c-command domain to find a matching active goal). English expletive constructions (where the finite verb agrees with a lower VP-internal associated subject) (2), Icelandic quirky case constructions (where the verb agrees in number with a nominative object) (3), and various other constructions all underlie structures where the probe and the goal of agreement never appear in spec-head configuration:

(2) a. There seems to have arrived some student.
b. There seem to have arrived some students.

(3) Icelandic (Bobaljik 2008)
a. Jóni líkuđu thessir sokkar
Jon.DAT like.PL these socks.NOM
'Jon likes these socks.'
b. Mér virdast hestarnir vera seinir
me seem. PL the.horses be slow
'It seems to me that the horses are slow.'

If uninterpretable features can no longer be taken to trigger Internal Merge, the question arises as to what should do instead. Chomsky (2000; 2001) argues that movement should be thought of as an operation dependent on, and not triggered by, agreement. For him, probes, carrying uninterpretable features, could be equipped with an additional feature [EPP], which requires that the specifier of the probing head be filled. If no other suitable candidate could be merged externally in that position (such as an expletive subject like English *there*, or a dative subject, to the extent that such elements could be externally merged in this position in the first place; cf. Chomsky 2000; Deal 2009 for different proposals and discussion), the goal would raise into that position.

Even though using the EPP-feature gets these facts right, its postulation has often been criticized for a lack of independent motivation. The EPP-feature is rather a movement-triggering diacritic and does not build upon any explanation as to why movement should take place in the first place, although it could be that the presence or absence of movement (diacritics) is really just formal arbitrariness (a position taken by Biberauer et al. 2009; 2014; Biberauer & Roberts 2015, among others). For this reason, others have proposed to reinstall uninterpretable features themselves, rather than EPP-features, to be the sole triggers of movement (e.g., Bjorkman & Zeijlstra 2019). Nevertheless, whether uninterpretable features or subfeatures of uninterpretable features are the trigger for movement, in both cases uninterpretable features still form necessary elements in movement-triggering configurations.

Naturally, it is not the case that EPP-features and (un-)interpretable features are the only candidates for being movement triggers. Richards (2016), for instance, has argued that phonological adjacency requirements trigger movement; and Neeleman & Van de Koot (2008) have argued that movement may feed various mapping rules. But it should be noted that this type of approaches also relates the necessity of movement to interface requirements, as do uninterpretable

feature approaches. This all suggests that, in cognitive systems that lack formal features mediating between phonological and semantic features, triggering of Internal Merge might not be possible.

3 Internal and External Merge in music

In this section, I discuss the extent to which Merge can be said to be the (sole) structure-building operation in music, as claimed by Katz & Pesetsky. In order to provide evidence for this claim, Katz & Pesetsky build upon the insights presented in Lerdahl & Jackendoff's (1983) *Generative theory of tonal music* (GTTM). I will first briefly illustrate the major components of GTTM that are relevant for the discussion in this paper, without doing justice to the richness of this theoretical framework (§3.1). Then, in §3.2, I will present a particular aspect of music, namely the existence of structural hierarchies in music, which, for Katz & Pesetsky, forms evidence for their claim that musical structures are generated by at least External Merge. In §3.3, I discuss how, according to Katz & Pesetsky, other musical properties provide evidence for Internal Merge in music.

3.1 Lerdahl & Jackendoff's Generative theory of tonal music

According to the GTTM model, there are four components that determine the proper analysis of a musical structure. These four components are listed/given in (4) below:

(4) a. grouping structure
    b. metrical structure
    c. time-span reduction (TSR)
    d. prolongational reduction (PR)

Following Lerdahl & Jackendoff (1983: 8–9), grouping structure "expresses the hierarchical segmentation of the piece into motives, phrases, and sections"; metrical structure "expresses the intuition that the events of the piece are related to a regular alternation of strong and weak beats at a number of hierarchical levels"; TSR "assigns to the pitches of the piece a hierarchy of "structural importance" with respect to their position in grouping and metrical structure"; and PR, finally, "assigns to the pitches a hierarchy that expresses harmonic and melodic tension and relaxation, continuity and progression".

For Lerdahl & Jackendoff (1983), each component can assign a set of structures to a given string of music; and an additional set of preference interface rules then

determines which of these analyses is the correct one (often just one). In this sense, the musical architecture forms a strong resemblance with Jackendoff's parallel architecture of grammar (Jackendoff 1997; 2002; Culicover & Jackendoff 2005), which treats phonology, syntax, and semantics as independent generative components whose structures are also linked by interface rules: each component generates (a number of) structures, and interface rules determine what the proper mappings between these structures are. Such interface rules, for instance, determine which prosodic and which syntactic structures correlate.

Jackendoff's parallel architecture differs from Minimalist grammar in the sense that parallel architecture grammar has multiple engines, whereas Minimalist grammar has only one engine: its output leading to different levels of representation (phonetic form (PF) and logical form (LF)). However, at least according to Katz & Pesetsky, and I follow them in this respect, it is not the case that every musical component may bi-directionally inform every other component. Rather, it turns out that the outputs of grouping structure and metrical structure both inform TSR, which, in turn, informs PR. But if that is the case, the model for a grammar of music can be thought of as these components being directionally ordered, much like different grammatical components are directionally ordered in Minimalist grammar (Figure 3.2). Katz & Pesetsky's implementation of GTTM (Figure 3.1) is the reverse of the reverse Y-model.

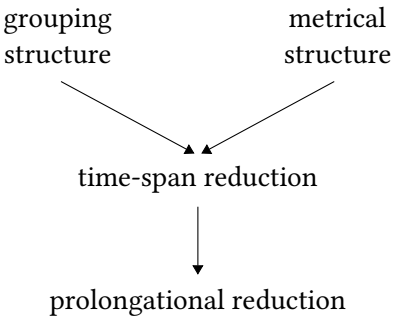


Figure 3.1: Katz & Pesetsky's (2011) (reverse reverse) Y-model of the grammar of music

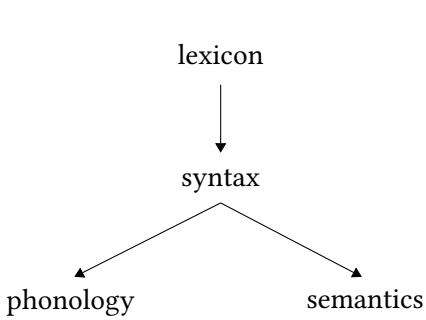


Figure 3.2: The reverse Y-model of the grammar of natural language

If this implementation is correct, the architecture of musical grammar forms a striking correspondence with the architecture of natural language grammar. A particular input is assigned an initial structure that can be derivationally transformed in subsequent structures, with particular well-formedness conditions holding at different levels of representation.

Under this architecture, it can indeed be investigated what the exact parallels are between the syntax of music and the syntax of natural language, and, most notably, whether the differences attested between language and music are merely a consequence of the differences in their building blocks or whether these differences are richer in nature.

3.2 External Merge in music

For Lerdahl & Jackendoff and for Katz & Pesetsky, the correspondence between language and music is stronger than merely being an architecture with various components that together are responsible for the analysis of a structure (irrespective of whether these components are derivationally or representationally connected by means of interface rules). As Lerdahl & Jackendoff already proposed, TSR in GTTM is very similar to prosodic structure in natural language, as both are formulated in terms of relative prominence. Moreover, Katz & Pesetsky take PR to align with linguistic syntax. The reason for them is that both PR and linguistic syntactic structures are binary branching, endocentric (i.e., headed) structures of the kind that is created by (External) Merge in Minimalist grammar. That such structures are headed can be witnessed by the fact that such structures are able to encode dependency relations between non-string-adjacent elements.

To see this, let us focus on the structure of PR. PR structures assign to the pitches a hierarchy that expresses harmonic and melodic tension and relaxation, continuity and progression. Simplifying things, every pitch that increases some kind of tension needs to be followed by some kind of relaxation. However, this need for tension followed up by relaxation is crucially not a string-adjacent condition. In fact, as we will see later on, it may very well be the case that the first tonic already induces a tension that is to be relieved by the final tonic, thus creating a constituent of two sisters whose heads span the entire musical piece. That means that tensions and relaxations in musical structures form non-local dependencies that are best explained as structural dependencies. This intuition is encoded in PRs by assigning head status to any sister of a node that is more relaxed. As an example, take the toy melody in Figure 3.3.

In this structure, the first event (the tonic C) establishes a sisterhood relation with the second event, the tonic being the head. In Western tonal music, tonics are always the most relaxed pitches, whereas pitches or chords based on pitches belonging to other scale degrees are felt to be tenser. Accordingly, the first event in this toy melody is the head of the merger with the second, third, fourth, and fifth events. The fifth event is the dominant (five degrees away from the tonic), which is tensed with respect to the tonic, but more relaxed with respect to the

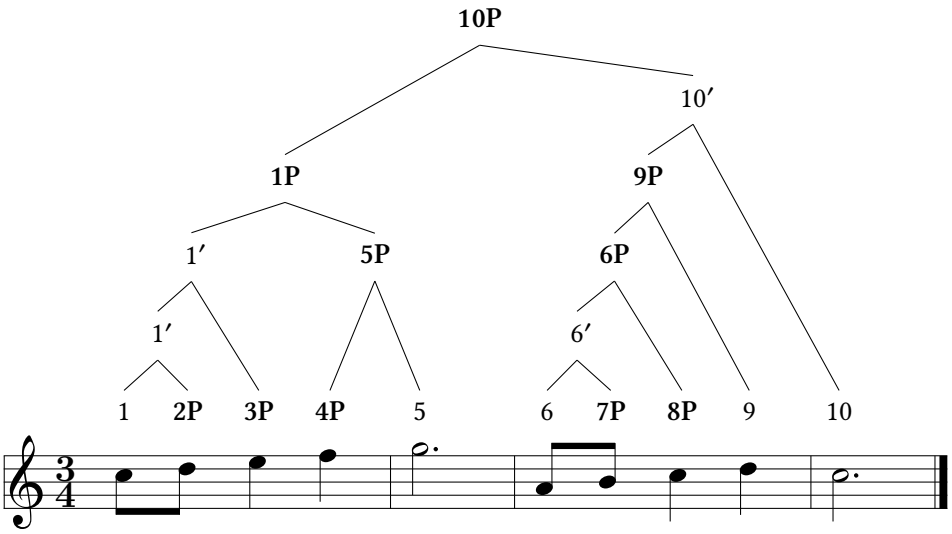


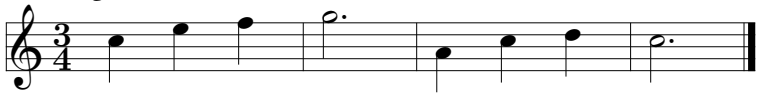
Figure 3.3: Toy melody (Katz & Pesetsky 2011: 16)

so-called subdominant (here, the fourth event), which is four degrees away from the tonic. Similarly, the final pitch (again, a tonic C) creates similar dependencies with the sixth till ninth events. The overall structure then consists of a constituent of two phrases: one in which the tonic in the first event is the head (1P) and one in which the tonic in the tenth event is the head (10P).

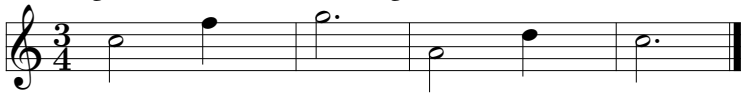
Evidence for this procedure of structure assignments comes from so-called *Schenkerian reductions* (see Forte 1959). Schenkerian reductions are best understood as musical summaries. Going bottom-up, removing every layer of non-heads will still yield a melody that feels like the same kind of melody as the intact structure. This process can in principle be continued until the most prominent chords are left. By contrast, if an event with higher prominence is left out, the piece is no longer perceived as a proper reduction. Examples, taken again from Katz & Pesetsky (2011), are presented below:

(5) Good reductions of Figure 3.3

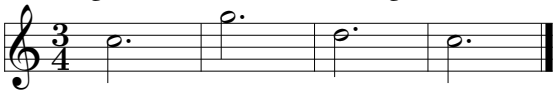
a. Deleting the non-heads of the lower 1′ and of 6′



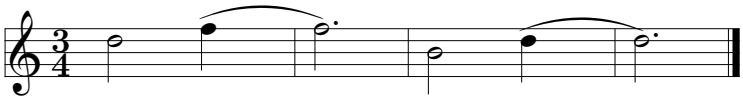
b. Deleting the non-heads of the higher 1′ and of 6P



c. Deleting the non-heads of the higher 1′, 5P and 9P



(6) A bad reduction of Figure 3.3



What does this tell us about Merge in music? The crucial comparison is that the structure-building operation appears to be similar to (External) Merge. Every two musical objects (being atomic or non-atomic) may merge and form a constituent of which the label is the same as that of one of its two daughters (the head). But if that is correct, it can be seen as evidence for there being a "syntactic engine" that is equally active in language and in music. This would, of course, be fully in line with Katz & Pesetsky's identity thesis for language and music. It is the module-specific properties of music that determine what elements can be merged and, once merged, which ones yield the heads (in terms of tension and relaxation, to be computed on the basis of scalar distance with respect to the tonic). But the combinatorial mechanism, Merge, applies to musical objects in exactly the same way as it applies to syntactic objects.

3.3 Internal Merge in music

The previous discussion of External Merge in music sets the ground for the next step in the discussion. If musical structures are indeed built by means of the single generative operation Merge (and the evidence for that claim, confirming the identity thesis for language and music, seems quite strong), then the question arises as to whether only External Merge applies or whether Internal Merge may apply as well. Formally, there is nothing in the combinatorial procedure that would exclude Internal Merge applying to music. Katz & Pesetsky argue that movement effects can indeed be attested in music. Let us first look at the arguments they present for that.

In order to assess whether musical pieces may display movement effects, one should first determine what the proper characteristics of movement in music would be. That task is far from trivial, as general diagnostics for movement (the surface position of some element does not correspond with the locus of its semantic interpretation) do not apply in music, for the simple reason that musical structures lack semantic interpretation (in the sense that musical structures

lack LF). Therefore, the diagnostics for movement should either be formal or PF-like. Moreover, such diagnostics are arguably different for phrasal movement and for head movement. Since Katz & Pesetsky do not provide any evidence for the existence of phrasal movement in music (even though they explicitly do not rule it out per se), but rather focus on head movement only, I will also only discuss what the characteristics of head movement in music would be. The characteristics that Katz & Pesetsky apply for head movement in language and music are given in (7) and (8), respectively:

(7) Head-movement in language (Katz & Pesetsky 2011: 40)
   a. Once the head H of a phrase HP has undergone head movement, H is pronounced string-adjacent to the head of a higher phrase, but at the same time …
   b. … the rest of HP remains an independent phrase that behaves just like a phrase whose head has not moved – even though:
   c. The movement is obligatory. Movement of finite V to T in French satisfies some need of an element in this structure […].
   d. The zero-level head that undergoes head movement to another zero-level head ends up tightly coupled to its new host. The two heads end up behaving like a single morphologically complex word for later processes of grammar (both syntactic and phonological).

(8) Head-movement in music (Katz & Pesetsky 2011: 41)
   a. Some chord X must be performed string-adjacent to a chord Y. But at the same time …
   b. … X has a normal set of syntactic dependents of its own, linearized normally – and thus apparently also heads its own phrase (an XP);
   c. The movement should be obligatory, insofar as it produces an alteration in the features of Y that is required in order for the derivation to succeed;
   d. Even though X may take a normal set of syntactic dependents, X is tightly coupled to its host Y, such that they function as an indivisible unit for other purposes (cf. the notion word).

Here, I will not contest these characteristics for movement, although I would like to point out that these characteristics should be interpreted in a uni-directional way. They are not diagnostics. Even if all effects attributed to head movement are indeed attested, this does not entail that the reverse must be the case

as well. If some α and β are both heads, pronounced string-adjacently, with α altering some feature of β and α and β together taken to form an indivisible unit (i.e., behaving word-like), this does not necessarily entail that α underwent head movement into β. I will come back to that in §4.

Katz & Pesetsky continue their argument by showing that so-called full cadences are a musical phenomenon that shows all the characteristics of head movement. In full cadences, the final chord, the tonic, which determines the key and counts as the head of the entire musical structure, must be preceded by a dominant, a chord whose root is five scale-steps away from the tonic and which has at least one dependent, generally headed by the so-called subdominant, often four scale-steps away from the tonic. In PR, the dominant is directly subordinate to the tonic and occupies a highly prominent position; metrically, it is often felt to be a much weaker chord that seems more deeply embedded in PR and seems to act as a weaker dependent of the tonic. This latter phenomenon is generally referred to as *cadential retention* – the phenomenon that the dominant and the tonic behave almost like a joint chord (and are even analysed as such in GTTM). An example is provided in Figure 3.4, where the dotted arrow (for now) indicates the stronger dependency of the dominant (δ) on the tonic (τ) (ν indicating the subdominant).

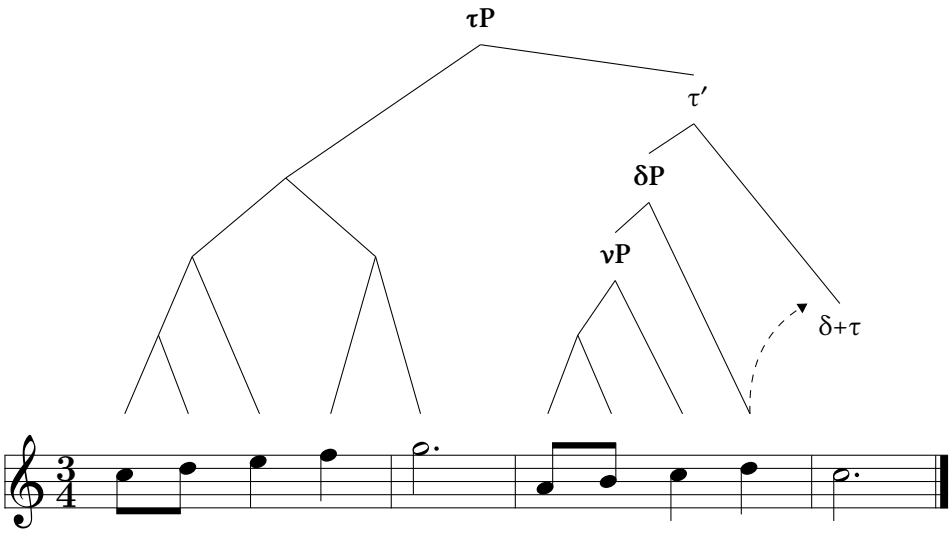


Figure 3.4: Example of a full cadence (Katz & Pesetsky 2011: 44)

Looking at the characteristics of head movement in music, Katz & Pesetsky conclude that full cadences indeed are the result of head movement, and, therefore, of the application of Internal Merge in music.

As for the first two characteristics, if the dominant indeed raises into the head position of the tonic (yielding the structure in (9), where angled brackets indicate lower copies of moved elements), the dominant is expressed string-adjacently to the tonic, even though the dominant still heads a phrase of its own (δP). This way, the construction behaves exactly like the first two clauses of the list of characteristics for head movement (in music).

(9) $[_{\tau P}\ [_{\delta P}\ [_{\nu P}\ \nu \ldots\ ]\ \langle\delta\rangle\ ]\ \delta{-}\tau]$

As for the third characteristic, Katz & Pesetsky claim that movement of the dominant into the tonic marks the tonic for establishing the key of the entire musical piece. They suggest that, in full cadences, movement of the dominant into the tonic head has the function of tonic-marking τ, i.e., assigning it the feature [+TON]. When the tonic head in a structure is tonic-marked, the terminal nodes of the phrase headed by the tonic are understood to belong to the key of τ. In this sense, head-movement of the dominant alters the tonic in having the feature [+TON].

As for the fourth characteristic, finally, Katz & Pesetsky argue that moving the dominant into the tonic position makes the joint dominant–tonic complex act more like a single unit in terms of metric position and makes the dominant look structurally less important than its PR position would legitimize. This joint behaviour, then, is what underlies the phenomenon of cadential retention.

On the basis of this analysis, Katz & Pesetsky conclude that musical structures are indeed generated by means of Merge, and the fact that Merge comprises both External and Internal Merge predicts that musical structures may indeed exhibit movement effects, of which full cadences are then an example. And, if musical structures indeed allow for movement, this forms additional evidence for Merge being the generator of musical structures. However, the reverse is not the case. If it turns out that head movement in music are absent (and that full cadences call for an alternative explanation), the claim that Merge is the sole generator of musical structures, and therefore also the identity thesis for language and music, can still be maintained. The evidence for structural (non-adjacent) dependencies in music and the structural mappings suffice as evidence for (External) Merge. The only question that would arise if (head) movement turns out to be absent in music, is: why is it absent in music despite the generative operation Merge being able to create structures involving movement, whereas (head) movement is so abundantly present in natural language? However, as argued for in §1 and §2, if so-called uninterpretable features are the sole triggers of Internal Merge and those features are absent in music, it is actually predicted that Internal Merge cannot apply in music.

4 Challenging movement in music

Full cadences are the sole cases of alleged (head) movement in music that Katz & Pesetsky present. That means that the validity of the claim that music exhibits movement rests solely on the validity of the argumentation behind their analysis of full cadences as involving head movement. Consequently, in order to maintain that Internal Merge applies in music, it must be shown that (i) full cadences indeed exhibit all the characteristics of head movement and (ii) that these constructions cannot be analysed in alternative terms (or that such an alternative analysis is much weaker). In this section, I argue that full cadences do not show a full parallel with instances of head movement in natural language and that the construction itself calls for an alternative analysis.

One fact that already casts doubt on the claim that music exhibits movement effects is that, outside full cadences, no other clear cases of movement in music have been attested. This is not because Katz & Pesetsky have been the first to look at those effects (although, admittedly, there have been few studies of the kind). Rohrmeier & Neuwirth (2014) discuss particular configurations that may involve movement in music as well, but crucially state that these constructions do not have to be analysed as syntactic movement and therefore do not form any evidence in favour of movement in music. The only other claim of movement in music that I am aware of is Temperley (1999), who notes a parallel between syncopation in rock music and head movement in syntax.

Strikingly, these cases of alleged movement in music are the linguistic equivalent of rightward, string-adjacent head-movement. That, of course, already triggers the question as to why other instances of movement (phrasal movement, non-string-adjacent movement and leftward movement) have so far not been attested in music.

It should be noted in this respect that the core cases of movement in language indeed are cases of leftward, non-string-adjacent movement. That phrasal movement has not been attested as such is not so telling. Both head movement and phrasal movement are indeed solid cases of movement, although head-movement has often been said to be an instance of PF-movement, instead of movement that takes place in narrow syntax (cf. Chomsky 1995; Boeckx & Stjepanović 2001; Harley 2004). However, even if head movement were an instance of PF-movement, this would not invalidate the claim that music exhibits movement effects, as musical structures, just like syntactic structures in language, are to be linearized. In fact, one might even argue that the specific nature of music (with its sole sound side and lack of a meaning side) would rather call for head movement only.

Things are different, however, when it comes to rightward, string-adjacent movement, which has received more scepticism in the linguistic literature. Rightward movement, especially in comparison to leftward movement, is heavily constrained (cf. Ross 1967; Kayne 1994; Cinque 1996; Ackema & Neeleman 2002; Abels & Neeleman 2012). For instance, Kayne (1994) observes that there are verb-second languages but no so-called verb-penultimate languages (where the finite verb appears in the penultimate position). Neither are there languages where *Wh*-terms consequently move to the right (with the possible exception of certain sign languages, cf. Cecchetto et al. 2009). According to Abels & Neeleman (2012), rightward phrasal movement is only possible for full extended projections (that do not strand any parts of it), and according to Ackema & Neeleman (2002), rightward head movement is restricted to moving heads that do not cross any of their dependents. If that is correct, then rightward head movement can only be string-adjacent.

But string-adjacent movement perhaps even calls for more scepticism. How can one determine whether a particular element underwent movement if the linear position of the moved element is the same as its base position? Already in linguistics this is far from clear. In the case of string-adjacent phrasal movement, there might be good reasons to assume that some particular elements indeed undergo movement. For instance, Pesetsky (1987) and Bobaljik (1995; 2002) have argued that subject *Wh*-phrases (like *Who* in *Who left?*) arguably undergo movement from Spec,TP into Spec,CP (to end up in A-bar position) (pace Grimshaw 1997). For head movement things are less clear. Do heads in head-final languages (the only candidates for rightward string-adjacent head movement), such as Korean and Japanese, undergo head movement or not? Is it the case that, in such languages in a configuration like (10), V moves into T and/or T into C?

(10) $[_{CP}\ [_{TP}\ [_{VP}\ V\ ]\ T\ ]\ C\ ]$

Whether languages like Japanese and Korean exhibit string-adjacent rightward head movement or not has been widely discussed in the literature. Various scholars have provided arguments in favour of it. Otani & Whitman (1991) have argued that, in Japanese, the verb must raise to account for various ellipsis effects. The same applies to Koizumi (1995; 2000), who has primarily discussed scrambling and coordination. Also, Yoon (1994) makes an argument in favour of string-adjacent head movement based on coordination of tensed and untensed conjuncts. Choi (1999), finally, formulates an account in terms of NPI licensing that calls for string-adjacent head movement. But as Han et al. (2007; 2016) have shown, basing themselves on arguments by Kim (1995), Chung & Park (1997), Hoji

(1998), Kim (1999), and Fukui & Sakai (2003), all these facts can also be accounted for by approaches that do not allude to rightward head movement. In turn, Han et al. (2007; 2016) argue that head-final languages (Korean is their example) may actually vary language-internally with respect to whether heads undergo raising or not (though see Zeijlstra 2017 for an argument against their claim that some varieties of Korean provide evidence for string-adjacent head movement).

But even if in some languages string-adjacent, rightward head movement can be attested, this does not predict that this is the case for every language. There may be particular language-specific reasons that call for such instances of string-adjacent, rightward head movement, but that does not entail that, in every head-final language, verbs raise into higher heads of the extended projection.

Under the null hypothesis that one should only postulate movement to take place if the data cannot be accounted for otherwise, the question really arises how strong the evidence for movement of the dominant into the tonic position is. What would go wrong if one were to analyse full cadences as instances where the dominant does not raise into the tonic-position but instead just stays in its string-adjacent PR position?

For this, we need to reinvestigate the characteristics of full cadences presented in §3.3. It turns out that, out of the four listed properties, three of them immediately follow by assuming that the dominant stays in situ (11). The fact that the dominant is expressed string-adjacently to the tonic, and the fact that the dominant still heads a phrase of its own (δP) are fully compatible with the analysis in (11).

(11) $[_{\tau P}\ [_{\delta P}\ [_{\nu P}\ \nu \ldots\ ]\ \delta\ ]\ \tau]$

Moreover, the fact that the dominant and the tonic are perceived as one unit (the musical counterpart of being a single word) can also be explained under string-adjacency. Here, the parallel with affixation comes up. Under more traditional concepts of head movement heads raise into higher head positions to ensure realization of the higher head as an affix on the lower head (or vice versa). In that sense, head movement is triggered by the so-called *stray-affix filter* (cf. Lasnik 1981; 1995; Baker 1988) (in any of its guises). For this stray-affix filter to apply, it suffices that the two relevant heads always appear in a string-adjacent position at PF. Now, in head-initial languages, this cannot be guaranteed without alluding to verb movement (due to intervening specifiers/adjuncts), but in head-final languages, where heads are already string-adjacent to each other, it can. Following Bobaljik (1995), an affix can be spelled out on the verb in an OV-language without the verb moving to it, since V and the affix are string-adjacent

at PF. But if that is the case, string-adjacency can suffice as a condition for the dominant and the tonic to be realized as a single unit. Consequently, the fact that the dominant and the tonic end up as one unit does not form evidence for head movement.

This leaves the obligatoriness of head movement as a final possible piece of evidence in favour of an analysis of full cadences in terms of head movement. Head movement in language is obligatory (e.g., movement of finite V to T in French must take place; the finite verb cannot stay in situ). This obligation for head movement is generally understood as a movement-triggering requirement: Some feature of the higher head must be altered for the derivation to proceed, and only raising of another head into this position can establish this feature alteration. For movement, Katz & Pesetsky argue that this feature alteration must be understood as tonic-marking. Movement of the dominant into the tonic position assigns a feature [+TON] to the tonic. Having a tonic feature, in turn, is responsible for this tonic to establish the key of the entire musical piece.

Two questions come to mind here. First, is it necessary that movement triggers such a feature alteration? Can't adjacency suffice here as well? It is known from various impoverishment facts that features present on one head can manipulate the features on a neighbouring head without undergoing movement. Hence, even if the tonic must be tonic-marked by the dominant, this does not have to be realized by means of movement.

Second, is it really the case that the feature of the tonic must be tonic-marked? After all, full cadences are not obligatory in music. Tonics do not require dominants to remerge into their head positions, and neither is it impossible for a dominant to remain in situ (which generally appears to be the case, except perhaps for full cadences). In that sense, head movement of the kind in music is not obligatory in the sense we understand movement to be obligatory in language. What appears to be the case under Katz & Pesetsky's analysis is that movement of the dominant into the tonic is only obligatory under string-adjacency, a much weaker requirement.

But if the structure underlying full cadences is not obligatory for tonic-marking, what one can say is that, at best, it facilitates key establishment. It may help the listener in determining what the key of the entire phrase or piece is. But naturally, other musical facts may play a similar role. For instance, the selection of pitches used in the musical piece already forms a strong (and often sufficient) cue for establishing the key of the entire piece. And also, if harmonic properties determine the PR of a musical piece and if TSR–PR mismatches may only take place under particular circumstances that follow from the underlying PR structure, such mismatches may also provide the listener with a cue of what the key

of the entire piece is. In other words, what full cadences seem to do is facilitate key recognition instead of establishing it.

This all calls for an alternative picture for an analysis of full cadences along the lines of (11), where the adjacency of the dominant and the tonic results in a confirmation of the tonic determining the key and where cadential retention is nothing but the result of an adjacency requirement (a string-adjacent dominant and tonic may or must be realized as a single unit). Already the existence of a viable alternative to the head-movement analysis undermines the status of full cadences as evidence for head movement in music. And this alternative analysis may equally well get the facts right, if not better. But if the only piece of evidence in favour of movement in music turns out to be inconclusive (and may be even incorrect), there is no evidence left any more for the claim that music triggers Internal Merge.

So where do we stand? If full cadences can be equally well, if not better, understood in terms of adjacency requirements, much like Bobaljik (1995) takes such requirements to suffice to establish dependencies between adjacent heads at PF, there appears to be no evidence for movement in music. This allows us to entertain a stronger and more powerful hypothesis, namely that musical structures, despite being generated by Merge, do not exhibit any kind of movement. There is only External Merge going on in music. That amounts to saying that, despite the principled availability of its application, Internal Merge never takes place in music. Given the discussion in §1, where I have argued that that musical building blocks crucially lack the type of features that may trigger Internal Merge and that, consequently, the identity thesis for language and music should predict that Internal Merge never takes place in music, I take this to be a welcome result.

5 Conclusions

In this paper, I have aimed at rethinking remerge. Starting from the premise that uninterpretable features are the sole trigger of Internal Merge, I have looked at another cognitive system, music, to see whether in such a system, where, clearly, (un)interpretable features are absent, Internal Merge may still apply. Focussing on Katz & Pesetsky's elaboration and modification of Lerdahl & Jackendoff's (1983) Generative theory of tonal music, I have evaluated Katz & Pesetsky's claim that musical structures also exhibit movement, and, in particular, their claim that full cadences are to be understood as involving string-adjacent, rightward head movement. My conclusion is that full cadences are equally well, if not better, understood in terms of linear adjacency requirements and that, therefore, the

presented evidence of movement in music does not hold. I have argued that this rather calls for a view of music where movement is absent. However, I have argued as well that this does not speak against Katz & Pesetsky's identity thesis for language and music, but rather speaks in favour of it. Musical structures indeed appear to be generated by means of Merge. However, the absence of uninterpretable features in music prevents Internal Merge from applying in the first place, at least under the assumption that uninterpretable features are the sole trigger for the application of Internal Merge. The reason why music lacks (un)interpretable features is that (un)interpretable features can only emerge in cognitive systems whose building blocks are multi-modular, such as linguistic building blocks. Musical building blocks, by contrast, are mono-modular and can therefore never consist of such (un)interpretable features. The absence of movement in music thus follows directly from the differences between musical and linguistic building blocks and is, therefore, fully in line with Katz & Pesetsky's identity thesis for language and music.

Abbreviations

EPP	extended projection principle	NOM	nominative
GTTM	Generative theory of tonal music	PF	phonetic form
		PR	prolongational reduction
LF	logical form	TSR	time-span reduction

Acknowledgements

I am much indebted to David Pesetsky, who triggered my interest in this topic during his 2009 class on language and music at the EGG summer school in Poznan, and during subsequent discussions. I am also very grateful to Jonah Katz for providing very helpful comments on earlier versions of this work. This paper is the outcome of one of my "Hot topics in language and cognition" classes, taught at the Cognitive Science Center Amsterdam and at the University of Göttingen in 2012–2013, where I discussed Katz & Pesetsky's paper. I thank my students for valuable feedback. Previous versions of this paper have been presented at GLOW in Asia X (held at the National Tsing Hua University, Taiwan), WCCFL 35 (held at Simon Frasier University, Vancouver) and at the Workshop "What drives syntactic computation? Alternatives to formal features", which was part of the annual DGfS meeting in Leipzig in 2015. I would like to thank the organizers and audiences of these events for the opportunity and their comments. All errors, of course, are mine.

References

Abels, Klaus & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74. DOI: 10.1111/j.1467-9612.2011.00163.x.

Ackema, Peter & Ad Neeleman. 2002. Effects of short-term storage in processing rightward movement. In Sieb Noteboom, Fred Weerman & Frank Wijnen (eds.), *Storage and computation in the language faculty*, 219–256. Dordrect: Kluwer.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2009. Linearization and the architecture of grammar: A view from the Final-over-Final Constraint. In Vincenzo Moscati & Emilio Servidio (eds.), *Proceedings of XXXV Incontro di Grammatica Generativa*, vol. 3 (University of Siena CISCL Working Papers), 77–89. MITWPL.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Biberauer, Theresa & Ian Roberts. 2015. Clausal hierarchies. In Ur Shlonsky (ed.), *Beyond functional sequence* (The cartography of syntactic structures 10), 295–313. Oxford: Oxford University Press.

Bjorkman, Bronwyn M. & Hedde Zeijlstra. 2019. Checking up on (φ-)Agree. *Linguistic Inquiry* 50(3). 527–569. DOI: 10.1162/ling_a_00319.

Bobaljik, Jonathan David. 1995. *Morphosyntax: The syntax of verbal inflection*. MIT. (Doctoral dissertation).

Bobaljik, Jonathan David. 2002. A-chains at the PF-interface: Copies and "covert" movement. *Natural Language & Linguistic Theory* 2(20). 197–267. DOI: 10.1023/A:1015059006439.

Bobaljik, Jonathan David. 2008. Where's phi? Agreement as a postsyntactic operation. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*, 295–328. Oxford: Oxford University Press.

Boeckx, Cedric & Sandra Stjepanović. 2001. Head–ing toward PF. *Linguistic Inquiry* 32(2). 345–355. DOI: 10.1162/00243890152001799.

Bošković, Željko. 2007. On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry* 38. 589–644. DOI: 10.1162/ling.2007.38.4.589.

Brody, Michael. 1997. Perfect chains. In Liliane Haegeman (ed.), *Elements of grammar*, 139–167. Dordrecht: Kluwer.

Cecchetto, Carlo, Carlo Geraci & Sandro Zucchi. 2009. Another way to mark syntactic dependencies: The case for right-peripheral specifiers in sign languages. *Language* 85(2). 278–320. DOI: 10.1353/lan.0.0114.

Choi, Young-Sik. 1999. Negation, its scope and NPI licensing in Korean. In Rebecca Daly & Anastasia Riehl (eds.), *ESCOL [the Eastern states conference on linguistics] '99*, 25–36. Ithaca, NY: CLC Publications.

Chomsky, Noam. 1995. *The Minimalist program.* Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.

Chung, Daeho & Hong-Keun Park. 1997. NPIs outside of negation scope. In Homin Sohn & John Haig (eds.), *Japanese/Korean linguistics*, vol. 6, 415–435. Stanford, CA: CSLI Publications.

Cinque, Guglielmo. 1996. The antisymmetric programme: Theoretical and typological implications. *Journal of Linguistcs* 32(2). 447–464. DOI: 10.1017/S0022226700015966.

Culicover, Peter W. & Ray Jackendoff. 2005. *Simpler syntax.* Oxford: Oxford University Press.

Deal, Amy Rose. 2009. The origin and content of expletives: Evidence from "selection". *Syntax* 12(4). 285–323. DOI: 10.1111/j.1467-9612.2009.00127.x.

Forte, Allen. 1959. Schenker's conception of musical structure. *Journal of Music Theory* 3(1). 1–30. DOI: 10.2307/842996.

Fukui, Naoki & Hiromu Sakai. 2003. The visibility guideline for functional categories: Verb-raising in Japanese and related issues. *Lingua* 113(4–6). 321–375. DOI: 10.1016/S0024-3841(02)00080-3.

Grimshaw, Jane. 1997. Projection, heads, and optimality. *Linguistic Inquiry* 28(3). 373–422.

Han, Chung-hye, Jeffrey Lidz & Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38(1). 1–47. DOI: 10.1162/ling.2007.38.1.1.

Han, Chung-hye, Julien Musolino & Jeffrey Lidz. 2016. Endogenous sources of variation in language acquisition. *PNAS* 113(4). 942–947. DOI: 10.1073/pnas.1517094113.

Harley, Heidi. 2004. Merge, conflation, and head movement: The First Sister Principle revisited. In Keir Moulton & Matthew Wolf (eds.), *NELS 34: Proocedings of the thirty-fourth annual meeting of the North East Linguistic Society*, 239–254. Amherst, MA: GLSA Publications.

Hoji, Hajime. 1998. Null object and sloppy identity in Japanese. *Linguistic Inquiry* 29(1). 127–152.

Jackendoff, Ray. 1997. *The architecture of language*. Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of language*. Oxford: Oxford University Press.

Katz, Jonah & David Pesetsky. 2011. The identity thesis for language and music. lingBuzz/000959.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kim, Jong-Bok. 1995. On the existence of NegP in Korean. In Susumo Kuno (ed.), *Harvard studies in Korean linguistics: VI. Proceedings of the 1995 Harvard International Symposium on Korean, January 13–15, 1995*, 267–282. Cambridge, MA: Harvard University.

Kim, Soowon. 1999. Sloppy/strict identity, empty objects, and NP ellipsis. *Journal of East Asian Linguistics* 8(4). 255–284. DOI: 10.1023/A:1008354600813.

Koizumi, Masatoshi. 1995. *Phrase structure in minimalist syntax*. Cambridge, MA: MIT. (Doctoral dissertation).

Koizumi, Masatoshi. 2000. String vacuous overt verb raising. *Journal of East Asian Linguistics* 9(3). 227–285. DOI: 10.1023/A:1008311420647.

Lasnik, Howard. 1981. Restricting the theory of transformations: A case study. In Norbert Hornstein & David W. Lightfoot (eds.), *Explanation in linguistics*, 152–173. London: Longman.

Lasnik, Howard. 1995. Verbal morphology: Syntactic structures meets the minimalist program. In Hector Campos & Paula Kempchinsky (eds.), *Evolution and revolution in linguistic theory*, 251–275. Georgetown: Georgetown University Press.

Lerdahl, Fred & Ray Jackendoff. 1983. *A generative theory of tonal music*. Cambridge, MA: The MIT Press.

Neeleman, Ad & Hans Van de Koot. 2008. Dutch scrambling and the nature of discourse templates. *Journal of Comparative Germanic Linguistics* 11(2). 137–189. DOI: 10.1007/s10828-008-9018-0.

Otani, Kazuyo & John Whitman. 1991. V-raising and VP-ellipsis. *Linguistic Inquiry* 22(2). 345–358.

Pesetsky, David. 1987. Wh-in-situ. In Eric J. Reuland & Alice G. B. ter Meulen (eds.), *Movement and unselective binding*, 98–129. Cambridge, Massachusetts: MIT Press.

Richards, Norvin. 2016. *Contiguity theory*. Cambridge, MA: MIT Press.

Rohrmeier, Martin & Markus Neuwirth. 2014. Towards a syntax of the classical cadence. In Markus Neuwirth & Pieter Bergé (eds.), *What is a cadence? Theoretical and analytical perspectives on cadences in the classical repertoire*, 285–336. Leuven: Leuven University Press.

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT. (Doctoral dissertation).

Schlenker, Philippe. 2016. Prolegomena to music semantics. lingBuzz/002925.

Starke, Michal. 2001. *Move dissolves into Merge: A theory of locality*. University of Geneva. (Doctoral dissertation).

Svenonius, Peter. 2007. Interpreting uninterpretable features. *Linguistic Analysis* 33(3–4). 375–413.

Temperley, David. 1999. Syncopation in rock: A perceptual perspective. *Popular Music* 18. 19–40. DOI: 10.1017/S0261143000008710.

Yoon, James. 1994. Korean verbal inflection and checking theory. In Heidi Harley & Colin Phillips (eds.), *The morphology-syntax connection*, vol. 22 (MIT Working Papers in Linguistics), 251–270. Cambridge, MA: MITWPL.

Zeijlstra, Hedde. 2008. On the syntactic flexibility of formal features. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 143–174. Amsterdam: John Benjamins. DOI: 10.1075/la.132.06zei.

Zeijlstra, Hedde. 2012. There is only one way to agree. *The Linguistic Review* 29(3). 491–539. DOI: 10.1515/tlr-2012-0017.

Zeijlstra, Hedde. 2017. *Two varieties of Korean or the ban against rightward head movement*. Paper presented at NELS [North East Linguistics Society] 48, University of Iceland, Reykjavik, 27–29 October, 2017.

Chapter 4

Life without word classes: On a new approach to categorization

István Kenesei
Research Institute for Linguistics, Budapest, & University of Szeged

This is an attempt to redefine word classes, or more precisely, to replace the concept of word class with clusters of properties much like the notion of the phoneme is dissolved into the various combinations of distinctive features. It is claimed that word classes are but comfortable generalizations not supported by hard evidence as seen in examples from a select group of languages and illustrated in detail by the list of auxiliaries in Hungarian.

1 Introduction and overview

The problem of the definition of word classes has been with us since the very beginnings of linguistics. The first grammars already provided terms according to which to classify words. Dionysius Thrax (BCE 170–90) lists the following eight classes: noun, verb, participle, article, pronoun, preposition, adverb, conjunction. The definitions are simple, familiar, and of course mostly notional, e.g.,

> A Noun is a declinable part of speech, signifying something either concrete or abstract (concrete, as stone; abstract, as education); common or proper (common, as man, horse; proper, as Socrates, Plato). It has five accidents: gender, species, forms, numbers, and cases.
>
> (*The grammar of Dionysios Thrax*, this citation from Davidson 1874: 331)

The classical definitions have followed us well into the 20th century. To quote another example, this is what the Port-Royal philosophers had to say about parts of speech in the 17th century:

István Kenesei. 2020. Life without word classes: On a new approach to categorization. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 67–80. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280633

> Les objets de nos pensées, sont ou les choses, comme *la terre*, *le Soleil*, *l'eau*, *le bois*, ce qu'on appelle ordinairement *substance*. Ou la manière des choses; comme d'estre *rouge*, d'estre *dur*, […] & c. ce qu'on appelle *accident*. […] Car ceux qui signifient les substances, ont esté appellez *noms substantifs*; & ceux qui signifient les accidens […], *noms adjectifs*.
>
> (Lancelot & Arnauld 1660/1967: 30–31)

This type of definition was widespread until about the middle of the 20th century. In his otherwise highly original *Grammar of spoken English*, Palmer (1924) lists more or less the same eight classes, viz., nouns, pronouns and determinatives, qualificatives (i.e., adjectives), verbs, adverbs, prepositions, connectives ("together with interrogative words"), and interjections and exclamations. In the "logical classification of nouns", for instance, he gives an inventory of subtypes, rather than a classical definition, namely, concrete nouns (including proper and common nouns, with the latter further divided into class, i.e. countable, and material nouns, etc.) and abstract nouns (Palmer 1924: 28–32).

However, due to the influence of Saussure's *Cours* (1916), American descriptive linguists, and in particular Leonard Bloomfield, who was the first of them to appreciate Saussure's achievements (cf., e.g., Koerner 1995), started to concentrate on the formal features of parts of speech. "The noun is a word-class; like all other form-classes, it is to be defined in terms of grammatical features […] When it has been defined, it shows a class-meaning which can be roughly stated as "object of such and such a *species*"; examples are *boy*, *stone*, *water*, *kindness*." (Bloomfield 1935: 202) One of Bloomfield's more dogmatic followers had this to say in his widely used textbook:

> [The pattern of interchangeability] defines a form-class which includes *she*, *he*, *it*, *John*, *Mary*, *the man at the corner*, *my friend Bill*, and so on endlessly, but which by no means includes all forms, since we can name many which are excluded: *her*, *him*, *them*, *me*, *yes*, *no*, *ripe*, *find her*, *go with us tomorrow*.
>
> (Hockett 1958: 162)

Note that Hockett's form-classes include not only words proper, but entire phrases, and there is no "class-meaning" mentioned, since the most important feature is mutual substitutability.

But if distributional analysis is closely observed, its negative consequences are unavoidable, as was seen as early as the 1960s. According to one British linguist "as many classes are set up as words of different formal behaviour are found" (Robins 1980 [1964]: 174), and another maintains in an article on the definition of word classes that "[…] very few words have an overall identical formal behaviour […]. One would end up with a multitude of single member classes" (Crystal 1967:

28). Or to cite a more recent article: "Whatever identifying criteria we use for parts of speech – meaning, syntactic function, or inflection – the relationship between particular criteria and particular parts of speech is typically many-to-many" (Anward 2000: 3).

Neither do alternative approaches fare better in this respect. Functionalist linguists, as shown by Simon Dik (1989) or Kees Hengeveld (1992), differentiate word classes by two prototypical functions or parameters, such as predication vs. referentiality, and head vs. modifier, with the resulting four classes arranged in an implicational hierarchical order in (1) that corresponds to the sequence verb > noun > adjective > adverb (Hengeveld 1992).

(1)	Head of pred. phrase	>	Head of ref. phrase	>	Modifier of head of ref. phrase	>	Modifier of head of pred. phase

The "radical constructionist" William Croft (2005) also notes the futility of the distributional method, and, instead of language specific word classes, proposes restricted typological universals based on "propositional acts", such as reference, predication, and modification, that define "lexical semantic classes" like objects, actions, and properties, respectively (Croft 2005: 438).

As I will try to show, neither the approach based on the introduction of a new or different set of criteria for the same small number of word classes nor the opposing view stemming from otherwise well-established criticism based on the failure of distributional analysis is viable. Instead, I will suggest a compromise solution that benefits from both without their possible drawbacks.

Research into the typology of word classes has come up with observations differentiating between part-of-speech systems depending on whether or not the categories of lexical items are fixed or not. Languages can thus be grouped into one of three sets: (a) differentiated, as English, in which all four word classes are clearly displayed, and two subtypes in which such dedicated lexical items are missing: (b) flexible, like Turkish, in which non-verbs can belong to any one of the three classes nouns, adjectives, and adverbs, and (c) rigid, like Krongo (Kadu, Sudan), in which there are nouns and verbs, but the rest of the lexical categories are rendered by syntactic means, e.g., relative clauses (Hengeveld 2013: 32ff.).[1]

[1]Due credit must be given here to the polyglot phonologist and theoretical linguist Ferenc Mártonfi (1945–1991), who had expressed similar thoughts well ahead of the recent upsurge of interest in word class typology, as illustrated in the following passage. "From the point of view of parts-of-speech this means that there are languages in which syntactic features like 'verbal' or 'nominal' must be marked for all or most of the words (e.g., in Hungarian, German, etc.), and there are languages where this would be redundant, non-distinctive marking, which is omissible (and this holds for the large majority of words in, e.g., Chinese, Vietnamese, etc. […]). In other words, this means that lexical word classes are not universal." (Mártonfi 1973: 201; my translation)

It is true that Distributed Morphology offers an attractive solution to the problem of word classes by merging a functional category with an unspecified root (cf. Halle & Marantz 1993; Marantz 1997; Arad 2003; Panagiotidis 2015, among others). In this approach, categorization is a syntactic process. Items, whether heads or phrases, have no categories of their own determined by their lexical characterization, but acquire them, as it were, by becoming complements of functional heads, such as the nominalizer *n*, the verbalizer *v*, or the adjectivizer *a* (Panagiotidis 2015: 17). However, Baker's (2003: 266ff.) arguments are persuasive in attributing syntactic categories to roots or stems, particularly, as I would focus on his proposal in the light of the above typology, in the case of a number of languages in the "differentiated" type, which will be the subject of our discussion below. Baker claims that "where there is less functional structure, we find more categorial distinctiveness" (Baker 2003: 268).

2 Properties rather than definitions

Traditional part-of-speech characterizations usually list the most general properties and illustrate them by prototypical examples, which serve practically as ostensive definitions, thus rendering the characterization itself redundant since the examples are a sufficient ground for any competent native speaker by means of which to classify the words of the language in question. The criteria, which usually rely on distributional and/or semantic factors, are usually too soft or porous, and the classes set up do not directly follow from the definitions.

At the same time these very definitions preclude the establishment of, for example, the uniform class of verbs in English or in other languages of the differentiated type since intransitive verbs are as a rule incapable of substituting for transitive ones, or mass nouns for countable nouns, and so forth. If, however, we are satisfied with partial overlapping, then the class of adjectives will in part coincide with that of nouns, cf. *Italian* or *(the) blind*, or even adjectives will subsume two partially overlapping subsets, relational and qualitative ones, cf. *(*more) naval (exercise)* vs. *(more) interesting exercise*. In addition to flexible word classes (cf. Rijkhoff & van Lier 2013), some dispute the distinction between inflection and derivation as well, positing a continuum for them (Dressler 1989). What is to blame in this state of affairs is the metric applied; if we have a single scale, the difficulties will inevitably resurface again.

Moreover, it follows from a unidimensional system of criteria that whenever some word class is defined by a set of characteristics, then a given item belongs to that word class if it has precisely those characteristics. If any item has some

property that it shares with another item, the property will serve to determine the class formed by them. This is clearly circular and if we insist on this approach the circle cannot be broken.

Note that the notion of word class applies only to linguistic items that can combine with other such items. Utterance-sized words, such as interjections, greetings, etc., even though they may be listed and categorized in dictionaries, do not partake in syntactic constructions (except in citation forms), thus, theoretically speaking they have no properties comparable to those of "ordinary" word classes, while the labels attached to them certainly have a practical advantage for users of these dictionaries.

It is precisely the (morphological, syntactic, semantic, or pragmatic) properties of combinable lexical items relevant from the viewpoint of categorization that control their cooccurrence with other lexical items. Consequently, there will be as many classes as there are properties, thus vindicating Robins's (1980 [1964]), Crystal's (1967), or Anward's (2000) views of a multitude of word classes. But these definitions will no longer be circular since the criteria they are based on will figure in various levels of grammar in determining the combination of items, that is, in morphology, syntax, semantics, and pragmatics.

Consequently, what we understand by a word class will be a set of instructions specifying what other lexical or syntactic objects, whether affixes, words or syntactic phrases, a given word can combine with. "Traditional" word classes, i.e., nouns, verbs, adjectives, adverbs, satisfy various clusters of properties. In effect, the unidimensional category of word class has been replaced by multidimensional matrices of sets of properties.

A similar suggestion is inherent in Crystal's (1967: 46) list of criteria for nouns in English, reproduced in Figure 4.1.

Gross (1986) gives a classification of French verbs according to the types of subjects, complements and the properties of their complements, based on 4 subject and 32 complement types, setting up a matrix of 36 verb types.

In a discussion of the problems of universal and language specific classification Haspelmath (2012: 94) presents the overlapping system of word classes in Chamorro, following Topping (1973) and Chung (2012), according to the properties and classes as in Table 4.1.

In contrast with more "regular" languages like Latin, which has the two major classes of verbs and nouns, with the two subclasses nouns (*nomen substantivum*) and adjectives (*nomen adjectivum*) in the latter group as distinguished by properties of having case and (in)variable gender, Haspelmath argues that Chamorro has six possible word class systems in view of the properties in Table 4.1, as illustrated in Figure 4.2.

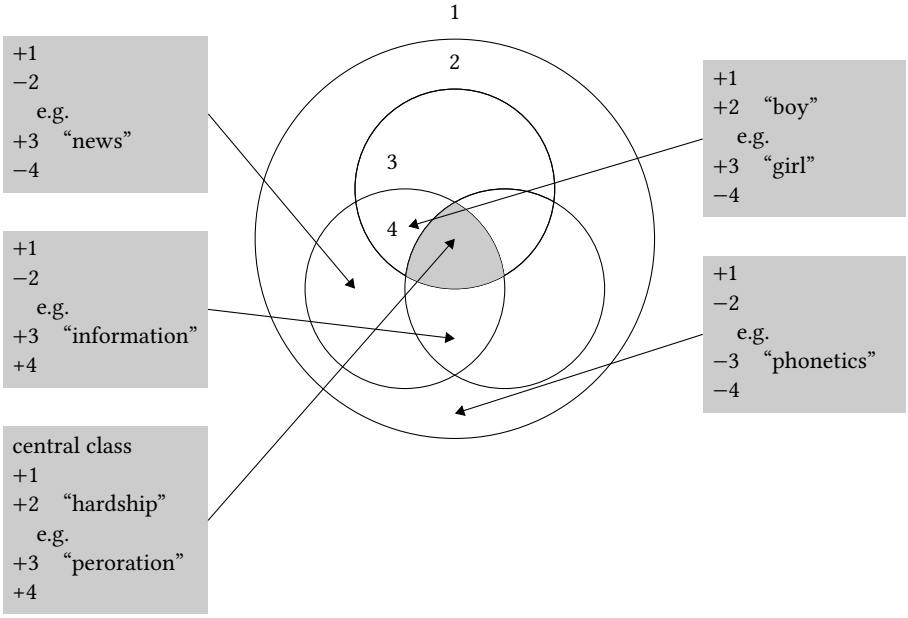


Figure 4.1: Crystal’s (1967) criteria for nouns. *Legend*: 1 – May act as subject; 2 – Inflect for number; 3 – Co-occur with article; 4 – Morphological indication.

Table 4.1: Haspelmath’s (2012) extension of Chung’s (2012) table of grammatical properties and clauses in Chamorro

	Word type			
Property	‘see’	‘go’	‘big’	‘person’
passive	+	−	−	−
Yo’-type pronoun subject	−	+	+	+
Infinitive	+	+	+	+
Incorporation	−	−	−	+
Prefixation with *mi-*	−	−	−	+
Subject-predicate agreement	+	+	+	−
Specific external agreement	+	+	−	−
Person-number agreement (realis)	+	−	−	−
	Transitival	Intransiverb	Adjectival	Nominal

A	Transitival + Intransiverb	Adjectival + Nominal	
B	Transitival	Intransiverb + Adjectival + Nominal	
C	Transitival + Intransiverb + Adjectiverb		Nominal
D	Transitival + Intransiverb	Adjectival	Nominal
E	Transitival	Intransiverb	Adjectival + Nominal
F	Transitival	Intransiverb + Adjectival	Nominal

Figure 4.2: The six possible word class systems of Chamorro according to Haspelmath (2012)

The properties in question can be of various ranks and significance, as claimed by Crystal (1967), since some may extend to more items than others, e.g., whether or not it can be a subject, take a definite article, etc. Then there are classes that can easily adopt new items, whereas others do not – a familiar distinction between open and closed classes. But closed classes, i.e., grammatical words or functional categories, do not form unified classes at all.

This was shown, for example, by Radford (1976) in classifying English auxiliaries by listing six properties distinguishing auxiliaries from verbs, such as the ability to take negative clitics, to take *do*-support, to nominalize, to occur in untensed clauses, to occur in untensed clauses, to take *to* before a following infinitive, and to display concord, all of which, except for the first, are properties characterizing verbs.

Aarts (2007) differentiates between subjective and intersective gradience, where the former is a case of "categorial shading in prototypicality from a central core to a more peripheral boundary" in a single category, while in the latter "there are two categories on a cline" (p. 97). Rendered in the framework presented here, it is the relevance and/or number of features from one or the other word class that determine to what degree the item in question belongs to one or the other category in Aarts' intersective gradience.

If we examine auxiliaries in Hungarian, we can identify the following properties that distinguish them from main verbs that also take infinitives as their complements.[2]

[2]Note that the first two properties (2) and (3) below lump together subclasses of main verbs with (some) auxiliaries.

(2) Split complex verbs, i.e., a combination of particle + verb construction, see (2b,c)[3]

i. *utál* 'hate', *szégyell* 'be ashamed to', ...

ii. *akar* 'want', *próbál* 'try', *tud* 'know, can', ...

iii. *fog* 'will', *szokott* 'usually does', *kell* 'must', *szabad* 'may, is allowed to', *talál* 'happen to', passive *van* + V-*va/ve*

a. * be utál-sz jön-ni
in hate-2SG come-INF

a′. utál-sz be jön-ni
hate-2SG in come-INF
'you hate to come in'

b. be akar-sz jön-ni
in want-2SG come-INF
'you want to come in'

c. be fog-sz jön-ni
in want-2SG come-INF
'you will come in'

(3) Has tense/modal meaning
tud 'be.able', *bír* 'can', *fog* 'will', *kell* 'must', *szabad* 'may, is allowed to', *lehet* 'may, is possible'

(4) Has no present or past tense forms
szokott 'usually does'; *fog* 'will', *szabad* 'may, is allowed to'

(5) Has no person or number agreement
kell 'must', *szabad* 'may, is allowed to', *lehet* 'may, is possible'

(6) Has no infinitival form
fog 'will', *szokott* 'usually does', *szabad* 'is allowed to', *lehet* 'is possible'

(7) Has no thematic subject (external argument)
fog 'will', *szokott* 'usually does', *talál* 'happen to'; passive *van* + V-*va/ve*, *kell* 'must', *szabad* 'is allowed to', *lehet* 'is possible'

[3] As is illustrated in (2i) and (2a,a′), not all verbs can split the complex verbs in their complement infinitivals. Those that do are listed in (2ii–iii) and illustrated in (2b,c), where (2ii) are examples of main verbs and (2iii) those of auxiliaries, as seen in Table 4.2. The phenomenon was first described by Prószéky et al. (1984) and in more detail by Kálmán C. et al. (1989), though their conditions are not followed here, cf. also Kenesei (2000).

(8) Has no potential inflection (i.e., missing *-hat/het* 'may' affixation)
fog 'will', *szokott* 'usually does', *talál* 'happen to'; *szabad* 'is allowed to'

(9) Has no conditional inflection (i.e., missing *-na/ne* 'would' affixation)
fog 'will', *szokott* 'usually does'

(10) Has no imperative/subjunctive forms
fog 'will', *szokott* 'usually does', *talál* 'happen to'

(11) Has person-marked infinitival complements, see (11a)
kell 'must', *szabad* 'is allowed to', *lehet* 'is possible'

a. Hungarian
jön-ni-ük kell
come-INF-3PL must
'they must come'

Moreover, the above list is augmented by restrictions on syntactic positions, i.e., what complement VPs each verb in the list can take, cf. (12).

(12) Hungarian

a. be fog kell-eni tud-ni jön-ni
in will must-INF be.able-INF come-INF
'it will be necessary to be able to come in'

b. * be fog tud-ni kell-eni jön-ni
in will be.able-INF must-INF come-INF
intended: '(someone) will be able to have to come in'

c. * be fog/szokott talál-ni jön-ni
in will/usually happen-INF come-INF
intended: '(s/he) will/usually happen/s to come in'

These properties set apart main verbs (in bold type, with each exemplifying a large array) and the single items of auxiliaries (in normal type). And, what is more important, there are no two auxiliaries that are characterized by the same set of features, as shown in Table 4.2, in which the lack of a property is marked by a minus sign.[4]

Starting with the fourth column there are only "classes" containing single items, and it is precisely these words that qualify as auxiliaries, which points

[4]The star in the last cell indicates the irrelevance of the property. The ± sign in column 2 shows that some verbs in this group have modal meanings, and in column 3 that speakers vary as to the acceptability of the past tense form of *szabad*.

Table 4.2: Feature matrix for Hungarian verbs and auxiliaries

Verb	Splitting	Modal meaning	Past tense	Person/number marking	Infinitive	Thematic subject	Potential	Conditional	Imperative	Person-marked infinitive
utál 'hate'	−	−	+	+	+	+	+	+	+	−
akar 'want'	+	±	+	+	+	+	+	+	+	−
bír 'can'	+	+	+	+	+	+	+	+	+	−
fog 'will'	+	+	−	+	−	−	−	−	−	−
szokott 'usually'	+	+	+	+	−	−	−	−	−	−
talál 'happen'	+	−	+	+	+	−	−	+	−	−
kell 'must'	+	+	+	−	+	−	−	+	+	+
szabad 'may'	+	+	±	−	−	−	−	+	+	+
lehet 'may'	+	+	+	−	−	−	−	+	+	+
PASSIVE *van* 'be'	+	−	+	+	+	−	+	+	+	*

at property (7) as the one distinguishing them from main verbs, or more precisely, main verbs that take infinitival clauses as complements.[5] Note, however, that the lack of a thematic subject/external argument is a property found also in unaccusative verbs, but they, in turn, do not take infinitival complements, and Table 4.2 was set up to include verbs with infinitival complements only. Again, it is another instance of cross-classification, as is generally the case with the open class of (main) verbs, but the ultimate lesson is that the word class of auxiliaries does not seem to emerge, because the rest of the features are not shared by any two of the items listed in Table 4.2.

3 Conclusion: Life without word classes

We could go on to demonstrate similar one-member classes in case of articles, conjunctions, and other functional categories, but, as was seen above, categories in open classes are also prone to a limitless multiplication of classes. The way out of this impasse is at hand: word class is an epiphenomenon, it is not a basic

[5] See Kenesei (2006) for a full set of arguments.

concept but a derivative notion in linguistics. There are no word classes; what we have to do with is properties and their combinations, clusters, or matrices. The morphological and syntactic environment, including the complements of individual functional or notional items, can be determined also by various combinations of properties, spelling them out as the characterizations of individual items as we have seen in the case of the auxiliaries.

Morphological or syntactic processes rely and work on properties rather than (classes of) words or morphemes, which renders the discussion on whether word classes are universal or language-specific irrelevant (Hengeveld 1992; Croft 2005; Haspelmath 2012 etc.). What can be universal is not some word class but a set of distinctive properties, some of which were illustrated above. Since there are probably no languages without subjects, Crystal's (1967) feature of "May act as subject" is probably universal.[6] It is likely that all languages have a property of "May have a complement", and if there are cases in a language, then it makes sense to posit the feature "Assigns (structural) case". But just as the consonantal phonological feature for clicks may be relevant only in Bantu languages, it is possible that the syntactic feature of incorporation, which is significant in Chamorro, is missing in a large number of languages. And with reference to the languages with "flexible word classes", as well as to the decomposition of categories in Distributed Morphology, it may very well be the case that the syntactic categorizing heads, i.e., the "categorizers" that merge with categorially unspecified lexical items, are themselves bundles of properties along the lines discussed here.

There is hardly anything surprising in this development, especially if we take into account the fact that it is no longer the phoneme that is the basic unit in phonology but distinctive features and the term phoneme is but shorthand for sets of distinctive features, as seen in the following passage:

> In recent years it has become widely accepted that the basic units of phonological representation are not segments but features, the members of a small set of elementary categories which combine in various ways to form the speech sounds of human languages. (Clements & Hume 1995: 245)[7]

[6] One anonymous reviewer contests my reliance on this property, cf.: "The author says 'there are probably no languages without subjects' but that is a statement which has frequently been contested by those who work on so-called 'topic prominent' languages". My studies of topic-prominent languages, which include Hungarian, among others, do not, however, confirm this statement, but cf. also e.g., É. Kiss (2002) for a more complete overview. This reviewer also maintains that "various theories do without a core concept of 'subject' (including most if not all versions of generative grammar), while others such as Lexical-Functional Grammar (LFG) and Relational Grammar make it a theoretical primitive." While this is indeed the case, the fact that 'subject' is a derived notion, rather than a core concept, in generative grammars does not preclude reference to it by the properties invoked here.

[7] See also Siptár (2006).

And finally, just as phonologists have not got rid of the term "phoneme", so syntacticians or morphologists need not throw out the notion of "word class" – if they are aware that it is a convenient abbreviation without any consequence or theoretical relevance.

Abbreviations

2	second person	LFG	Lexical-Functional Grammar
3	third person	PL	plural
INF	infinitive	SG	singular

Acknowledgements

This article has grown out of a number of presentations to various audiences, e.g., at the 14th *morphology meeting*, and the *beyond dichotomies* conference, both in Budapest, 2010, the Research Institute for Linguistics, and the Linguistics and Literature Section of the Hungarian Academy of Sciences. I am grateful to the audiences there, and in particular to László Kálmán and Péter Siptár. My special thanks go to the two anonymous reviewers of the current version. Research reported here was supported by Grant NKFIH K120073 "Open access book series on the syntax of Hungarian".

References

Aarts, Bas. 2007. *Syntactic gradience: The nature of grammatical indeterminacy*. Oxford: Oxford University Press.

Anward, Jan. 2000. A dynamic model of part-of-speech differentiation. In Petra M. Vogel & Bernard Comrie (eds.), *Approaches to the typology of word classes*, 3–45. Berlin: De Gruyter. DOI: 10.1515/9783110806120.3.

Arad, Maya. 2003. *Natural Language & Linguistic Theory* 21(4). 737–778. DOI: 10.1023/a:1025533719905.

Baker, Mark C. 2003. *Lexical categories: Verbs, nouns and adjectives*. Cambridge: Cambridge University Press.

Bloomfield, Leonard. 1935. *Language*. 2nd revised edition. London: George Allen & Unwin.

Chung, Sandra. 2012. Are lexical categories universal? The view from Chamorro. *Theoretical Linguistics* 38(1–2). 1–56. DOI: 10.1515/tl-2012-0001.

Clements, George N. & Elizabeth Hume. 1995. The internal organization of speech sounds. In John Goldsmith (ed.), *Handbook of phonological theory*, 245–306. Oxford: Blackwell.

Croft, William. 2005. Word classes, parts of speech, and syntactic argumentation. *Linguistic Typology* 9(3). 431–441. DOI: 10.1515/lity.2005.9.3.391.

Crystal, David. 1967. English. *Lingua* 17. 24–56. DOI: 10.1016/0024-3841(66)90003-9.

Davidson, Thomas. 1874. The grammar of Dionysios Thrax. Trans. by Thomas Davidson. *The Journal of Speculative Philosophy* 8(4). 326–339. http://www.jstor.org/stable/25665891.

Dik, Simon C. 1989. *The theory of functional grammar: Part I: The structure of the clause*. Dordrecht: Foris.

Dressler, Wolfgang U. 1989. Prototypical differences between inflection and derivation. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42. 3–10.

É. Kiss, Katalin. 2002. *The syntax of Hungarian*. Cambridge: Cambridge University Press.

Gross, Maurice. 1986. *Lexicon – grammar: The representation of complex words*. Paris: LDL Université de Paris 7. https://halshsarchives-ouvertes.fr/halshs-00278082/document.

Halle, Moris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth L. Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 111–176. Cambridge, MA: MIT Press.

Haspelmath, Martin. 2012. Escaping ethnocentrism in the study of word-class universals. *Theoretical Linguistics* 38(1–2). 91–102. DOI: 10.1515/tl-2012-0004.

Hengeveld, Kees. 1992. *Non-verbal predication: Theory, typology, diachrony*. Berlin: De Gruyter.

Hengeveld, Kees. 2013. Parts-of-speech systems as a basic typological determinant. In Jan Rijkhoff & Eva van Lier (eds.), *Flexible word classes: Typological studies of underspecified parts of speech*, 31–54. Oxford: Oxford University Press.

Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.

Kálmán C., György, László Kálmán, Ádám Nádasdy & Gábor Prószéky. 1989. A magyar segédigék rendszere. *Általános Nyelvészeti Tanulmányok* 17. 49–103.

Kenesei, István. 2000. Szavak, szófajok, toldalékok [Words, word classes, suffixes]. In Ferenc Kiefer (ed.), *Strukturális magyar nyelvtan [The structure of Hungarian]*, vol. 3: Morfológia [Morphology], 75–136. Budapest: Akadémiai Kiadó.

Kenesei, István. 2006. Szófajok [Word classes]. In Ferenc Kiefer (ed.), *Magyar nyelv [The Hungarian language]*, 80–119. Budapest: Akadémiai Kiadó.
Koerner, E.F.K. 1995. *Professing linguistic historiography.* Amsterdam: John Benjamins. DOI: 10.1075/sihols.79.
Lancelot, Claude & Antoine Arnauld. 1660/1967. *Grammaire générale et raisonnée.* Facsimile edition. Menston: Scolar Press Ltd.
Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. *UPenn Working Papers in Linguistics* 4. 201–225.
Mártonfi, Ferenc. 1973. Vannak-e szófajok a kínai nyelvben? [Are there word classes in Chinese?] *Nyelvtudományi Közlemények* 75. 177–203.
Palmer, Harold E. 1924. *A grammar of spoken English.* Cambridge: Heffer & Sons.
Panagiotidis, Phoevos. 2015. *Categorial features: A generative theory of word class categories.* Cambridge: Cambridge University Press.
Prószéky, Gábor, László Kálmán, György C. Kálmán & Ádám Nádasdy. 1984. Topic, focus and auxiliaries in hungarian. *Groninger Arbeiten zur germanistischen Linguistik* 24. 162–177.
Radford, Andrew. 1976. On the non-discrete nature of the verb-auxiliary distinction in English. *Nottingham Linguistic Circular* 5(2).
Rijkhoff, Jan & Eva van Lier (eds.). 2013. *Flexible word classes: Typological studies of underspecified parts of speech.* Oxford: Oxford University Press.
Robins, Robert H. 1980 [1964]. *General linguistics.* Third edition. London: Longman.
Saussure, Ferdinand de. 1916. *Cours de linguistique générale.* Lausanne: Librairie Payot.
Siptár, Péter. 2006. A fonéma tündöklése és … [The rise of the phoneme and its…] *Magyar Nyelv* 102. 407–420.
Topping, Donald. 1973. *Chamorro reference grammar.* Honolulu: University Press of Hawaii.

Chapter 5

The matrix: Merge and the typology of syntactic categories

Andrea Moro
University School for Advanced Studies IUSS, Pavia

In recent works (Moro 2000; 2009; Chomsky 2013; 2017; Chomsky et al. 2019; Rizzi 2015; 2016) a new type of phrasal structure has been assumed resulting from Merging two XPs where neither XP projects: the unlabelled [XP YP]. This structure stands out as an exception with respect to the typical X^0s and XPs. I will show that by considering some basic properties of Merge in an abstract combinatorial framework the stipulative character of this category is absorbed along with some potential redundancies of UG.

1 The X^0 vs. XP distinction and the lexicon

A basic opposition is manifested in syntax between X^0s and XPs. A traditional way of distinguishing between these two categories is to refer to the lexicon: an X^0 directly comes from the lexicon, whereas an XP does not. In fact, this opposition can also be captured by referring to Merge by reasoning as follows.

2 The matrix or beyond the X^0–XP taxonomy

An X^0 cannot be targeted by Internal Merge (IM) whereas an XP can; call this property "atomicity". Interestingly, this not the only way to cast X^0 and XPs into two disjoint classes by referring to Merge. An X^0 cannot appear as a specifier whereas an XP can. Since a specifier is an XP which is Merged to another XP

Andrea Moro. 2020. The matrix: Merge and the typology of syntactic categories. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 81–92. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280635

without projecting, one can say that an XP is an optional projector whereas an X^0 is not; call this property "incapsulation".[1]

(1) A syntactic entity S is:
   a. atomic ([+a]) iff no parts of it can be targeted by IM.
   b. incapsulable ([+i]) iff it can be merged to an XP without projecting.

Let us now construe a combinatorial square matrix based on these two independent properties displaying both positive and negative polarities and start by representing the two opposite and already recognized entities, namely an X^0 as [+a, −i] and an XP as [−a, +i]:[2]

(2)

	+i	−i
+a		X^0
−a	XP	

This matrix raises a new question, namely whether there exist any [+a, +i] and [−a, −i] syntactic entities, i.e. homopolar syntactic entities, or whether there exist only the heteropolar ones. I will show that the answer is affirmative and this matrix solves the problem raised by unlabeled [XP YP] structures. Let us first consider the case of a syntactic entity with all negative polarity features.

[1]This operation can in principle be reiterated generating "multiple specifiers" or one specifier and multiple adjuncts; I will maintain Kayne's (1994) LCA-based principle according to which there can be only one element merged with a phrase to preserve the possibility of linearization. This is only partially true since there could be multiple subjects provided that only one is spelled-out at phonetic form (PF). The existence of these configurations is provided by inverse copular sentences in Italian. In this case, the preverbal phonologically overt DP is mutually c-commanding *pro* without violating the LCA since *pro* is not visible to linearization. Clear support for this analysis comes from cases where the preverbal subject is singular and the postverbal one plural: in this case, the copula anomalously agrees with the postverbal DP showing that there must be a *pro* (in fact a "null predicate") mediating the agreement relation as in *la causa sono Pietro e Giovanni* (the cause-sing.fem. are Peter and John). The intervening subject is *pro* as proposed in Moro (1997) as in *la causa pro sono io* (the cause pro am I; 'the cause is me') or just *sono io* (am I; 'it's me'). Indeed, if more than one adjunct/subject is generated: all but one must move, as a consequence of the principle of dynamic antisymmetry.

[2]Matrices are typical structuralist tools that have their origin in phonological models. In syntax, they have been used less massively; two major examples are Chomsky's (1970) and Jackendoff (1977) – both incorrectly assuming that noun phrases cannot be predicates – and Muysken & van Riemsdijk 1986 relying on features pertaining to X-bar levels. In fact, perhaps the first use of derivative categories in linguistics can be traced to at least the Hellenistic models of grammar, witness the term "participium" (lit: that takes part) related to a verbal form which displays adjectival morphology.

2.1 Bare small clauses

A natural candidate to occupy the [−a, −i] slot is the so-called "bare small clause" (BSC), prototypically represented by the complement of the copula. Two separate issues must be addressed here: a preliminary one is whether there is any empirical reason to assume that such non-atomic constituents exist; the other is whether there is any empirical reason to exclude them from the specifier position. In fact, they have both already be answered positively. I will just sketchily remind here the data upon which the answer is built.

Originally, the complement of the copula was considered to be the same as the complement of *believe*-type verbs and labelled "small clause" (SC): namely, a non-inflected predicative structure (see Williams 1978 and Stowell 1978 for the first proposals and Graffi 2001 for a critical survey). It has been later proposed that these two types of complements have two distinct structures (see Moro 1997 for the original proposal; and Moro 2017a,b for a synthetic update): the complement of *believe*-type verbs is a phrase headed by a predicational head – whose precise categorical nature is still under discussion – whereas the complement of the copula is an unlabeled phrase resulting from the direct merge of two phrases. The minimality of the latter structure is what justifies the term "bare"; accordingly, these phrases are represented as [XP YP] merged without any intervening head.[3] The specificity of this construction is not the merging of two phrases but rather the fact that *neither* phrase project, unlike the case of specifiers that yield $[_{\alpha}$ XP YP] where the label α coincides with either phrase and the specifier is the phrase which does not project.[4]

The empirical reasons supporting the distinction between SC and BSC are based on several distinct domains. For the sake of simplicity, three distinct types of domains can be reminded here and exemplified in (3): the distribution of predicative markers (3a,b);[5] intervening effects on cliticization, more specifically vio-

[3]This analysis revives Williams's (1980) original proposal for the analysis of SCs which was abandoned partially because of the influential proposal by Chomsky's (1986) to uniform clause structures to the XP format, normalizing all phrases to endocentric structures.

[4]Notice that in this analysis of predicative structures both the subject and the predicated are incapsulated; this independent fact shows that incapsulation is more general than "specifierhood" which is inherently asymmetrical.

[5]The presence of a predicative marker in the complement of *believe*-type verbs was taken by Moro (1988) as the spell-out of an abstract predicative head ($Pred^0$); its absence in copular constructions, instead, led to hypothesis that the clausal constituent was better analyzed as an AgrP and – correspondingly – the copula as the expression of tense (and aspect) features (T^0) yielding a first version of the so-called "Split-Infl" hypothesis. This analysis preceded and was empirically distinct from the influential version proposed by Pollock (1989) and was later partially abandoned in favor of the unheaded BSC hypothesis, while maintaining the idea that IPs were in fact to be analyzed as TPs.

lations of Rizzi's (1990) relativized minimality (3c,d);[6] instability, i.e. the necessity of movement out of the embedded clausal structure both in English (3e–g) and in *pro*-drop languages (3g):[7]

(3) a. Mary considers [John (as) the culprit] (cf. also *John is considered* t *(as) the culprit*)
b. John is [*t* (*as) the culprit]
c. Italian
*lo ritengo [Maria H^0 *t*]
so-CL believe Maria
d. Italian
Maria lo è [*t t*]
e. Mary considers [John stupid]
f. *is [John stupid]
g. Italian
*è [Gianni stupido]
is Gianni stupid

All these facts converge toward the analysis according to which the complement of the copula consist of merging two phrases without the intervention of a head. This analysis has proved to be consistent across languages; a strong support to the existence of BSCs along with SCs comes from Pereltsvaig's analysis of Russian (Pereltsvaig 2007). Moreover, it has also been proposed that BSCs also

[6] I have simplified the representation in (3d): for locality reasons, a BSC can never be completely evacuated (see Moro (1993) elaborating on Rizzi's (1990) notion of head-government. The clitic is rather sub-extracted from a DP as an N^0. The same D^0/N^0 distinction holds for wh-elements where *which* corresponds to D^0 while *what* to N^0, witness cases like *what a party*! where the wh-element co-occurs with an overt D^0; this also explains the possibility to extract *what* but not *which* in existential sentences (see Moro 1997 revising Heim's (1987) semantic account of this contrast and the locality conditions on extraction; see also Moro 1993 for locality issues within a Minimalist framework).

[7] Notice that the *pro*-drop parameter is totally irrelevant here: movement in required in Italian on a par with in English. No "expletive" can rescue the structure where neither phrase moves, not even *ci* (there), reinforcing the hypothesis that movement is required to solve the instability of the lower BSC rather than satisfy some specific condition of the subject position; for the impact of this phenomenon on discharging the extended projection principle see Moro (1997; 2000) and, in particular, Moro (2009) for a detailed discussion involving the role of $Focus^0$ in post-verbal positions.

occur in nominal domains, as complements of P^0 heads playing the same role as the copula in that they provide a landing site for either the subject or the predicative phrase (Moro 2000; see also Kayne 1994; den Dikken 1997; Zamparelli 2000). Simple examples are pairs like *these types of books* vs. *books of this type* which are generated by the same underlying structure containing a BSC, namely [of [$_{BSC}$ [books] [this type]]], by raising either the subject [books] or the predicative nominal [this type] to the specifier of P^0 (cf. *books are of these types*). We can now turn to the second issue, namely as to why BSCs cannot be specifiers.

One of the special properties of BSCs – witness examples like (3f,g) – is that they force movement of either XP: if the two XPs constituting the BSC are both noun phrases then either movement is possible, yielding a canonical vs. inverse copular sentence depending on whether the subject or the predicate raises (and similarly, mutatis mutandis, in nominal constructions); if the predicate of the copular sentence is not a noun phrase – say an adjectival phrase – then the only viable rescue strategy is for the subject to raise, because of the morphological restrictions imposed on the landing site (arguably related to Case assignment). The reason of the instability of this structure is inherently related to the symmetrical nature of this configuration; there are two alternative explanations, one based on the LCA (Moro 2000) – movement is necessary to allow linearization of two mutually c-commanding phrases – the other on labeling algorithm (Moro 2009) – movement is necessary to provide a label to the BSC (see also Moro 2000; 2009; Chomsky 2013; 2017; Chomsky et al. 2019; Rizzi 2015; 2016 for further support to this explanation and in general for the principle of dynamic anti-symmetry). It could well be that both explanations are valid and that this phenomenon reveals a twofold nature of instability depending on the test adopted. Duality is not to be avoided per se in empirical science if it is grounded and impinges on separate empirical reasons.

However, for what matters here, even if only one explanation will turn out to be true, still the instability – hence, the necessity of movement out of a BSC – remains as an undisputed fact. And it is this very fact that offers a straightforward explanation for the second issue addressed in this section, namely as to why BSCs cannot be specifiers. An obvious case study is the impossibility for BSC to be clausal subjects, i.e. specifiers of TP. The crucial fact is that movement is banned from within this position unless some specific conditions are realized which do not apply here (for the locality conditions on the subject position see in particular the discussion in Rizzi 2015, Stepanov 2007 and references cited there). All in all, the impossibility for a BSC to occur as a subject follows for principle reasons

without ad hoc stipulations: on the one hand its instability requires movement; on the other, movement is impossible for locality conditions.[8]

Eventually, the homopolar negative slot [−a, −i] generated by the matrix in (2) can then be filled in by BSCs:

(4)

	+i	−i
+a		X^0
−a	XP	BSC

The matrix, in fact, completely eliminates the stipulative character of BSCs: these acentric phrases are not exceptions as they are now framed in the same two property based grid generating the other two categories, namely words and endocentric phrases. The exception would now rather be if they did *not* exist.

2.2 Expletives

There is a residual empty slot in the matrix in (4), namely the homopolar positive syntactic entity: [+a, +i]. Is there a reason for assuming that there exist atomic entities that can occur as the specifiers of a phrase, that is that can be incapsulated? I would like to suggest that this category exists and coincides with expletives.[9] In a sense, this assumption is trivially proved. Elements like *there* in English existential sentences, for example, are clearly atomic but they cannot further project when merged with a phrase – in fact, they prototypically end up occupying the position canonically reserved to clausal subjects – hence [+i]. Nevertheless, they do qualify as exceptions since atomic entities, i.e. X^0s, do project and they cannot occupy the subject position: expletive appear like "inert heads". One possibility

[8]Interestingly notice the following contrast:

(i) a. * [John the culprit] is strange
b. [for John to be the culprit] is strange

This shows that what prohibits for a clausal structure to be clausal subject is not related to the finiteness of tense and aspects features. As for the possibility of a local movement to a focal position to solve instability (see Moro 2009). Notice also that being BSC [−i] it must project when merged with an XP: this is consistent and in fact it derives the solution to the instability of these constituents as predicted by the principle of dynamic anti-symmetry (see Moro 2000; 2009; Chomsky 2013; 2017; Chomsky et al. 2019; Rizzi 2015; 2016).

[9]I refer to "expletives" in general but a more fine-grained terminology would distinguish between subject-expletives as in *it was clear that John left* and predicative-expletives as in *it's that John left*, just to remain to pro-CPs, along the lines of Moro (1997).

would of course be to assume that expletives are not real heads but rather "monolithic" phrases which exceptionally contain no parts visible to Internal Merge but this would of course be a way just to rephrase the situation. On the other hand, however, the capacity of expletives to share *some* properties with heads can indeed be independently supported, by considering more fine-grained and hidden empirical data, such as those manifested in copular constructions. Consider the following contrast taken from Moro (1997; see also Stepanov 2007 for an analysis of the same data in (5a):[10]

(5) a. which wall do you think there was [a picture of *t*]
b. * which wall do you think the cause of the riot was [a picture of *t*]

Following Moro (1988; 1997), I will assume that *there* is a not a subject expletive which is inserted late in the derivation; this element is rather a pro-predicate expletive raised from a lower position or, equivalently, that existential sentences like (5b) belong to the more general class of inverse copular sentences: cf. [there was [[a picture of the wall] *t*]]. In (5b), instead, the phrasal predicate *the cause of the riot* is raised to the pre-verbal position. The major difference between the two sentences, then, is that the head of the predicate is embedded in (5b) (namely, *cause*) whereas it edges the TP phrase in (5a) (namely, *there*).

This distinction allows to explain this contrast by appealing to the notion of L-marking. More specifically, Moro (1997) adopted the version of L-marking as formulated in Cinque (1990) which differed from Chomsky's (1986) original proposal: Cinque's version is based on the selectional capacities of a head rather than its theta-marking ones. Synthetically, a phrase is an island (or a barrier to movement) unless it enters into a local relationship with a head selecting it, where by "local relationship" a minimal dominance relation is intended canonically expressed in terms of c-command. An interesting remark on L-marking highlights its persistence in Minimalist frameworks: "Though varieties of government would be 'imperfections', to be avoided if possible, the closer-to-primitive notion of L-marking should pass muster, hence also notions of barrier that are

[10]This contrast was also discovered with respect to quantifier raising:

(i) a. there weren't pictures of many girls
b. the cause of every riot wasn't pictures of many girls

The embedded quantifier *many* can have scope over negation, hence be extracted from the subject DP at logical form (LF), only in a *there*-sentence (ia). Notice that the example in (ia) falsifies Williams's (1984) analysis of *there* as a scope marker: for a full discussion, see Moro (1997: Ch. 2).

based on nothing more than L-marking" (Chomsky 2000, 117; for a critical review of the notion of L-marking and the empirical and historical reasons behind it see Roberts 1988).

All in all, the impossibility to extract from within the post-verbal subject in (5b) is immediately explained by the fact that it is not L-marked: the element selecting it is the predicative head *cause* and it fails to c-command it; the only other head c-commanding the subject is the copula: although it qualifies in terms of local configuration, it does not select the subject: thus the subject is not L-marked and extraction from it yields an ungrammatical sentence. This parallels the case of a preverbal subject of an embedded sentence: it is in a proper local configuration with a complementizer c-commanding it but it is not selected by it (see Rizzi 1990; 2015; see also again Stepanov 2007 for critical considerations on extractions from the subject position). In (5a), instead, the head *there* (locally) c-commands the lower subject and it selects it in its capacity as a pro-predicate: thus, the subject is L-marked and extraction is viable. The special head-like relation between the expletive *there* in subject position and the copula is also manifested in the fact that the copula anomalously shows rightward agreement, reasonably a sign that the number features of the subject have been transmitted by the pro-predicative element selecting it:[11]

(6) a. there were many pictures of the wall
b. the cause of the riot was/*were many pictures of the wall

Similar considerations concerning *there* would hold for pre-verbal *it* in quasi-copular sentences such as *it seems that Mary left* as well as in inverse copular sentences with clausal subjects like *it's that Mary left*, whose common structure is: [it V^0 [[that Mary left] *t*]. There are also other occurrences of *there* with other verbs than the copula which would lead to the same conclusion, namely unaccusative constructions but illustrating them here would take us too far (see Moro 1997 and the crucial extensions suggested in the comprehensive theory of argument structure proposed in Hale & Keyser 2002).

[11]That there are cases where the nominal head of a predicate *must* agree with its subject is independently attested in cases like:

(i) I consider John and Peter my best friend*(s)

However, agreement is by no means obligatory in all cases. In fact, there can be a complete mismatch in gender and number as in:

(ii) considero i libri la mia passione
consider-1SG the-M.PL books-M.PL the-F.SG my-F.SG passion-F.SG

See Moro (1988; 1997; 2017a) for further considerations.

Crucially, for what matters here, there is a further piece of evidence in favor of the fact that expletives have a twofold nature. In the previous examples, I have provided evidence that they share the same selectional properties as *heads*; it can be also proved that they do behave like *phrases* by reasoning as follows. expletives are only merged with other phrases; as [+i] elements they cannot project, thus the resulting phrase can either be a full endocentric phrase (where the other element projects) as in [$_{TP}$ Expl TP] or it can be a BSC (where neither phrase projects) as in the [$_{BSC}$ DP Expl] generating (5a) where neither phrase projects. In the latter case, either phrase must be further moved as predicted by dynamic anti-symmetry:[12]

The very existence of atomic and incapsulated syntactic categories (expletives) is ultimately well-grounded empirically and this allows us to fill in the last available slot in the two property based grid:[13]

(7)

	+i	−i
+a	Expl	X^0
−a	XP	BSC

[12]For the reasons why the expletive raises and the impact it has on semantic structure see Moro (1997: Ch. 3; 2000; 2009); Chomsky (2013; 2017); Chomsky et al. (2019); Rizzi (2015; 2016) if the expletive did not have phrasal properties and they were just like heads, it would be hard to explain why the structure is unstable and it requires movement. All in all, expletives appear to share some properties with both X^0 and XPs.

[13]Notice that the BSC analysis originally proposed for existential sentences, quasi-copular sentences, and unaccusative constructions has been extended to cover previously unrelated constructions. In particular, the same analysis has been proposed to include wh-phrases to explain split interrogatives, including the classic "*was-für* split phenomena" and its equivalent in Romance languages (see Moro 2000 and Ott 2012 for a further and original extensions of this proposal). In Italian, for example, we get the following case study where the particle *di* ('of') plays the same role as a nominal copula in *questi tipi di libri* ('these types of books') forcing movement of the wh-element *cosa* ('what') to the specifier position of the proper CP-slot:

(i) Cosa legge [*t* di [$_{BSC}$ libri *t*]]?
what reads-3SG of books
'What books does s/he read?'

For what matters here, examples like (i) show that the twofold nature of elements like *there* is not isolated to canonical expletives: it is rather unexpectedly shared by wh-elements like *cosa* ('what') which constitute an unstable structure with another full phrase, revealing their phrasal nature, but do not contain any part accessible to Internal Merge, i.e. they behave like X^0. We should perhaps speak of "generalized expletives" to include clausal and non-clausal ones.

3 On evaluating the matrix: Suggestions for the future agenda

The fourfold taxonomy generated by the matrix absorbs the exceptionality of BSC and expletives framing them along X^0 and XP in a natural way within the same grid generated by two syntactic properties formulated by referring to Merge.

In principle, this may not be the only welcome result: the matrix could also be exploited to capture further empirical generalizations. For example, it reveals natural classes – i.e. agreement is possible only with a [+i] category – or it allows to identify grammatical functions in a more comprehensive way – i.e. predicative structures coincide with [−a,−i] category (see Moro 2000; 2004 for further discussion) or simplifications – i.e. two homopolar entities (namely, expletives and BSCs) cannot be merged. Whether or not this matrix will be theoretically useful for formulating new questions is left for future research to answer.

Abbreviations

1	first person	LF	logical form
3	third person	M	masculine
BSC	bare small clause	PF	phonetic form
CL	clitic	PL	plural
F	feminine	SC	small clause
IM	Internal Merge	SG	singular
LCA	linear correspondence axiom	UG	Universal Grammar

Acknowledgments

My special thanks go to Robert Frank, Raffaella Zanuttini, Giorgio Graffi, Cristiano Chesi, Andrew Nevins and Alessandra Tomaselli for their illuminating remarks and two anonymous reviewers: the errors remain all mine. I wish I wrote this paper in Italian for if Ian translated it for me it would have become much better.

References

Chomsky, Noam. 1970. Remarks on nominalization. In Roderick A. Jacobs & Peter S. Rosenbaum (eds.), *Readings in English transformational grammar*, 184–221. Waltham, MA: Ginn & Company.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.
Chomsky, Noam. 2017. The language capacity: Architecture and evolution. *Psychonomic Bulletin & Review* 24(1). 200–203. DOI: 10.3758/s13423-016-1078-6.
Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions and challenges. *Catalan Journal of Linguistics*. DOI: 10.5565/rev/catjl.288.
Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, MA: MIT Press.
den Dikken, Marcel. 1997. Predicate inversion in DP. In Artemis Alexiadou & Chris Wilder (eds.), *Possessors, predicates and movement in the determiner phrase*, 177–214. Amsterdam: John Benjamins. DOI: 10.1075/la.22.08dik.
Graffi, Giorgio. 2001. *200 years of syntax: A critical survey*. Amsterdam: John Benjamins. DOI: 10.1075/sihols.98.
Hale, Kenneth L. & Samuel Jay Keyser. 2002. *Prolegomena to a theory of argument structure*. Cambridge, MA: MIT Press.
Heim, Irene. 1987. Where does the definiteness restriction apply? Evidence from the definiteness of variables. In Eric Reuland & Alice ter Meulen (eds.), *The representation of (in)definiteness*, 21–42. Cambridge, MA: MIT Press.
Jackendoff, Ray. 1977. *X-bar syntax: A study of phrase structure*. Cambridge, MA: MIT Press.
Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
Moro, Andrea. 1988. Per una teoria unificata delle frasi copulari. *Rivista di Grammatica Generativa* 13. 81–110.
Moro, Andrea. 1993. Heads as antecedents: A brief history of the ECP. *Lingua e Stile* 28. 31–57.
Moro, Andrea. 1997. Dynamic antisymmetry: Movement as a symmetry-breaking phenomenon. *Studia Linguistica* 51(1). 50–76. DOI: 10.1111/1467-9582.00017.
Moro, Andrea. 2000. *Dynamic antisymmetry*. Cambridge, MA: MIT Press.
Moro, Andrea. 2004. Linear compression as a trigger for movement. In Henk C. van Riemsdijk & Anne Breitbarth (eds.), *Triggers*, 387–430. Berlin: De Gruyter. DOI: 10.1515/9783110197433.
Moro, Andrea. 2009. Rethinking symmetry: A note on labelling and the EPP. In Paola Cotticelli Kurras & Alessandra Tomaselli (eds.), *La grammatica tra*

storia e teoria: Scritti in onore di Giorgio Graffi, 129–131. Alessandria: Edizioni dell'Orso.

Moro, Andrea. 2017a. Copular sentences. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley–Blackwell companion to syntax*, Revised edition, vol. 2, 1210–1232. Hoboken, NJ: Wiley–Blackwell. DOI: 10.1002/9781118358733.wbsyncom043.

Moro, Andrea. 2017b. Existential sentences and expletive *there*. In Martin Everaert & Henk C. van Riemsdijk (eds.), *The Wiley–Blackwell companion to syntax*, Revised edition, vol. 3, 1483–1508. Hoboken, NJ: Wiley–Blackwell. DOI: 10.1002/9781118358733.wbsyncom044.

Muysken, Pieter & Henk C. van Riemsdijk. 1986. Projecting features and featuring projections. In Pieter Muysken & Henk C. van Riemsdijk (eds.), *Projecting features and featuring projections*, 1–30. Dordrecht: Foris.

Ott, Dennis. 2012. *Local instability: Split topicalization and quantifier float in German*. Berlin: De Gruyter.

Pereltsvaig, Asya. 2007. *Copular sentences in Russian*. New York: Springer.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2015. Cartography, criteria, and labeling. In Ur Shlonsky (ed.), *Beyond functional sequence* (The cartography of syntactic structures 10), 314–338. Oxford: Oxford University Press.

Rizzi, Luigi. 2016. Labelling, maximality and the head-phrase distinction. *The Linguistic Review* 33(1). 103–127. DOI: 10.1515/tlr-2015-0016.

Roberts, Ian. 1988. From rules to constraints. *Lingua e stile* 23(3). 445–464.

Stepanov, Arthur. 2007. The end of CED? minimalism and extraction domains. *Syntax* 10(1). 80–126. DOI: 10.1111/j.1467-9612.2007.00094.x.

Stowell, Timothy. 1978. What was there before *there* was there. In Donka Farkas, Wesley M. Jacobsen & Karol W. Todrys (eds.), *CLS 14: Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, 458–471. Chicago: Chicago Linguistic Society.

Williams, Edwin. 1978. Small clauses in English. In John P. Kimball (ed.), *Syntax and semantics*, vol. 4, 249–273. New York: Academic Press.

Williams, Edwin. 1980. Predication. *Linguistic Inquiry* 11(1). 203–238.

Williams, Edwin. 1984. *There*-insertion. *Linguistic Inquiry* 15(1). 131–153.

Zamparelli, Roberto. 2000. *Layers in the determiner phrase*. New York: Garland.

Chapter 6

On a difference between English and Greek and its theoretical significance

George Tsoulas
University of York

This paper offers a comparative study of the coordinator *and* and the comitative preposition *with* in its coordinating function. Greek is shown to behave differently from English in this respect and this is accounted for in terms of labelling potential of a syntactic/lexical object. The more general claims are that labelling is a locus of variation and that labelling is (still) a syntax internal process.

1 Introduction

One of the major proposals concerning the possible *loci* of syntactic variation is the so-called Borer–Chomsky conjecture which Baker (2008) formulates as follows:

> All parameters of variation are attributable to differences in features of particular items (e.g. the functional heads) in the lexicon.

In general, it is a somewhat more restricted version that is more widely accepted, namely that syntactic variation and parametric properties are restricted to properties of inflectional heads only.[1]

In this note, I would like to suggest that the potential of a category to supply a label to a constituent that it *heads* is also a property that, though not strictly inflectional and clearly not restricted to functional heads, is a locus of variation across languages. The empirical argument in favour of this position comes from

[1]This is more in line with both Chomsky's and Borer's formulations.

George Tsoulas. 2020. On a difference between English and Greek and its theoretical significance. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 93–111. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280637

the behaviour of certain coordinated structures in English and Greek (and to a much lesser extent French). It is well known that the preposition *with* in English also functions as a coordinator. The same is true in Greek, but coordinations with *with* pattern differently in the two languages. In a nutshell, while in English the first conjunct must raise out of the *with* phrase, there is no such requirement in Greek.

In this paper I consider more closely these patterns and argue that they are better understood if we extend Chomsky's (2013) proposal on structured coordination with *and* to the case of coordination with *with* and argue, contra Kayne (1994), that movement of the first conjunct is driven not by Case but by the requirements of the labelling process, and more specifically the idea that while some categories may be able to label in some languages they may not in others. Taking Chomsky's idea that some categories may be assigned a feature [LABEL] that nothing can remove more seriously than he probably intended, we can imagine that this feature is an integral part of lexical items. It follows that for categories that lack that feature, the labelling algorithm cannot identify any of their properties for externalisation and the conceptual-intentional system.[2]

The paper is structured as follows: in §2 I present the facts of English concerning *with*-coordinations. §3 develops the account of *with*-coordinations in English in labelling terms. In §4 I turn to the Greek data and show that the patterns follow from the simple proposal that Greek *me* ('with') is a labelling category. I also discuss some interpretive issues relating to distributivity. §5 spells out some consequences of the analysis.

2 Coordination: *and* and *with*

The following paradigm in English is well known:

(1) a. Sue and Sy are friends
b. * Sue is friends and Sy

(2) a. Sue is friends with Sy
b. * Sue with Sy are friends

Examples like those in (2) are found with a variety of symmetric predicates, as Lakoff & Peters (1969) as well as Dong (1970) have discussed (cf. 3), although with varying degrees of acceptability.

[2]This is an important point to which we will return in §5.

(3) a. Sue is co-workers with Sy
b. Sy is mates with Sue
c. Sue is school/bandmates with Sy
d. ? Sy is siblings with Sue
e. Sue is twins with Sy
f. Sy is co-authors with Sue

Compare now (3) with its version where *with* is replaced by *and*.

(4) a. Sue and Sy are co-workers
b. Sue and Sy are mates
c. Sue and Sy are school/bandmates
d. Sue and Sy are siblings
e. Sue and Sy are twins
f. Sy and Sue are co-authors

The main difference between the paradigm in (3) and that in (4) is that with *and*-coordinations the whole constituent remains together while with *with* the first conjunct must move out.

Beyond nominal predicates, as above, the pattern extends to verbal symmetric predicates such as *collide* or *fuck*:

(5) a. Rosetta collided with comet 67P
b. Rosetta and comet 67P collided
c. * Rosetta with comet 67P collided
d. * Rosetta collided and comet 67P
e. Sue fucks with Sy every Wednesday evening
f. * Sue with Sy fuck every Wednesday evening
g. Sue and Sy fuck every Wednesday evening
h. * Sue fucks and Sy every Wednesday evening

Lakoff & Peters 1969 suggested first that the preposition *with* was functioning here as a coordinator and, moreover, the *and*- and *with*-coordinations were related and should be transformationally linked through a process of replacing *and* by *with* and extraposing *with NP*. The issue of the relatedness of the two constructions as well as the basis for Lakoff & Peters's (1969) account was revisited, in light of the LCA, by Kayne 1994: §6.3, who proposed that the reason for the commonalities between (1a) and (2a) is that they both derive from the same underlying structure, namely (6).

(6) [DP1 [[and/with] DP2]]

What sets the two constructions apart, for Kayne, is that there is a requirement for the first conjunct to move out of the conjoined phrase in (2a) because it cannot be adequately Case licensed in situ. More specifically, while a phrase coordinated with *and* allows both conjuncts to be Case licensed by virtue of the fact that the whole coordinated constituent is in a Case-licensing position, this is not true of coordinated phrases with *with*. A somewhat different way of putting this restriction is that, from a Case theoretic point of view, DP coordination is only licit if Case can be distributed to both conjuncts. In the case of *and* this appears to be so. In the case of *with*, however, this does not happen because the second conjunct is case licensed by *with* while the first one has to get Case from an external source.

The latter way of putting the relevant constraints can be made to work further, in the sense that a constituent of the type *A and B* does distribute like its conjuncts whereas a constituent like *A with B* does not. But again, if we assume that the construction is headed by the coordinator, we would have to suggest that in the case of *with* it is still a Case assigning preposition rather than a coordinator, which in turn casts doubt on the analysis of these two constructions as deriving from identical underlying structures. Moreover, under this analysis it is not clear why with different predicates it is impossible to extract the first conjunct of a *with* coordination:

(7) * Sue is French with Sy.

For this, Kayne suggests that in order to obtain a distributive reading a coordinated phrase must be preceded by a distributor which may be overt or covert. This distributor, noted BOTH following Kayne's convention, forces the distributive reading on the coordinated phrase, which is, of course equivalent to a sentential coordination.

(8) BOTH [John and Mary] love cats → John loves cats AND Mary loves cats.

And, of course, these cases are also fine with an overt distributor:

(9) Both John and Mary love cats.

In the case of *with*-coordinations, however, the distributor induces a barrier to the movement of the first conjunct. Thus, sentences with the following representation are out.

(10) (Kayne 1994: 66, example 56)
John$_i$ is human beings [BOTH [[e$_i$] with Bill]]

But it is unclear why this should be so. After all *both*, as a floating quantifier, does not induce a barrier to the movement of its complement (cf. Sportiche 1988). Equally, a modifying adjunct usually does not induce a barrier to movement of the specifier of the category to which it attaches. I will set aside the issues relating to interpretation and distributivity and revisit them briefly in §4.1.

As we can see, Kayne's analysis is problematic in various respects, and yet, it remains both plausible and attractive. In the following sections I will claim that the basic insights can be maintained and find more elegant and general expression in terms of the labelling requirements and possibilities in these structures.

3 Labelling and coordination

Chomsky (2013) puts forward a particular proposal regarding structured coordination (with *and*), according to which coordinate structures start as (11):

(11) $[_{\alpha}$ and $[_{\beta}$ DP$_1$ DP$_2$]]

As β cannot be labelled because configurations of the type [XP YP] are problematic for the labelling algorithm (both heads are equally prominent), one of DP$_1$ or DP$_2$ must raise (say DP$_1$) and β receives the label of DP$_2$. Importantly, however, α receives the label of DP$_1$, reflecting the fact that the distribution of these coordinated structures is determined by the shared label of the two coordinated elements. As Chomsky notes, though, the construction remains headed by the conjunction which remains visible in order to determine the structure but is not available as a label. This entails that the whole constituent can be the target for movement yielding (12) as an instance of DP movement:[3]

(12) [$_{\text{DP}}$ Peter and Susan] are [$_{\text{DP}}$ ~~Peter and Susan~~] teachers

Assuming this to be on the right track, let us turn to the case of *with*-coordinations. Given that (13), modelled on (12) is ungrammatical, it is clear that this proposal will not be applicable to *with*-coordinations.

[3] To be sure, there are various questions surrounding Chomsky's proposal on coordination. For example, it is unclear what it means for the construction to headed by the coordinator, which determines structure but does not supply a label. This requires further clarification on the assumption that the labelling algorithm identifies heads. We set this aside for now.

(13) * [$_{DP}$ Peter with Susan] are [$_{\sout{DP}}$ ~~Peter with Susan~~] teachers

In these cases the distribution of the coordinate structure does not reflect the distribution of their shared label (DP); in fact, it does not constitute a well-formed constituent at all, as the data show. It follows that the derivation will also be somewhat different. Keeping, however, as close as possible to the proposal on *and* will allow us to pinpoint the difference. The following is a reasonable approximation of their derivation that preserves full parallelism between the *and* and the *with* case. Let us assume that DP_1 and DP_2 merge again like before yielding an unlabellable [XP YP] structure. Next, *with* merges with that syntactic object just like in the case of *and*. The difference, I claim, is that unlike *and*, *with* can provide a label for the resulting object, and we have the following configuration:

(14) [$_{withP}$ with [$_{\alpha}$ DP_1 DP2]]

At this point, DP_1 must raise so that α receives the label of DP_2, yielding (15):

(15) [$_{\beta}$ DP_1 [$_{withP}$ with [$_{DP_2}$ ~~DP_T~~ DP2]]]]

Of course, the question that arises now is what label will β receive. As the two elements of β are [DP_1 withP] we are in the same situation as before where we have a [XP YP] configuration and one of the two elements must raise. DP_1 does and following merging of further material we obtain the initial contrast repeated here:

(16) a. Sue is friends with Sy
b. Sue and Sy are friends

If this is correct it is not Case but the requirement for the whole constituent to be labelled that is responsible for the movement of the first conjunct. The lack of label also accounts for the fact that the whole constituent cannot be targeted for movement, yielding the ungrammaticality of (2b). Whether the constituent remains unlabelled is an important question that we will pick up in §5.

Although this analysis provides an account of the basic patterns, the ungrammaticality of (7) remains problematic. Within the analysis presented here, a covert distributor will not do the job – both because assuming that it induces a barrier to movement is not an idea that is easy to implement in the general framework I am assuming, but also because, in fact, even in cases like (2a) the reading is *distributive* in the sense that the following is a contradiction:

(17) # Sue is friends with Sy but Sy is not friends with Sue.

With a predicate like *being French*, however, this reading is not possible. Furthermore, the distributive reading is not really what matters, but rather the symmetric/reciprocal one. Thus, observe the following contrast:

(18) a. Both Sebastien and Julie are French
b. * Both Sebastien and Julie are friends

With verbal predicates the contrast is perhaps even more telling:

(19) a. Both Sue and Sy fucked (every/on Wednesday evening)
b. Both Rosetta and Galileo collided *(with comet 67P)

Clearly what is missing in the meanings of the examples above is this reciprocal/symmetrical meaning. There is no suggestion that Sue and Sy fucked (with) each other or that Rosetta and Galileo collided with each other. Of course, with an overt reciprocal the sentences are perfect:

(20) a. Sue and Sy fucked each other
b. Rosetta and Galileo collided with each other

The sentences become significantly degraded by the addition of an overt distributor:

(21) a. ???/* Both Sue and Sy fucked (with) each other
b. ???/* Both Rosetta and Galileo collided with each other

One way to extend the account presented here is to focus on the fact that while *and* and *with* appear to perform the same function and give rise to the same structures, it is also not true that they are synonymous.[4] Specifically, I assume that *with* even as a coordinator retains its comitative meaning and θ licenses its DP complement (DP_2 in our examples). We can then ask how is DP_1 θ-licensed.[5] I propose here that a derivation involving a *with*-coordination will converge only if both coordinated DPs can be independently θ licensed.[6] This means that they will work only with two-place predicates, either verbal (like *collide*, *fuck*, *dance*), in which case the DP will receive a thematic role in the subject position, or with symmetric relational nouns like *friends, co-workers* and so on where the thematic role will be available in the nominal extended projection.[7] The idea, therefore, is

[4] In §4.1 I revisit this issue and propose that even if we stick with distributivity, the results will come out right if we look more closely at the morphology of distributivity.

[5] This is a legitimate question even if we have a coordination where we generally assume that θ licensing involves the whole constituent. The distribution of Case inside the *with*-coordination also does not work in the same way.

[6] Again, in parallel with Case.

[7] The actual mechanism is not relevant here.

that, unless the DP that moves out in order to allow the [DP withP] constituent to be labelled can be thematically licensed in its derived position, the sentence will be ungrammatical, not as a result of lack of Case (Case can be assigned) or of lack of label, but as a violation of the θ-criterion. Labelling is important, however, as it is the label that allows thematic licensing in the case of *and*-coordinations and prevents it in the cases of *with*, with the results that we saw earlier. As noted earlier, there is lexical variation in the range of elements that allow the patterns involving *with*-coordination. So, while with a relational, symmetric noun like *friends* it works fine, with others speakers find it less acceptable at first. Interestingly, with a noun like *enemy* which allows for a non-symmetrical reading the *with* coordination is possible only in the symmetrical reading:[8]

(22) She is mortal enemies with John

Assuming now this analysis, I turn to the corresponding Greek facts.

4 Greek

And-coordinations in Greek show a behaviour similar to that of their English counterparts in the relevant respects, witness (23–24):

(23) Greek
O Kiriakos ke o Aris ine fili.
The Kiriakos and the Aris are friends
'Kiriakos and Aris are friends.'

(24) Greek
*O Kiriakos ine fili ke o Aris.
The Kiriakos are friends and the Aris
'Kiriakos and Aris are friends.'

Greek *me* 'with' also functions as a coordinator, as in (25–26):

(25) Greek
O Kostas me ton Ari ine fili.
The Kostas with the Ari are friends
'Kostas and Aris are friends.'

[8]Example (22) is taken from http://www.davidagler.com/teaching/criticalthinking/handouts/Handout3_AdHominemFallacy.pdf.

(26) Greek
O kostas ine filos me ton Ari.
The Kostas is friend with the Aris
'Kostas is friends with Aris.'

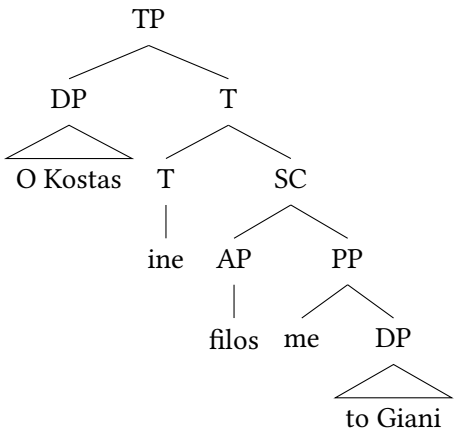
At first sight, taking Greek and English to be basically the same, it looks like in Greek the first conjunct may remain in situ. From a Case theoretic perspective this is somewhat problematic. One would wonder why the same mechanism is not available in English. One approach could suggest that while we may unify Greek and English in terms of Case assignment in these constructions, the EPP requirement of C–T must be satisfied by DP movement in English while in Greek V-to-T suffices. This is a reasonable approach but raises the question why is it impossible to raise the whole withP to [spec T]. The labelling account developed here provides an explanation for that. However, this question may be moot, at least in part, given the evidence on agreement to which we now turn. There are some differences between *with* and *me*. Consider the following:

(27) Greek
*O Kostas ine fili me ton Ari.
The Kostas is friends with the Aris
'Kostas is friends with Aris.'

(28) Greek
*Ego ime fili me ton Ari.
I am friends with the Ari
'I am friends with Aris.'

The agreement contrast between (25) and (26) on the one hand and (27) and (28) on the other is interesting when compared to the agreement found in the English *friends with* construction. In the Greek case, plural agreement on the predicate nominal is only triggered when the first conjunct of the [A with B] element stays in situ. If, however, the first conjunct raises to [Spec T], then agreement is in the singular both on the copula in T and the predicate nominal. Compare this to the English *friends with* construction (2a) where the predicate nominal shows plural agreement but T bears singular features (from agreement with the subject). Now, given that the plural on the predicate nominal is pretty much the only tangible evidence we can lay our hands on in favour of the idea that the underlying structure involves a coordination, we can take the absence of plural agreement (together with the absence of any other factor that blocks plural agreement) as evidence that there is no underlying coordination in Greek, and the right analysis of (26) is roughly (29):

(29)



Friends with construction is not available in Greek. Under a Case theoretic approach, this is problematic given that *me* assigns Case to its complement DP while DP_1 has its Case valued externally. So even pursuing that path one would have to find out why Greek allows this type of Case valuation in cases that look otherwise equivalent.

Given the discussion above and the agreement facts, it is, I suggest, reasonable to propose that the difference between Greek and English regarding *with*-coordinations should be located in the labelling potential of *with/me*.

In the previous section we saw that in English *with* was different from *and* in that it could supply a label. I want now to propose that in Greek *me* is exactly the same as *ke* 'and' in terms of labelling potential,[9] i.e. neither can supply a label (in other words nether carries the feature [LABEL]), and, as a result, it is not surprising that the behaviour of *me*-coordinations in Greek is similar to that of *and*-coordinations (in Greek and English). Assuming this, the patterns follow.

Consider first the fact that the whole constituent will be labelled DP and as a result can be targeted for EPP driven movement and for Case valuation. Concerning Case, as we saw above, *me* will Case license DP_2 while DP_1 will have its Case valued via Agree with T. The following examples show that the whole DP can appear preverbally in subject position with different nominal or prepositional predicates:

(30) Greek
Ego me ton patera mu imaste sinehia se sigrusi.
I with the father mine are always in collision
'I am always fighting with my father.'

[9]They are different in other ways, see §4.1.

(31) Greek
Ego me ton Kosta imaste aderfia.
I with the Kostas are siblings
'Kostas and I are siblings.'

(32) Greek
Ego me ton Apostoli imaste panda antipali.
I with the Apostolis are always rivals
'Apostolis and I are always rivals.'

Assuming further that in some way coordinated phrases are marked as formally plural, agreement both with the predicate nominal and T is expected to be in the plural. This prediction is borne out.

Furthermore, we predict that these coordinated structures will be available with a wide variety of verbal predicates too; in other words, not just with the symmetric ones with which they co-occur in English. Again the prediction is borne out as the following examples show:[10]

(33) Greek
O tragudistis me ti sizigo tu tu ehun megali adinamia.
The singer with the spouse his to-him have great weakness
'The singer and his wife have a weak spot for him.'

(34) Greek
O Kostas me ti Marina, pu ehun molis padrefti, benun mesa sto
The Kostas with the Marina, who have just married, enter in the
saloni.
living-room
'Kostas and Marina, who just got married, enter the living room.'

(35) Greek
O Nikos me ti Maria ehun dio pedia.
The Nikos with the Maria have two children
'Nikos and Maria have two children.'

(36) Greek
O Sakis me ti Frini apoktisan pedi.
The Sakis with the Frini obtained child
'Sakis and Frini had a child.'

[10]The examples (33–38) were found with a simple Google search.

(37) Greek
O Panagiotis me ti Hrisa ehun anagagi to kreopolio tus se
The Panagiotis with the Hrisa have elevated the butcher's theirs to
horo sinathrisis.
space rally
'Panagiotis and Hrisa have turned their butcher's shop to a major gathering place.'

(38) Greek
O Grigoris me ton Petro kserun pos tha se odigisoun.
The Grigoris with the Petros know how will you drive
'Grigoris and Petros know how to drive you around.'

(39) Greek
Telika i Rihana me to Saudarava ine mazi edo ke mines.
Finally the Rihana with the Saudi are together here and months
'In the end Rihana and the Saudi man have been together for months.'

The interpretation of these examples is dependent on the predicate; if the predicate allows for a symmetric reading like (34), where if A is married to B then B is also married to A, then this is what we obtain. If the predicate allows or requires a group reading, like (37–38), this what we get. And finally, if the predicate allows or requires a distributive reading, like (33) or one reading of (35) this is again what we have.

Under the simple proposal that *me* is a non-labelling head the data above are all expected. Let me now turn to a somewhat complicating factor, namely distributivity.

4.1 A complication: Distributivity

There seems to be one significant difference between *ke* and *me* in Greek. It is well known that in Greek, like in French, the coordinator can appear in front of both coordinated constituents:

(40) French
Pierre connaît et Isabelle et Marie.
Pierre knows and Isabelle and Marie
'Pierre knows both Isabelle and Marie.'

(41) Greek
O Kostas gnorizi ke ti Maria ke tin Eleni.
The Kostas knows and the Maria and the Eleni
'Kostas knows both Maria and Eleni.'

Kayne (1994: 146, fn. 16) for French and Chatzikyriakidis et al. (2015) for Greek have argued that the initial (outer) occurrence of the coordinator is in fact a distributive operator. Although this is generally true in the sense that the initial *ke/et* yields a distributive reading it is also true that this is only the case when the second (inner) coordinator is *and/ke/et*. Thus, in Greek, with a *me*-coordination no distributive readings are induced by the presence of an initial *ke*, compare:

(42) Greek
Ke o Sakis ke i Sula sikosan ena trapezi.
And the Sakis and the Sula lifted a table
'Both Sakis and Sula lifted a table.' DISTRIBUTIVE

(43) Greek
Ke o Sakis me ti Sula sikosan ena trapezi.
And the Sakis with the Sula sikosan ena trapezi
'Sakis and Sula lifted a table.' COLLECTIVE

Now perhaps it is the comitative meaning of *me* (which was suggested in §3 for English and is presumably also valid for Greek) that somehow blocks the distributive reading. One way of putting this is to suggest that, semantically, the output of a *me*-coordination is a *group* individual, acting in part as an atom, whereas this is not necessary for *ke*-coordinations, whose semantic value may be that of a *group* (in which case there is no difference with *me*) but can also be an individual of type SUM, which would be an appropriate argument for the distributive operator. However, examples like (44) seem to suggest otherwise, in the sense that, as things stand, there is no immediate suggestion that the two teams form a group in a relevant sense:[11]

(44) Greek
O Olimpiakos me ton Panathinaiko kserun pia apenandi se pies
The Olimpiakos and the Panathinaikos know at-last against to which
omades tha agonistun.
teams will play
'Olimpiakos and Panathinaikos have at last found out which teams they will face.'

[11]This is perhaps too strong. The two teams might form a group in the sense that they are the two Greek teams in the relevant international championship. I will set this aside for this paper.

The reading of (44) is distributive in the sense that it corresponds to a sentential conjunction (45):

(45) Olympiakos knows which team it will face AND Panathinaikos knows which team it will face.

Now adding an initial *ke* to (44) does not have the desired effect:

(46) Greek
Ke o Olimpiakos me ton Panathinaiko kserun pia apenandi se
And the Olimpiakos with the Panathinaikos know at-last against to
pies omades tha agonistun.
which teams will play
'Olimpiakos and Panathinaikos **ALSO** have at last found out which teams they will face (as well as some other group of teams).'

In this case the reading is that of the additive *ke*.[12]

Another issue with the idea that the initial *ke* is the distributive operator applying to an argument of SUM type is that *ke*, *qua* distributive operator, is not available with plurals, which are routinely thought of as carrying the type of sums (Link 2002 and many more after him). Interestingly this is not true for English *both*:[13]

(47) Greek
Ke ta pedia efagan gemista.
And the children ate gemista
'The children TOO ate gemista.'

[12]For more details on the additive *ke*, see Chatzikyriakidis et al. (2015) and references therein.

[13]In French the relevant sentences are altogether ungrammatical so we will not pursue the comparison further although the question why the distributive *et* cannot appear with plurals in any position is an intriguing one:

(i) French
*Et les enfants ont soulevé une table
and the children have lifted a table
'The children have lifted a table.' (intended: each)

(ii) French
*Jean connaît et les enfants
Jean knows and the children
intended: 'Jean knows each child.'

(48) Both children ate gemista

Again the *ke* on (47) is the additive *ke* and does not give the desired distributive reading, unlike what we see in (48).

Setting aside this concern, these patterns can be understood in two ways which probably boil down to the same insight. On the one hand, as suggested earlier, we can think of inner *and/ke/et* as sum forming operators and outer *ke/et* as distributors acting upon these sums. In contrast *with/me* are group forming operators whose outcome behaves in the relevant respects as an atom and therefore the distributor cannot act on them in the same way. This would mean that the reason why initial *ke* followed by a *with* coordination can only be read as additive falls together with (49):

(49) Greek
Ke i epitropi apofasise tin isvoli stin Amorgo.
And the committee decided the invasion to-the Amorgos
'The committee (as well as some other organisation) decided the invasion of Amorgos.'

The alternative way of analysing these patterns is to suggest that the distributive operator is in fact the discontinuous morpheme:

(50) a. Both … and
b. Ke … ke
c. Et … et

Again this idea predicts that adding *both* or *ke* in front of a *with/me*-coordination will not yield a distributive reading simply because, at least in these cases, it is just not the right morpheme for the intended meaning. I think that in this way the ungrammaticality of Kayne's example (10), repeated here, is explained too:

(51) $John_i$ is human beings [BOTH [[e_i] with Bill]]

While Kayne is right that distributivity is the key to understanding the judgement, it is not because a covert BOTH blocks the extraction. Rather, it is because the distributive reading does not arise in these cases because the lexical material is just not right.

5 Some consequences

Let us take stock. I argued so far in this paper that a number of differences in the syntax of coordination both within and across languages can be understood in terms of the labelling potential of different categories and the labelling algorithm. The account developed here raises a number of questions primarily about the role of labels in syntactic derivations.

A particular point of debate regarding labelling going back to the early days of minimalism is whether labels are mere tags onto pieces of structure serving to identify them as a potential targets for operations such as internal Merge or AGREE at least,[14] or active drivers of the derivation. Chomsky (1993; 1995) took the former view. A different view was taken by Adger & Tsoulas 1999, who proposed that labels are complex and include category determining features from both merged elements, i.e. Merge(α, β) $\rightarrow$ [$_{\{\alpha,\beta\}}$ α, β]. Crucially, the label {α,β} was taken to be semi-uninterpretable in the sense that one of the two categorial features that make it up (α and β) had to be eliminated. Eliminating that feature was done in the standard way, by seeking a goal in the numeration or the sub-array, agreeing, and merging it with the existing structure or, by internal merge, raising an element with the required specification. In that proposal, computation was driven by the labels, whether on heads or intermediate projections. Although Chomsky's recent proposals on labelling and the one from Adger & Tsoulas (1999) differ in many respects, they converge on the idea that determining the label of a particular part of the structure is a driving force for computation and that in principle labelling need not obey endocentricity. They diverge on two important conceptual points, namely (a) whether the output of merge needs to be always labelled, and (b) what are labels required for. Regarding the former, Chomsky (2015: 6) is particularly clear on this point:

> Crucially, LA does not yield a new category as has been assumed in PSG and its various descendants, including X′ theory. Under LA, there is no structure [$_{\alpha}$ X], where α is the label of X. LA simply determines a property of X for externalization and CI. It is therefore advisable to abandon the familiar tree notations, which are now misleading. Thus in the description of an [XP, [YP, ZP]] structure, there is no node above either of the two merged constituents. There is no label for the root of the branching nodes.

Taking this at face value, it means that not every output of merge operations will be labelled. A question we might ask about this approach is what happens

[14] The question of external merge is also relevant in terms of the elements that are identified for Merge.

to elements such as [α, β] when LA has not identified a property for externalisation and CI. The issue is puzzling. Imagine that there is some element X for which the Labelling algorithm as identified no property (I suppose that this would be its label) for externalisation and CI. What would that actually mean? In terms of externalisation it would mean that the element would not be pronounced. This is the reasonable understanding of the idea (from Chomsky (2015) that copies do not label. In other words the algorithm will identify no property of copies relevant to externalisation. *Wanna* contraction aside, this seems correct. But what of CI? Would one expect that such an element would be invisible also to the interpretive mechanisms? This seems problematic. Focusing on the cases of interest in this paper, both *and/ke-* and (in Greek at least) *me*-coordinations would be such that the coordinator would provide no relevant property for externalisation and CI. If the reasoning based on copies is on the right track, then the non-labelling nature of the coordinators is a clear counterexample (they are after all externalised). But setting externalisation aside, in the case of CI it is unclear, in this case, how a structure [DP_1 and DP_2] would be interpreted. What does seem clear is that it is a property of the conjunction that is preeminent in the interpretation, namely whatever it is that turns that constituent into a plural (sum) entity. Assume for concreteness that the semantics for DP conjunction corresponds to set formation, or more precisely set-product formation, defined in its general form as follows (Heycock & Zamparelli 2005: 241):

(52) Set product (SP)
$\text{SP}(S^1, \ldots, S^n) =_{\text{def}} \{X : X = A^1 \cup \cdots \cup A^n, A^1 \in S^1, \ldots, A^n \in S^n\}$

The way this works is by taking one element from the denotation of each of the two conjoined elements and yielding their union for all elements of these sets. This is the property that is relevant to CI, rather than the DP label that, as we saw, is assigned by the labelling algorithm. The DP label (or at the very least the lack of label deriving from the conjunction), however, is precisely what accounts for the syntactic patterns. Thus, if the reasoning is correct, we are led to rethink the labelling process as follows: labels in part drive syntactic computation but in crucial respects do not represent properties for CI and externalisation. There is a mismatch between the label relevant to the derivation itself and the CI/semantically relevant one. Labels are necessary and the labelling algorithm is a tool that affords insightful understandings of syntactic patterns, but labels do not determine interface interpretation and do not reflect interface properties. Often in fact, as in the cases analysed in this paper, the syntactic label is at odds with the semantically relevant one.

6 Conclusion

In this paper I tried to rethink the properties of two types of coordination in English and Greek. I argued that the different behaviour of *and* and *with*-coordination in English are the result of the fact that while *and* does not provide a syntactic label *with* does. In Greek, however, neither did, resulting in different behaviours. If I am correct we probably also have to accept two higher level conclusions. First, that the (non)-labelling nature of a category can capture linguistic variation and perhaps is a parametric property. Given that this is not an inflectional category, if I am correct, then there is evidence for variation that, although ultimately located in the lexicon if we assume that there is a feature [LABEL], concerns the only thing that is determined internally to the computational system. The second conclusion, connected directly to the first, is that labelling is a process necessary for the syntactic computation and is neither determined by nor determines interface properties.

Abbreviations

EPP	extended projection principle	LCA	linear correspondence axiom
LA	labelling algorithm		

Acknowledgements

A number of people have very patiently discussed with me the material in this paper, have shared very generously their judgements, and have occasionally stopped me from making some important mistakes. In alphabetical order, I want to thank: Kook-Hee Gil, Nino Grillo, Ekali Kostopoulos, Margarita Makri, Dimitris Michelioudakis, Gillian Ramchand, Peter Sells, Hanna de Vries, Rebecca Woods, and Norman Yeo. Unless the mistakes they prevented were not mistakes then they are not responsible for any shortcomings. Theresa Biberauer's combination of encouragement and understanding has been more instrumental to the completion of this paper than anything else (but she should not be blamed for it). I am extremely happy to offer this to Ian on his birthday and raise *n* glasses to many happy rethinks.

References

Adger, David & George Tsoulas. 1999. The syntax and semantics of Merge. In Patrick Sauzet (ed.), *Proceedings of the third Langues et Grammaire*, 5–20. Departement des Sciences du Langage, Université Paris 8.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–373. Amsterdam: John Benjamins. DOI: 10.1075/la.132.16bak.

Chatzikyriakidis, Stergios, Dimitris Michelioudakis & Giorgos Spathas. 2015. Greek focus operators and their associates. In G. Papanastasiou, D. Koutsogiannis & M Theodoropoulou (eds.), *Studies in Greek linguistics*, vol. 35, 167–179. Thessaloniki: Institute of Modern Greek Studies, Manolis Triandaphyllidis Foundation.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth L. Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.

Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins. DOI: 10.1075/la.223.01cho.

Dong, Quang Phuc. 1970. A note on conjoined noun phrases. *Conneries Linguistiques* 17. Reprinted in: *Studies out in left field: Studies presented to James D. McCawley on the occasion of his 33rd or 34th birthday*, ed. by Arnold M. Zwicky et al., 11-18. Edmonton: Linguistic Research Inc., 1971., 94–101.

Heycock, Caroline & Roberto Zamparelli. 2005. Friends and colleagues: Plurality, coordination, and the structure of DP. *Natural Language Semantics* 13(3). 201–270. DOI: 10.1007/s11050-004-2442-z.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Lakoff, George & Stanley Peters. 1969. Phrasal conjunction and symmetric predicates. In David A. Reibel & Sanford A. Schane (eds.), *Modern studies in English: Readings in transformational grammar*, 113–142. Englewood Cliffs, NJ: Prentice-Hall.

Link, Godehard. 2002. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In Paul Portner & Barbara H. Partee (eds.), *Formal semantics: The essential readings*, 127–146. Oxford: Blackwell.

Sportiche, Dominique. 1988. A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19(3). 425–449.

Chapter 7

Rethinking linearization

Kyle Johnson
University of Massachusetts, Amherst

The reason "movement" is used to describe the relationship between an interrogative phrase in English and the syntactic position is binds a variable in, is because that variable is silent. Impressionistically, the interrogative phrase has changed location – it has moved from the position interpreted as a variable. To derive this feature of the relationship while maintaining a semantics that correctly captures the nature of the variable is not trivial. The presently best model is one that claims that the interrogative phrase is, at least partially, in both positions – the position it is spoken in and the position the variable is in. Jairo Nunes has suggested a method of using that model and an algorithm that converts syntactic representations into strings – a linearization algorithm – to derive the fact that a change of location is how being in two positions is manifest. I develop this idea in a framework that expresses the "be in two positions" syntax with phrase markers that allow a term to be dominated by more than one mother. This interpretation of movement does not fit well with the execution Jairo Nunes had of his idea. I develop an alternative implementation that preserves his leading idea.

1 Introduction

In a series of papers, a book, and a dissertation, Jairo Nunes (1995; 1996; 1999; 2004) has provided a compelling way of deriving a signature property of movement, a property I will call *terseness*.

(1) Terseness
When a term is moved from one position to another, it gets spoken in only one of those positions.

Kyle Johnson. 2020. Rethinking linearization. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 113–135. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280639

There are exceptions to terseness, and some of these Nunes' account predicts. This venue doesn't provide the space to consider these exceptions, or how they fit Nunes' project, so I will set them aside and concentrate on the normal case, in which terseness holds. Nunes' leading idea is that movement creates a structure that the linearization algorithm can interpret only if terseness holds.

Nunes' account has two parts. First, he adopts the copy theory of movement (2).

(2) Copy theory of movement
   a. From a term X is made a copy: X′
   b. X′ is merged into a position higher than X

On this view, movement could take the structure in Figure 7.1, form a copy of *which flower* and form the structure in Figure 7.2.[1]

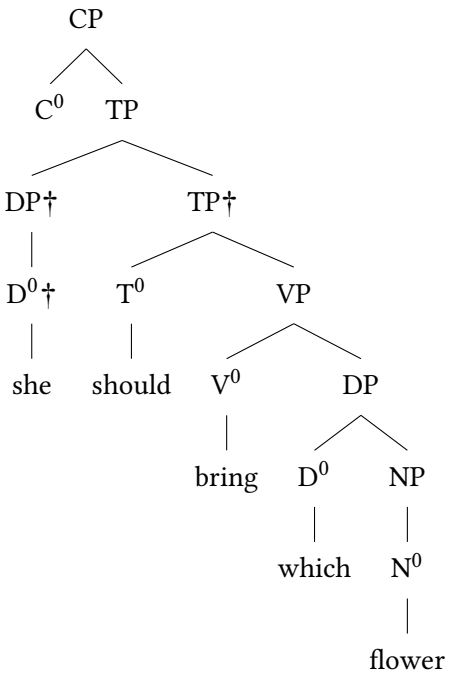


Figure 7.1: Pre-move structure

[1]In order to focus on just one movement operation at a time, I will only consider cases of embedded constituent questions, where movement of the T^0 to C^0 doesn't occur.

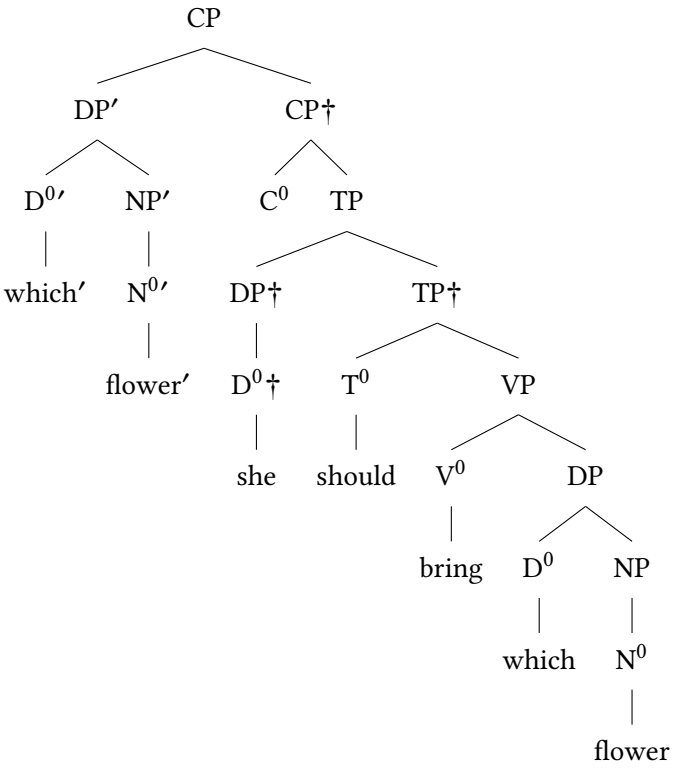


Figure 7.2: Post-move structure

The second part relies on a standard condition on how phrase markers are linearized into strings that Kayne 1994 calls *antisymmetry*.

(3) Antisymmetry
A linearization cannot contain both $a < b$ and $b < a$.

Antisymmetry assumes that a linearization is a set of ordered pairs $x < y$, where x and y are words and "<" is the precedence relation. Antisymmetry simply states that no word can both follow and precede another. Nunes' second proposal, then, is that antisymmetry cannot distinguish one word from. The structure in Figure 7.2 is not pronounced with two instances of *which* and *flower* because a linearization that contains both *which′*<*she* and *she*<*which* will be a violation of antisymmetry. This is terseness.

One goal of this paper is to define copies so that they have the effect of invoking antisymmetry in the way that Nunes envisions. That definition will use the idea broached in Engdahl 1980 that a moved term is a term in two syntactic positions.[2] This can be represented by letting phrase marker trees allow multidominance. Another goal of this paper is to devise a linearization algorithm that can handle such trees.

2 Nunes' proposal

Nunes couches his idea with a slightly modified version of the linearization algorithm in Kayne 1994. The key departure from Kayne's algorithm concerns the items that are linearized. Kayne's algorithm linearizes morphemes – including subword material – and Nunes' doesn't. I'll adopt Nunes' view, which is useful in accounting for certain exceptions to terseness. A goal of Kayne's work is to derive (4) from the linearization algorithm.

(4) If XP asymmetrically c-commands YP, then the words dominated by XP (= d(XP)) will precede the words dominated by YP (= d(YP)) (modulo the effects of movement).

(5) α c-commands β if every phrase dominating α dominates β too, and α doesn't dominate β. α asymmetrically c-commands β if α c-commands β and β doesn't c-command α.

This is achieved by building (4) into the linearization algorithm along the lines of (6).

(6) a. Let L be the set of pairs of heads and phrases, $\langle A, \mathrm{B}\rangle$, in a phrase marker P such that A asymmetrically c-commands B.

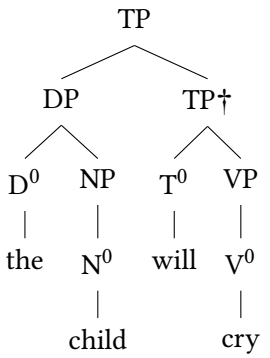
b. The linearization of P is the union of $\mathrm{d}(A) < \mathrm{d}(B)$ for every $\langle A, B\rangle$ in L.[3]

As Kayne notes, (6) needs to be weakened if it is to work for phrase markers that have specifiers. To see this, consider how (6) applies to (7).

[2] Engdahl cites the unpublished Peters & Ritchie (1981) as her source for the idea.

[3] More explicitly: the linearization of P is $\{a < b$: $\forall a \in \mathrm{d}(A)$ and $\forall b \in \mathrm{d}(B)$ if $\langle A, B\rangle$ is in L of P$\}$. Note that "<" is the precedes relation.

(7)



The *L* for (7) is (8a), and this produces the linearization in (8b).

(8) a. $L = \{\langle D^0, N^0\rangle, \langle DP, T^0\rangle, \langle DP, VP\rangle, \langle DP, V^0\rangle, \langle T^0, V^0\rangle, \langle TP\dagger, D^0\rangle, \langle TP\dagger, N^0\rangle\}$

b. $\left\{ \begin{array}{lll} the < child & child < can & can < cry \\ the < can & child < cry & \\ the < cry & & can < the \\ & & cry < the \\ & & can < child \\ & & cry < child \end{array} \right\}$

≡ can cry the child can cry

(8b) violates antisymmetry.

(9) Antisymmetry
A linearization cannot contain both $a < b$ and $b < a$.

The problem with (6) is that it allows too many asymmetric c-commanding pairs to enter *L*. Because TP† is part of some of the pairs in *L*, the orderings *can*<*the*, *can*<*child*, *cry*<*the* and *cry*<*child* get into the linearization. But because DP is also part of some of the pairs in *L*, the linearization contains *the*<*can*, *the*<*cry*, *child*<*can* and *child*<*cry*. To address this problem, Kayne proposes a way of limiting the class of items that can be in *L* so that it achieves certain goals his system has for ordering sub-word morphemes. Because that is not a feature of the procedure needed to derive terseness, I will take a slightly different tack. I will limit *L* to just maximal and minimal projections.

(10) a. Let *L* be the set of pairs of heads and maximal projections, $\langle A, B\rangle$, in a phrase marker P such that *A* asymmetrically c-commands *B*.

b. The linearization of P is the union of $d(A) < d(B)$ for every ordered pair in L.

Because TP† is neither a minimal nor a maximal projection it will be jettisoned from L. (10) will produce the L in (11a), and this generates the correct linearization in (11b).

(11) a. $L = \{\langle D^0, N^0\rangle, \langle DP, T^0\rangle, \langle DP, VP\rangle, \langle DP, V^0\rangle, \langle T^0, V^0\rangle\}$

b. $\left\{ \begin{array}{lll} the < child & child < can & can < cry \\ the < can & child < cry & \\ the < cry & & \end{array} \right\}$

≡ the child can cry

(10) correctly linearizes a wide array of syntactic structures and provides a way of deriving (4).

We are now ready to see how Nunes proposes to derive terseness. His proposal amounts to adopting (12).

(12) A term, X, and , X', cannot be distinguished by antisymmetry.

A consequence of (12) is that a linearization which contains both $X < Y$ and $Y < X'$ will violate antisymmetry. Applying (10) to the result of movement in Figure 7.2 produces the linearization in (13b).

(13) a. $L = \left\{ \begin{array}{lllll} \langle D^{0\prime}, N^{0\prime}\rangle & \langle DP', C^0\rangle & \langle DP', TP\rangle & \langle DP', DP\dagger\rangle & \langle DP', D^0\dagger\rangle \\ \langle DP', T^0\rangle & \langle DP', VP\rangle & \langle DP', V^0\rangle & \langle DP', DP\rangle & \langle DP', D^0\rangle \\ \langle DP', NP\rangle & \langle DP', N^0\rangle & \langle C^0, DP\dagger\rangle & \langle C^0, D^0\dagger\rangle & \langle C^0, T^0\rangle \\ \langle C^0, VP\rangle & \langle C^0, V^0\rangle & \langle C^0, DP\rangle & \langle C^0, D^0\rangle & \langle C^0, NP\rangle \\ \langle C^0, N^0\rangle & \langle DP\dagger, T^0\rangle & \langle DP\dagger, VP\rangle & \langle DP\dagger, V^0\rangle & \langle DP\dagger, DP\rangle \\ \langle DP\dagger, D^0\rangle & \langle DP\dagger, NP\rangle & \langle DP\dagger, N^0\rangle & \langle T^0, V^0\rangle & \langle T^0, DP\rangle \\ \langle T^0, D^0\rangle & \langle T^0, NP\rangle & \langle T^0, N^0\rangle & \langle V^0, D^0\rangle & \langle V^0, NP\rangle \\ \langle V^0, N^0\rangle & \langle D^0, N^0\rangle & & & \end{array} \right\}$

b. $\left\{ \begin{array}{lllll} which' < flower' & flower' < should & should < she & she < bring & bring < which \\ which' < should' & flower' < she & should < bring & she < which & bring < flower \\ which' < she & flower' < bring & should < which & she < flower & which < flower \\ which' < bring & flower' < which & should < flower & & \\ which' < which & flower' < flower & & & \\ which' < flower & & & & \end{array} \right\}$

≡ which′ flower′ should she bring which flower

Because of the existence of *which'*<*bring* and *bring*<*which* in (13b), along with many other such pairs, antisymmetry is violated.

This derives the impossibility of speaking a moved term in both of the places it occupies, but something more is needed to produce the string that actually arises. Nunes suggests that this involves a movement-specific deletion operation which removes orderings from a linearization. Applied to (13), this deletion operation could remove orderings to form one of the strings in (14), all of which satisfy antisymmetry.

(14) a. which flower should she bring
b. which should she bring flower
c. flower should she bring which
d. should she bring which flower

Nunes assumes, and so shall I, that (14a) and (14d) are possible outcomes – some languages choosing one or the other – but that (14b) and (14c) are not. To block these two outcomes, Nunes makes two assumptions. First the deletion operation in question applies not to a linearization – it doesn't remove elements of the set in (13) for instance – but to the syntactic structure being linearized. It removes the linearization statements corresponding to the phrases and heads that populate a syntactic representation. I'll formulate Nunes' condition, which he calls *chain reduction*, to reflect this.

(15) Chain reduction
Chain reduction applied to d(X) deletes every ordered pair in a linearization that contains a word in d(X), X a head or phrase.

To form the strings in (14), chain reduction will delete from L the ordered pairs indicated in (16).

(16) a. To form (14a), chain reduction applies to d(DP).
b. To form (14b), chain reduction applies to d(NP′) and d(D^0).
c. To form (14c), chain reduction applies to d($D^{0\prime}$) and d(NP).
d. To form (14a), chain reduction applies to d(DP′).

The second assumption Nunes makes is that there is an economy condition that favors fewer targets for chain reduction.

(17) Economy
Let N be the number of terms that an instance of chain reduction, R, applies to. Block R if its N is greater than the N for another R that satisfies antisymmetry.

Economy will block the applications of chain reduction in (16b) and (16c) because of the equally antisymmetry compliant applications of chain reduction in (16a) and (16d).

There are a variety of successes for this method of deriving terseness, and I will not challenge it. Instead, I will focus on understanding (12). Why is antisymmetry unable to distinguish a term from its copy?

3 Multidominance

A simple way of explaining why a term and its copy are the same thing for antisymmetry is that they *are* the same thing. Rather than modeling movement as an operation that creates a copy of a term and puts that term in an additional position, we could model movement as an operation that puts one term in two positions. This is a thesis that Engdahl (1980), Starke (2001), de Vries (2007), Gärtner (2002), among others, have suggested.

An immediate problem with this view, though, is that it leads to the expectation that the denotation a phrase has will be the same in both of the positions movement relates it to. Consider, for instance, a way of representing this thesis that allows one term to have two positions in a phrase marker. That would give Figure 7.2 the representation in Figure 7.3.

There is evidence that the semantics of constituent questions of this kind must be able to involve a binder/variable relation. In principle, we want phrasal movement to be able to cause a moved phrase to bind a variable in the position it moves from. The representation in Figure 7.3 makes that possibility obscure. The single phrase, *which flower*, would not seem to be able to simultaneously have the meaning of a variable and the meaning of the term that binds that variable.[4] We want to define "copy of" so that it gives the equivalent of Figure 7.3 for antisymmetry, but not for the meanings involved.

In Johnson 2012, I argue that the solution to this dilemma comes from recognizing that there can be material in the higher position that is not part of the term that has moved. If we represent this additional material with "Q," then Figure 7.3 can be replaced by Figure 7.4.

Depending on the kind of semantic relation involved, we can credit the denotation of Q^0 with being responsible for creating a binder out of the higher phrase. See Johnson 2012 for details. I will assume that movement is an operation that puts one term in two positions, but that it does so always in a way parallel to Figure 7.4. The moved item is part of a larger term in the higher position.

[4] But see Engdahl (1986) for a method.

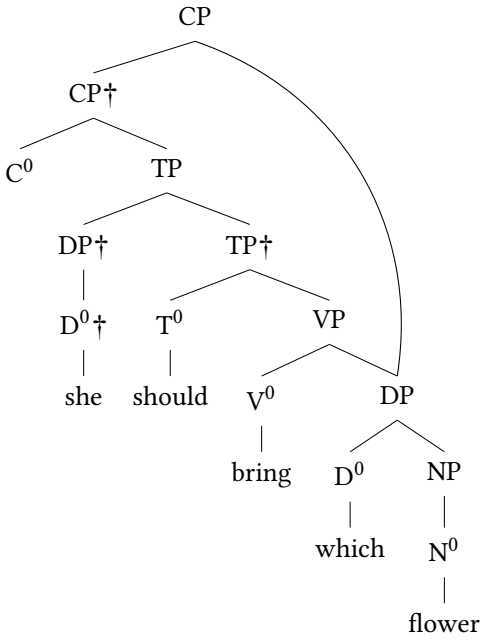


Figure 7.3: Remerge structure

Adopting this view requires a recasting of Nunes' method of deriving terseness. We cannot rely on an operation like chain reduction to fix the violations of antisymmetry that movement will create as it will overshoot. To see this, consider how (10) will apply to Figure 7.4; it produces the linearization in (18).

(18) a. $$L = \left\{ \begin{array}{lllll} \langle X^0,D^0\rangle & \langle X^0,NP\rangle & \langle X^0,N^0\rangle & \langle XP,C^0\rangle & \langle XP,C^0\rangle \\ \langle XP,TP\rangle & \langle XP,DP\dagger\rangle & \langle XP,D^0\dagger\rangle & \langle XP,T^0\rangle & \langle XP,VP\rangle \\ \langle XP,V^0\rangle & \langle C^0,DP\dagger\rangle & \langle C^0,D^0\dagger\rangle & \langle C^0,VP\rangle & \langle C^0,V^0\rangle \\ \langle C^0,DP\rangle & \langle C^0,D^0\rangle & \langle C^0,NP\rangle & \langle C^0,N^0\rangle & \langle DP\dagger,T^0\rangle \\ \langle DP\dagger,VP\rangle & \langle DP\dagger,V^0\rangle & \langle DP\dagger,DP\rangle & \langle DP\dagger,D^0\rangle & \langle DP\dagger,NP\rangle \\ \langle DP\dagger,N^0\rangle & \langle T^0,V^0\rangle & \langle T^0,DP\rangle & \langle T^0,D^0\rangle & \langle T^0,NP\rangle \\ \langle T^0,N^0\rangle & \langle V^0,D^0\rangle & \langle V^0,NP\rangle & \langle V^0,N^0\rangle & \langle D^0,N^0\rangle \end{array} \right\}$$

b. $$\left\{ \begin{array}{lllll} Q < which & which < flower & flower < she & she < should & should < bring \\ Q < flower & which < she & flower < should & she < bring & should < which \\ Q < she & which < should & flower < bring & she < which & should < flower \\ Q < should & which < bring & flower < which & she < flower & bring < which \\ Q < bring & which < which & flower < flower & & bring < flower \end{array} \right\}$$

≡ Q which flower should she bring which flower

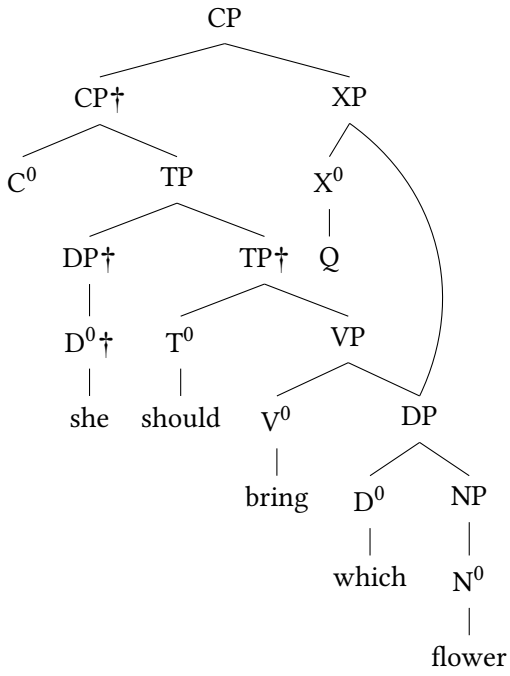


Figure 7.4: Parallel Merge structure

There are numerous violations of antisymmetry in (18b) (e.g., *which*<*bring* and *bring*<*which*) as well as the arguably anomalous *which*<*which* and *flower*<*flower*. For chain reduction to remove these violations, it would have to apply to either d(XP) or d(DP). If it applies to d(DP), (18) will lose all ordered pairs that have either *which* or *flower* in them, producing a linearization that is equivalent to (19).

(19) Q she should bring

If movement puts one thing in two places, thereby explaining (12), then something must replace chain reduction in Nunes' explanation for terseness.

A minimal modification of Nunes' system would be to allow the pairs that go into *L* to be partial in a way that mimics chain reduction. Rather than removing ordering statements that produce a violation of antisymmetry, we can allow the linearization to avoid introducing them to begin with. (10) becomes (20).

(20) a. Let *L* be a set of pairs of heads and maximal projections, $\langle A, B \rangle$, in a phrase marker P such that *A* asymmetrically c-commands *B*.

b. The linearization of P is the union of $d(A) < d(B)$ for every ordered pair in L.

Unlike (10a), which required that L contain $\langle A, B\rangle$ for every A that asymmetrically c-commands B, (20a) allows L to contain a proper subset of such ordered pairs: all it requires is that L contain $\langle A, B\rangle$ only if A asymmetrically c-commands B. (20) allows partial orderings, and so it will have to be coupled with something that ensures that every word in a syntactic representation end up in the linearization. This can be achieved by adopting another of Kayne's (1994)'s well-formedness conditions:

(21) Totality
If a and b are words in P, then either $a < b$ or $b < a$ must be in the linearization of P.

(20) will allow for the English linearization of Figure 7.4 – in (22) – and totality will prevent incomplete outcomes like (19).

(22) a. $$L = \begin{Bmatrix} \langle X^0, D^0\rangle & \langle X^0, N^0\rangle & \langle D^0, N^0\rangle & \langle XP, C^0\rangle & \langle XP, DP\dagger\rangle \\ \langle XP, T^0\rangle & \langle XP, V^0\rangle & \langle C^0, DP\dagger\rangle & \langle C^0, T^0\rangle & \langle C^0, V^0\rangle \\ \langle DP\dagger, T^0\rangle & \langle DP\dagger, V^0\rangle & \langle T^0, V^0\rangle & & \end{Bmatrix}$$

b. $$\begin{Bmatrix} Q < which & which < flower & flower < she & she < should & should < bring \\ Q < flower & which < she & flower < should & she < bring & \\ Q < she & which < should & flower < bring & & \\ Q < should & which < bring & & & \\ Q < bring & & & & \end{Bmatrix}$$

≡ Q which flower she should bring

Moreover, (20) will also correctly block (14b) and (14c), in which *which* and *flower* are linearized in non-contiguous positions. This is because for totality to be satisfied, XP must be in L. Only if XP is in L will Q get linearized with all the words that are not in XP. But once XP is in L, all of the words in XP (i.e., *Q*, *which* and *flower*) will be linearized in the same way to every word not in XP. A feature of (20) is that it enforces contiguity on any phrase that enters L.[5]

(23) Contiguity
A linearization is contiguous if for every phrase, XP, in L, if $b \notin d(XP)$, then $b < a$ or $a < b$ for every $a \in d(XP)$.

[5]There is a very close resemblance between contiguity and the central condition in Lisa Selkirk's (2011) match theory, which requires that phrases map onto prosodic units that contain every word within them. A tantalizing prospect is to reduce contiguity to this condition on the syntax/prosody mapping.

An interesting feature of movement is that it creates structures which violate a stronger form of contiguity, one that holds of every phrase in a structure, not just those used to form a linearization. This stronger form of contiguity is quite widely honored by linearization; we should have an account for why it is relaxed just for movement structures. (20) takes a step towards doing this by letting contiguity hold not of the entire phrase marker, but of the subset of phrases chosen from that phrase marker to base a linearization on. Totality forces this subset to be sufficiently representative, spreading contiguity among the non-moved parts of the phrase marker. The moved parts of a phrase marker are allowed to violate contiguity because there is a way of satisfying totality without considering all the positions they are in.

Unfortunately, this feature of (20) prevents any other linearization of Figure 7.4, including the one Nunes' theory countenanced in (24).

(24) Q she should bring which flower

In general, if phrasal movement creates a structure in which, like Figure 7.4, the moved phrase is part of a larger phrase in the higher position, then (20) will not allow covert movement.

What this section shows is that it's possible to preserve much of the linearization algorithm that Nunes uses to explain terseness, while giving a natural and simple explanation for why antisymmetry should treat a moved term as if it's one thing in two positions. Kayne called his linearization algorithm the *linear correspondence axiom*, or LCA. Let's know this modified version of his algorithm as the *multidominant-friendly linear correspondence axiom*, or MLCA.

(25) MLCA
   a. Let L of P consist of pairs of minimal and maximal projections, $\langle A, B\rangle$, where A asymmetrically c-commands B in P.
   b. A linearization of P is the union of $d(A) < d(B)$ for every $\langle A, B\rangle$ in L of P.
   c. $d(\alpha) =_{def}$ all the words dominated by α.

(26) Antisymmetry
A linearization of P cannot contain both $a < b$ and $b < a$.

(27) Totality
A linearization of P must contain $a < b$ or $b < a$ for every pair of words a, b in P.

The MLCA has properties which should be regarded as features. Some of them are (28).

(28) MLCA Features
  a. Preserves the goal of Kayne's LCA, i.e. the generalization in (4).
  b. Enforces contiguity on a moved phrase (i.e., blocks 14b–c).
  c. Derives terseness.
  d. Produces linearizations corresponding to overt movement.

It also has a property that could be regarded a bug. If movement has the properties I argued for (Johnson 2012), then it will not allow for a linearization that corresponds to covert movement. I regard that as a bug, and so I will offer an alternative linearization scheme in the next section.

4 Paths

If a structure like Figure 7.4 is to be able to linearize into covert movement, i.e. a string in which *which flower* follows *bring*, then it will be necessary to allow *Q* and *which flower* to end up non-contiguous. This means that the linearization algorithm cannot prevent *Q* from getting into the linearization unless everything else in d(XP) gets ordered the same way to the things that XP asymmetrically c-commands. We must let *Q* get into the linearization without using XP's position to do so. I cannot see a way of doing that which preserves Kayne's program, so I will abandon (4) as a goal of the linearization scheme.[6] What shouldn't be abandoned, though, is contiguity which seems to be a general truth about how syntactic structures map onto strings. If movement employs multidominant representations, contiguity must be relaxed, but only just where multidominance arises. So my goal will be to devise a linearization algorithm which preserves contiguity in all those cases where multidominance (aka movement) doesn't arise and explain why it selectively permits violations where multidominance does arise.

Contiguity is typically conceived of as a relationship between dominance relations and contiguous strings and this is how I've stated it in (23). It enforces the law in (29).

(29) If words $a_1, \ldots, a_n$ are dominated by a phrase XP (= d(XP)), then $a_1, \ldots, a_n$ will form a contiguous substring in the linearization.

[6] See Abels & Neeleman 2012 for another direction to pursue.

For standard phrase markers that don't have multidominance in them, an equally valid way of stating the law that contiguity enforces is (30).

(30) If phrase XP_1 dominates phrase XP_2, then the words in XP_2 (i.e., d(XP_2)) will form a contiguous substring of the string formed by the words in XP_1 (i.e., d(XP_1)).

Indeed, the transitive closure of (30) holds for phrase markers that obey contiguity and don't contain multidominance.

(31) Let $p = (XP_1, XP_2, \ldots, XP_n)$ be a series of phrases such that every XP_i in p is dominated by every $XP_{j \leq i}$ in p. For every p in a phrase marker, d(XP_i) must be a contiguous substring of d($XP_{j \leq i}$) for every XP in p.

(NB: "dominance" and "substring" are reflexive.)

I will call a series of phrases that form a p, a *path*.

Interestingly, (31) isn't obeyed in a phrase-marker that allows for multidominant representations. To see this, consider Figure 7.5 and the linearization of Figure 7.5 that corresponds to overt movement, in (32).

(32) *Overt movement linearization:*

Q which flower she should bring here

Two paths that contain DP and NP in Figure 7.5 are (33).

(33) a. *Paths for NP:*
   i. (NP,DP,VP†,VP,TP†,TP,CP†,CP)
   ii. (NP,DP,XP,CP)

b. *Paths for DP:*
   i. (DP,VP†,VP,TP†,TP,CP†,CP)
   ii. (DP,XP,CP)

(32) makes (33a-i) and (33b-i) violate (31); neither *flower* (=d(NP)) nor *which flower* (=d(DP)) are contiguous substrings of d(TP) (=*she should bring which flower here*), d(TP†) (=*should bring which flower here*), d(VP) (=*bring which flower here*) or d(VP†) (=*bring which flower*). If contiguity were to be expressed in a way that derives (31), then only covert movement operations would be permitted. That's not a desirable outcome. Notice, however, that if the paths in (33a-i) and (33b-i) are ignored, the linearization in (32) doesn't violate (31). Conversely, the paths in (33a-ii) and (33b-ii) violate (31) if the linearization is (34).

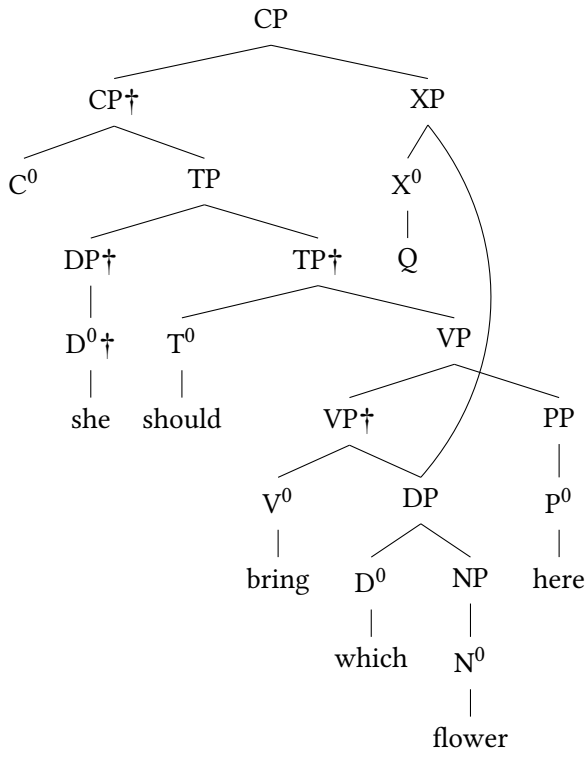


Figure 7.5: Wh-movement structure

(34) Q she should bring which flower here

Under this linearization, neither d(NP) (=*flower*) nor d(DP) (=*which flower*) are contiguous substrings of d(XP) (=*Q which flower*). This linearization doesn't violate (31), however, if the paths in (33a-ii) and (33b-ii) are ignored. Paths give us a way, then, of linearizing a phrase that is in two positions in either one of those positions. We can use paths to make movement overt or covert.

The linearization algorithm I will propose is based on paths. As we've seen, framing contiguity in terms of paths in the way that (31) does leaves its effects unchanged for phrase markers that don't have multidominance in them, but has useful effects in situations where multidominance arises. The role that asymmetric c-commanding phrases have in the MLCA will be taken up by paths in my algorithm. Words will get into a linearization by virtue of the paths they have, and so I will state totality in terms of paths too. This will also allow a phrase marker that has multidominance, and therefore more than one path for a word

or group of words, to satisfy totality by choosing just one of those paths. Finally, because the formalism for representing linearizations is a set of ordered pairs, (31) will have to be expressed in a way that references those ordered pairs rather than the strings they correspond to. Here, then, is a system that does those things.[7]

(35) Path correspondence axiom (PCA)

a. Let $p(w)$=(XP_1, XP_2,..., XP_n), a path, be the set of phrases that dominate w, a word, and include the root phrase such that every XP_i is dominated by every $XP_{j \leq i}$.

b. $\Pi(P)$ is a set of paths formed from the words in P.

c. d(XP) is the set of ws such that XP is in $p(w)$. d(w) is w.

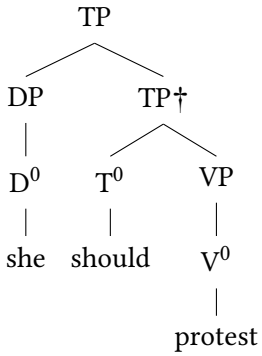
d. If p, a path, is in Π, then for every $XP \in p$, either $a < b$ or $b < a$ is in the linearization, for all $a \in d(\text{XP})$ and $b \in d(\beta)$, β XP's sister.

e. Totality
For every w in P, $\Pi(P)$ must contain $p(w)$.

Totality requires that every word in a sentence be associated with a path that is used to linearize it. The sum of these paths is Π. For each of these paths, (35d) then introduces contiguity-preserving ordered pairs into the linearization. (35d) doesn't make the language particular correct choices – that must come from a part of the linearization scheme that fixes the choices among the cross-linguistic word-orders – but it limits those choices to just ones that satisfy contiguity.

We'll look at two case studies to see how the PCA does its job. Consider first a vanilla phrase-marker with no multidominance.

(36)



For each of the words in (36), there is only one path. Consequently, the smallest Π that satisfies totality is (37).

[7]Note that the PCA does not need antisymmetry to derive terseness. It follows from the part of the PCA that enforces contiguity. Indeed, it could be that the PCA derives antisymmetry.

(37) a. p(*she*) = {DP, TP}
b. p(*should*) = {TP†, TP}
c. p(*protest*) = {VP, TP†, TP}

From these paths, we can calculate d, which relates phrases to the words that are linearized by (35d). The d of a phrase are all the words that contain that phrase in its path.

(38) a. d(TP) = {*she*, *should*, *protest*}
b. d(DP) = {*she*}
c. d(TP†) = {*should*, *protest*}
d. d(VP) = {*protest*}

(35d) requires that each of the sets in (38) map onto a contiguous substring in the linearization. For instance, for (35d) to hold of TP†, all of the words in d(TP†) (i.e., *should* and *protest*) must be ordered in the same way to the words in TP†'s sister: d(DP) (i.e., *she*). Every phrase that is in some word's path will be subject to this requirement, and so every word will be part of a series of phrases that are contiguous, each larger phrase in that path mapping onto a larger contiguous superstring containing that word.

The PCA therefore allows for the linearizations of Figure 7.5 in (39).

(39) a. she should protest
b. should protest she
c. she protest should
d. protest should she

This is probably more possibilities than should be allowed – (39d) is a sufficiently rare way for a language to linearize this structure that we might want to block it – but it comes close to what's cross-linguistically available. I will assume that the language particular choices narrow this set down to the particular outcomes appropriate for any particular language. English (a head initial, Specifier initial language) chooses (39a).

The second case study is shown in Figure 7.6. As we've seen, *which* and *flower* have two paths in Figure 7.6, and so the largest Π contains them both:

(40) a. p(*which*) = {DP, VP†, VP, TP†, TP, CP†, CP}
b. p(*which*) = {DP, XP, CP}
c. p(*flower*) = {NP, DP, VP†, VP, TP†, TP, CP†, CP}

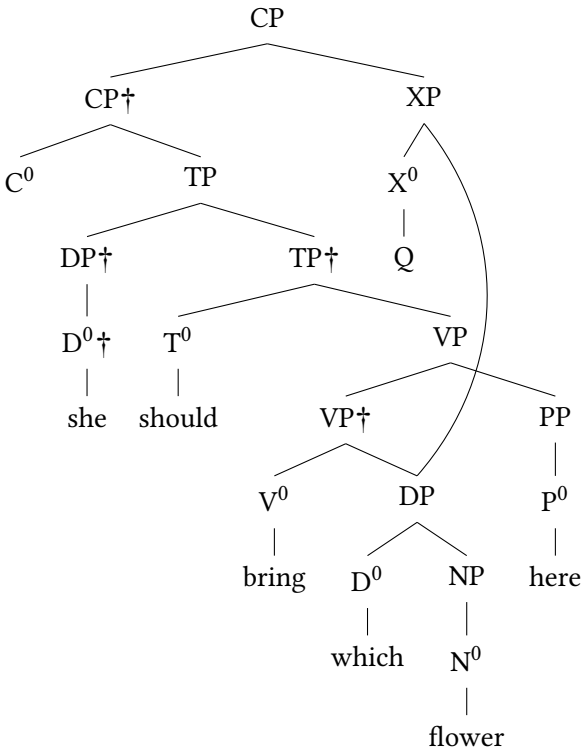


Figure 7.6: Wh-movement structured (repeated from Figure 7.5)

d. p(*flower*) = {NP, DP, XP, CP}
e. p(*bring*) = {VP†, VP, TP†, TP, CP†, CP}
f. p(*here*) = {PP, VP, TP†, TP, CP†, CP}
g. p(*should*) = {TP†, TP, CP†, CP}
h. p(*she*) = {DP†, TP, CP†, CP}
i. p(*Q*) = {XP, CP}

The values for d are:

(41) a. d(CP) = {*she*, *should*, *bring*, *here*, *Q*, *which*, *flower*}
b. d(XP) = {*Q*, *which*, *flower*}
c. d(CP†) = {*she*, *should*, *bring*, *here*, *which*, *flower*}
d. d(TP) = {*she*, *should*, *bring*, *here*, *which*, *flower*}
e. d(DP†) = {*she*}

f. d(TP†) = {*should*, *bring*, *here*, *which*, *flower*}
g. d(VP) = {*bring*, *here*, *which*, *flower*}
h. d(VP†) = {*bring*, *which*, *flower*}
i. d(DP) = {*which*, *flower*}
j. d(NP) = {*flower*}

(35d) prevents almost all linearizations of (40). It allows a linearization for this Π only under very narrow circumstances: when the language's word order settings would allow the multidominant phrase to be simultaneously contiguous to the sisters it has in both of its positions. Because of (41b), (35d) requires the linearization to have a contiguous string made from *Q, which* and *flower*. But because of (41g) and (41h), it also requires contiguous substrings made from {*bring, which, flower*} and {*bring, which, flower, here*}, which means the linearization must have one of the strings in (42) in it.

(42) a. i. bring which flower here
ii. bring flower which here
b. i. here bring which flower
ii. here bring flower which
c. i. which flower bring here
ii. flower which bring here

The strings in (42a) can't coexist in a linearization that also puts *Q* contiguous with {*which, flower*}. The strings in (42b) and (42c) can if nothing in larger phrases separates *Q*. For instance, the strings in (43) would satisfy (35d).

(43) a. Q which flower bring here should she
b. she should here bring which flower Q

I don't know of such a case, but I don't know of any harm in letting in this possibility. In general, though, (40) is too large to have a viable outcome. A smaller Π will have to be chosen.

There are four other Πs that satisfy totality. They all give to *which* and *flower* just one path. One such Π chooses paths for *which* and *flower* that go through XP; another chooses paths for *which* and *flower* that go through VP† instead. The first of these is (44) and the second (45).

(44) a. p(*which*) = {DP, XP, CP}
b. p(*flower*) = {NP, DP, XP, CP}

c. p(*bring*) = {VP†, VP, TP†, TP, CP†, CP}
d. p(*here*) = {PP, VP, TP†, TP, CP†, CP}
e. p(*should*) = {TP†, TP, CP†, CP}
f. p(*she*) = {DP†, TP, CP†, CP}
g. p(*Q*) = {XP, CP}

(45) a. p(*which*) = {DP, VP†, VP, TP†, TP, CP†, CP}
b. p(*flower*) = {NP, DP, VP†, VP, TP†, TP, CP†, CP}
c. p(*bring*) = {VP†, VP, TP†, TP, CP†, CP}
d. p(*here*) = {PP, VP, TP†, TP, CP†, CP}
e. p(*should*) = {TP†, TP, CP†, CP}
f. p(*she*) = {DP†, TP, CP†, CP}
g. p(*Q*) = {XP, CP}

The *d*s for (44) are in (46), and they correspond to the string in (47) in a head-initial and Specifier-initial language like English.

(46) a. d(CP) = {*she*, *should*, *bring*, *here*, *Q*, *which*, *flower*}
b. d(XP) = {*Q*, *which*, *flower*}
c. d(CP†) = {*she*, *should*, *bring*, *here*}
d. d(TP) = {*she*, *should*, *bring*, *here*}
e. d(DP†) = {*she*}
f. d(TP†) = {*should*, *bring*, *here*}
g. d(VP) = {*bring*, *here*}
h. d(VP†) = {*bring*}
i. d(DP) = {*which*, *flower*}
j. d(NP) = {*flower*}

(47) Q which flower she should bring here

The *d*s for (45) are in (48), and they correspond to the string in (49), in a head-initial, Specifier-initial language.

(48) a. d(CP) = {*she*, *should*, *bring*, *here*, *Q*, *which*, *flower*}
b. d(XP) = {*Q*}
c. d(CP†) = {*she*, *should*, *bring*, *here*, *which*, *flower*}
d. d(TP) = {*she*, *should*, *bring*, *here*, *which*, *flower*}
e. d(DP†) = {*she*}

f. d(TP†) = {*should*, *bring*, *here*, *which*, *flower*}
g. d(VP) = {*bring*, *here*, *which*, *flower*}
h. d(VP†) = {*bring*, *which*, *flower*}
i. d(DP) = {*which*, *flower*}
j. d(NP) = {*flower*}

(49) Q she should bring which flower

These are the desired outcomes; they correspond to the overt and covert movement possibilities.

The remaining two Πs that satisfy totality give to *which* and *flower* divergent paths. They are both blocked by the PCA. To see how, consider (50), where *flower* is given a path through XP and *which* is given a path through VP†.

(50) a. p(*which*) = {DP, VP†, VP, TP†, TP, CP†, CP}
b. p(*flower*) = {NP, DP, XP, CP}
c. p(*bring*) = {VP†, VP, TP†, TP, CP†, CP}
d. p(*here*) = {PP, VP, TP†, TP, CP†, CP}
e. p(*should*) = {TP†, TP, CP†, CP}
f. p(*she*) = {DP†, TP, CP†, CP}
g. p(*Q*) = {XP, CP}

The ds for (50) are (51).

(51) a. d(CP) = {*she*, *should*, *bring*, *here*, *Q*, *which*, *flower*}
b. d(XP) = {*Q*, *flower*}
c. d(CP†) = {*she*, *should*, *bring*, *here*, *which*}
d. d(TP) = {*she*, *should*, *bring*, *here*, *which*}
e. d(DP†) = {*she*}
f. d(TP†) = {*should*, *bring*, *here*, *which*}
g. d(VP) = {*bring*, *here*, *which*}
h. d(VP†) = {*bring*, *which*}
i. d(DP) = {*which*, *flower*}
j. d(NP) = {*flower*}

d(VP†) and d(VP) together require that the linearization produce the string *bring which here* (once English-specific choices are made). But d(DP) requires that the linearization also produce the string *which flower*. There is no way of

linearizing these words that preserves these two requirements. Exactly the same incompatibility arises if the path for *flower* goes through VP† and the path for *which* goes through XP – the other way of choosing divergent paths for these words. The reason these choices lead to a conflict is because all choices of paths for *which* and *flower* will contain DP, and (35d) will consequently require *which* and *flower* to be contiguous. This is how this system prevents the words in a moved phrase from getting linearized in different positions.

The PCA, then, allows for both overt and covert movement and, like the MLCA, explains why multidominant structures allow for selective relaxation of contiguity. It makes contiguity, rather than asymmetric c-command, the driving force behind a linearization. The formalization of contiguity involved enforces a particular kind of "nesting" condition on entire phrase markers. It allows multidominance in just those cases where that nesting condition can be satisfied for every word in the phrase marker without considering the complete structure of the sentence.

5 Summary

What I've shown here is a way of completing Nunes' method of deriving terseness that involves defining the "copy of α" as "giving α an addition position in the phrase marker." Traditional linearization schemes have stood in the way of such a move. I've offered two new linearization algorithms that don't, each with slightly different empirical footprints.

Abbreviations

LCA	linear correspondence axiom		linear correspondence axiom
MLCA	multidominant-friendly	PCA	Path correspondence axiom

Acknowledgements

This paper is largely a channeling of many people's thoughts. These include Sakshi Bhatia, David Erschler, Hsin-Lun Huang, Rodica Ivan, Jyoti Iyer, Petr Kusily, Deniz Ozyildiz, Ethan Poole, Katia Vostrikova, Michael Wilson, Rong Yin, Rajesh Bhatt, Peggy Speas, Nikolaos Angelopoulos, John Gluckman, Nicoletta Loccioni, Travis Major, Iara Mantenuto, Sozen Ozkan, Richard Stockwell, Carson Schutze and Tim Hunter. A special thanks to Leland Kusmer whose guidance has improved this paper in many ways.

References

Abels, Klaus & Ad Neeleman. 2012. Linear asymmetries and the LCA. *Syntax* 15(1). 25–74. DOI: 10.1111/j.1467-9612.2011.00163.x.

de Vries, Mark. 2007. Internal and external remerge: On movement, multidominance, and the linearization of syntactic objects. Ms., University of Groningen.

Engdahl, Elisabet. 1980. *The syntax and semantics of questions in Swedish*. Amherst, Massachusetts: University of Massachusetts, Amherst. (Doctoral dissertation).

Engdahl, Elisabet. 1986. *Constituent questions*. Dordrecht: Reidel.

Gärtner, Hans-Martin. 2002. *Generalized transformations and beyond: Reflections on minimalist syntax*. Berlin: Akademie Verlag.

Johnson, Kyle. 2012. Toward deriving differences in how *Wh* movement and QR are pronounced. *Lingua* 122(6). 529–553. DOI: 10.1016/j.lingua.2010.11.010.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Nunes, Jairo. 1995. *The copy theory of movement and linearization of chains in the minimalist program*. University of Maryland. (Doctoral dissertation).

Nunes, Jairo. 1996. On why traces cannot be phonetically realized. In Kiyomi Kusumoto (ed.), *NELS 26: Proceedings of the twenty-sixth annual meeting of the North East Linguistic Society*, 211–226. Amherst, MA: GLSA Publications.

Nunes, Jairo. 1999. Linearization of chains and phonetic relization of chain links. In Samuel David Epstein & Norbert Hornstein (eds.), *Working minimalism*, 217–249. Cambridge, MA: MIT Press.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

Peters, Stanley & Robert W. Ritchie. 1981. Phrase linking grammars. Ms., Stanford University.

Selkirk, Elisabeth. 2011. The syntax–phonology interface. In John Goldsmith, Jason Riggle & Alan C. Yu (eds.), *The handbook of phonological theory*, Second edition, 435–484. Malden, MA: Wiley-Blackwell.

Starke, Michal. 2001. *Move dissolves into Merge: A theory of locality*. University of Geneva. (Doctoral dissertation).

Chapter 8

Rethinking the reach of categorical constraints: The final-over-final constraint and combinatorial variability

Neil Myler
Boston University

This squib argues that categorical rules and constraints of the sort traditionally found in generative syntax can, in principle, make interesting and testable quantitative predictions about surface frequencies in language use, despite occasional claims to the contrary. Specifically, the final-over-final constraint (FOFC, Biberauer et al. 2014; 2009; Holmberg 2000; Walkden 2009; many others) is predicted to exert a specific influence on the likelihood of OV vs. VO word order in the language use of a speaker that allows both, given a COMBINATORIAL VARIABILITY approach to intra-speaker syntactic variation (Adger 2006 et seq.).

1 Introduction

Generative linguistics has traditionally employed categorical rules and constraints in its quest to understand the properties of the syntax of particular languages and the properties of the syntactic component of the language faculty more generally. For this reason, its theoretical postulates have often been taken to be either irrelevant to or at odds with the inherent variability of language use (see Guy 2005; Newmeyer 2005; inter alia).

In this squib, I will argue that categorical constraints can, in fact, make interesting and testable quantitative predictions about surface frequencies, given a certain theory of how intra-speaker syntactic variation is to be modeled. More

Neil Myler. 2020. Rethinking the reach of categorical constraints: The final-over-final constraint and combinatorial variability. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 137–147. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280641

specifically, I will show that the *final-over-final constraint*[1] (FOFC – Biberauer et al. 2014; 2009; Holmberg 2000; Walkden 2009, many others) should exert a specific influence on the likelihood of OV vs. VO word order in the language use of a speaker that allows both, given a *combinatorial variability* approach to intra-speaker syntactic variation (Adger 2006 et seq.).

The squib is structured as follows. In §2, I introduce the combinatorial variability approach, showing how it might be used to generate predictions concerning the expected baseline surface frequencies of OV vs. VO order in the speech of Quechua–Spanish bilinguals, focusing on DP complements and the head-directionality of VP and TP. In §3, I introduce FOFC and demonstrate that the surface frequencies predicted by the combinatorial variability approach change if FOFC is held to be valid. In §4, I outline the prospects and challenges for testing these predictions in a sociolinguistic study of actual Quechua–Spanish bilinguals in Cochabamba, Bolivia. §5 is a brief conclusion.

2 Quechua–Spanish contact and combinatorial variability

To make the discussion of combinatorial variability more concrete, I will frame this section around the specific example of language contact between speakers of Quechua and Spanish. Speakers of these two languages are in contact in Peru, Bolivia, Ecuador, parts of Colombia, and parts of northern Chile and northern Argentina. Many Quechua speakers in these places are bilingual in Spanish. As is well-known, Quechua and Spanish are almost typological opposites in terms of their basic word order. Quechua is predominantly head-final, as shown in the example from Cochabamba Quechua (a Bolivian variety) in (1). Spanish, on the other hand, is a head-initial language, as shown in (2).

(1) Cochabamba Quechua
Kay runa Cochabamba-man ri-q ka-rqa.
This man Cochabamba-to go-NMLZ be-PST
'This man used to go to Cochabamba.'

(2) Spanish
Este hombre ha ido a Cochabamba.
This man has gone to Cochabamba
'This man has gone to Cochabamba.'

[1]Note that FOFC is referred to as the *final-over-final condition/constraint* in some more recent work, including Sheehan et al. (2017).

Pre-theoretically, one might expect contact between Quechua speakers and Spanish speakers to give rise to mutual influence on word order, such that head-initial orders increase in Quechua usage, and/or head-final ones increase in Spanish usage, depending on the degree of bilingualism of the speaker, attitudes towards each language, and so on. Indeed, such has been reported in the literature on Andean Spanish (e.g., Muntendam 2008; Muysken 1984; Sánchez 2003) and in studies of the influence of Spanish on Quechua (Camacho 1999; Hintz 2009; Sánchez 2003, 2012). Let us now turn to the combinatorial variability approach, and how it might analyze such variation.

Comparative syntax research within the Minimalist program has pursued the idea that syntactic variation across languages/dialects should be analyzed only in terms of variation in the featural needs of functional items (the so-called Borer-Chomsky conjecture, as it is dubbed by Baker 2008; see Borer 1984; Chomsky 1995). This presents a generativist pathway to *orderly heterogeneity* in the sense of Weinreich et al. (1968): Suppose that an individual's lexicon contains function morphemes with the same categorial feature and the same contribution to truth conditions (and thus roughly the same distribution), but which differ in one or more of their morphosyntactic features. Then, the choice of one or the other lexical item in a derivation will result in somewhat different outputs, but with no difference in meaning. Thus, there will be an appearance of syntactic optionality, but in reality the only optionality is in lexical choice: once particular lexical items have been chosen, the syntactic derivation is fully determined. This is the essence of Adger's (2006 et seq.) proposed reconciliation of Minimalist syntax with sociolinguistic variation.

As Adger (2006) points out, it is possible to calculate quantitative predictions about variability which arise from the combinatorics of the relevant syntactic elements (hence the name *combinatorial variability* for the overall approach). Take lexical items A, B, and C; all with identical truth-conditional meaning but with distinct syntactic features. A and B, when chosen, give rise to a series of derivational steps S_1. C, on the other hand, differs in some aspect of its feature content from A and B, and thus gives rise to a distinct derivation S_2, whose output differs on the surface from S_1. This will give the appearance of syntactic variability. All else held equal, a prediction is made about the nature of that variability. Since two out of a possible three lexical choices give rise to S_1, but only one choice yields S_2, the prediction is that the output corresponding to S_1 should appear in usage two thirds of the time, and the output of S_2 should appear one third of the time.[2]

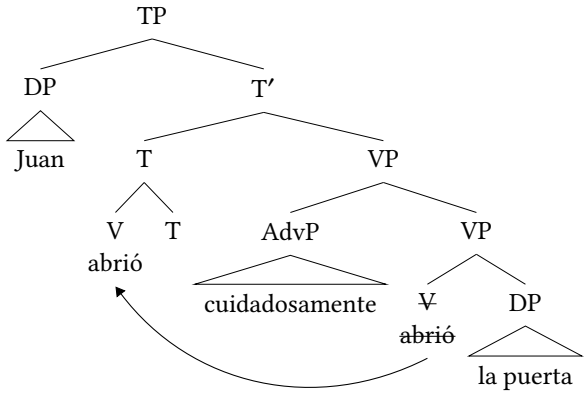
[2]This follows only if no other factors favor A, B, or C over the others, so that the choice is determined by chance. In actual use, of course, the probability distribution predicted by purely syntactic combinatorics will be modulated by sets of factors influencing lexical choice itself, including sociolinguistic factors. I return to this issue below.

Returning to our example from Quechua–Spanish contact, we will now examine the baseline frequencies of OV and VO word order that a combinatorial variability approach would predict. First, we need an inventory of the syntactic microparameters that are relevant to analyzing word-order differences between the two languages.

The first is HEAD-DIRECTIONALITY OF THE VP.[3] In Spanish, the head of VP is on the left (this value will be denoted "L" for short). In Quechua, the head of the VP is on the right ("R" for short).

The second parameter is HEAD-DIRECTIONALITY OF THE TP. This parameter, of course, is directly analogous to the first. Spanish T is on the left, and Quechua T is on the right. This parameter has a direct influence on where the verb surfaces relative to its complement, because T in these languages attracts the verb (i.e., there is V-to-T movement). V-to-T movement is known to apply in Spanish because of the placement of VP-peripheral adverbs relative to the verb and the direct object (Pollock 1989; Zagona 2002).[4]

(3) Spanish
Juan abrió cuidadosamente la puerta.
Juan opened carefully the door
'Juan carefully opened the door.'



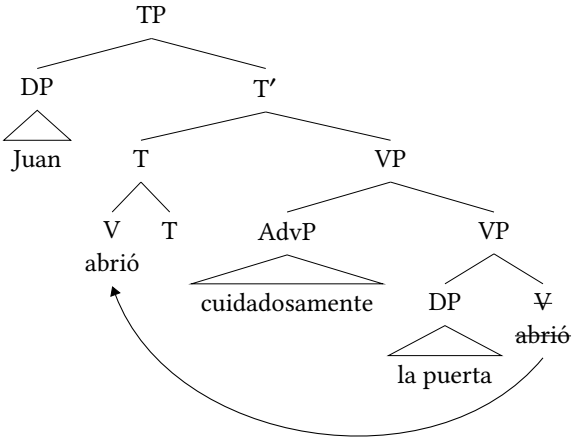
[3]For simplicity I will assume the traditional head parameter in the ensuing discussion, but nothing I have to say is incompatible with an antisymmetric approach to the relationship between structure and linearization (see Kayne 1994). Since Kayne's linear correspondence axiom is a key component of many existing approaches to deriving FOFC, this is good news.

[4]I assume here that T is the relevant landing site in all cases, but this is certainly an oversimplification. See Schifano (2015; 2018) for evidence that considerably more granularity is needed, with verb movement targeting different positions in the Cinquean extended IP (Cinque 1999 et seq.) in different languages. This does not affect the main point here, so long as verb movement is to a landing site higher in the structure than the final position of the direct object. Thanks to an anonymous reviewer for raising this issue.

It is much more difficult to ascertain whether or not there is V-to-T movement in Quechua, since both VP and TP are head-final in that language, and this makes it impossible to check whether the verb "crosses over" adverbs at the edge of VP. The empirical evidence we have to hand is therefore compatible with V-to-T movement being present or absent in Quechua. However, there is one typological consideration which weighs in favor of assuming that Quechua does have V-to-T movement. The syntactic literature has found that VO languages with rich agreement inflection on the finite verb always have V-to-T movement (Kosmeijer 1986; Pollock 1989; see Koeneman & Zeijlstra 2012 for a recent reaffirmation of this correlation). Since Quechua has extremely rich agreement inflection on its finite verbs, we may assume it has V-to-T movement also.[5]

To see why this matters for surface word-order, consider the case of a derivation in which VP-headedness has the Quechua "R" value, but TP-headedness has the Spanish "L" value. In such a case, the surface word order will be VO in spite of the fact that the structure is "underlyingly" OV, because of V-to-T movement.

(4) V-to-T movement obscures head-finality of VP



Given these basic assumptions about clause structure and the points of parametric variation which differentiate Spanish and Quechua, we can now ask about the predictions of combinatorial variability for the baseline frequencies of OV vs. VO order.

[5] An anonymous reviewer points out that there remain a number of potential problems for this conclusion (referring to Vikner 2005; Han et al. 2007; 2016). This must be borne in mind, because if it turns out that Quechua lacks V-to-T, then another test-bed for the quantitative predictions of FOFC would need to be found. The broader point of this squib, that such predictions are formulable and testable in principle, stands regardless.

Let us assume that a bilingual speaker is able to represent syntactic objects from each language in much the same way as a monolingual speaker. That is, a bilingual speaker has access to a left-headed VP structure much as a monolingual Spanish speaker does, and also has access to a right-headed VP structure in the same way that a monolingual Quechua speaker does. Similarly, the bilingual's functional lexicon will contain a lexical item T which takes its complement to its right, Spanish-style, and another lexical item T which takes its complement to the left, Quechua-style, and so on for other syntactic objects. Of course, in making utterances, bilingual speakers will have to make a choice between these options. It turns out that the different parameter settings discussed above, simply through the nature of their logically possible combinations, give rise to quantitative predictions about what the baseline frequencies of these different choices should be.

For the purposes of simplicity, I will concentrate on DP direct objects only. The calculations below would have to be somewhat different for QP and CP complements. In the case of QPs, the fact that Quechua allows overt scrambling for scope would somewhat increase the chance of OV order surfacing, relative to non-quantificational DPs. For CPs, the possibility of clausal extraposition in both languages would boost the predicted baseline frequency of VO order.

There are $2 * 2 = 4$ possible combinations of parameter settings relevant here, shown below.

(5) Combinations of parameter settings: DPs

Combination A	Output: VO
Parameter	Setting
VP-headedness	L
TP-headedness	L

Combination B	Output: OV
Parameter	Setting
VP-headedness	R
TP-headedness	R

Combination C	Output: VO
Parameter	Setting
VP-headedness	R
TP-headedness	L

Combination D	Output: OV
Parameter	Setting
VP-headedness	L
TP-headedness	R

Hence, the logically possible combinations predict a 50/50 split between VO orders and OV orders for DPs.

(6) VO vs. OV order with DP complements
VO = 2/4 outputs = 50%
OV = 2/4 outputs = 50%

3 Bringing in the final-over-final constraint (FOFC)

The *final-over-final constraint* of Biberauer et al. (2014: 171) has an interesting effect on this calculation.

(7) *The final-over-final constraint* (FOFC)
A head-final phrase αP cannot dominate a head-initial phrase βP, where α and β are heads in the same extended projection.

This constraint will, of course, make the categorical prediction that V-O-Aux orders will be absent from compound tenses in the Spanish and the Quechua of bilinguals. In addition, however, FOFC has a quantitative effect. In particular, it rules out combination D in (5), because that combination involves a head-final TP dominating a head-initial VP. In terms of the predicted baseline surface frequencies, we thus obtain the following results instead of the ones we saw in (6):

(8) VO vs. OV order with DP complements (if FOFC is valid)
VO = 2/3 outputs = 67%
OV = 1/3 outputs = 33%

This is an exciting finding, because it shows that categorical constraints can give rise to stochastic effects, meaning that such constraints *are* of potential relevance to variationist work after all. This result emerges from the fact that combinatorial variability derives quantitative predictions by looking at the interaction of different parameter settings, and universal constraints like FOFC take certain combinations of parameter settings out of the picture. Another intriguing consequence of this result is that it becomes possible, in principle, to use variationist data to test the predictions of such universal constraints. Since the baseline frequencies predicted are different if FOFC holds than they are if it does not, in principle it becomes possible to test FOFC by seeing how the variationist data pan out. In the next section, I examine the prospects for doing this.

4 Testing the predictions: Prospects and challenges

It is clear what the signature of FOFC should be in quantitative data: because FOFC bars one of the logically possible routes to OV word order, OV should be less common than VO all else held equal if FOFC is valid. If FOFC is not valid, then OV and VO should be equally frequent, all else held equal.

The challenge in testing predictions of this sort, of course, is that all else is seldom equal, and a range of social factors that have been discussed in the sociolinguistics literature will also influence the actual surface frequencies of the orders. These must be controlled for or accommodated somehow if the signature of FOFC is to be detected. Most obviously, although the literature reports mutual influence between Spanish and Quechua word orders, it still might be the case that speakers have some (presumably subconscious) sense that Quechua exhibits more head-finality. If so, language mode would be expected to favor OV when the speaker is talking in Quechua, and VO when the speaker is talking in Spanish. Such an effect would be especially likely if the VO vs. OV difference turned out to be a socially salient linguistic variable.

The issue of social salience raises the possibility that speakers might use OV vs. VO order as a way of indexing particular identity categories, including attitudes to Quechua and Spanish, orientation towards or away from indigenous culture, and so on. Since exposure to standard Spanish will favor VO order, degree of education is another factor to be considered. In addition, of course, degree of bilingualism/proficiency in each language would be expected to be relevant.

Finally, there is a presupposition of the combinatorial variability approach which itself has yet to be tested; namely, the idea that the probability that a given variable will be used is determined by chance if no other factor intervenes. This assumption is not unreasonable, but nor is it certain to be correct – we still await an empirical demonstration that it is on the right track.

In an ongoing collaboration, the sociolinguist Daniel Erker and I have carried out a pilot study involving demographic/attitudinal surveys, sociolinguistic interviews, reading passage data, and grammaticality judgments on both Spanish and Quechua as spoken in Cochabamba, Bolivia. The data set includes 19 speakers: 4 monolingual Spanish speakers, and 15 Quechua–Spanish bilinguals. For the bilinguals, we have interview data, reading passage data, and grammaticality judgment data on both languages. The analysis of this data is still in progress. As well as addressing a number of issues in the sociolinguistics of language contact, we hope that a full version of this study (including monolingual Quechua speakers, and many more speakers overall) will allow us to test the quantitative predictions of FOFC, and the predictions of the combinatorial variability approach more generally.

5 Conclusion

This squib has shown that categorical principles and constraints can make predictions about apparently non-categorical phenomena. Testing those predictions, however, is a difficult and delicate task, one that is not yet within our reach from a practical standpoint. Bringing it within our reach will require the collaboration of formal linguists and sociolinguists.

Abbreviations

FOFC	final-over-final condition/constraint	NMLZ	nominalizer/nominalization
		PST	past

Acknowledgements

For Ian Roberts, the syntactic Eric Cantona.

This squib is an outgrowth of ongoing joint work with Daniel Erker on Quechua–Spanish contact, including a poster we jointly presented at the annual meeting of the Linguistic Society of America in 2016. I would like to thank the audience of that poster presentation, a colloquium audience at UPenn, anonymous reviewers from the National Science Foundation, David Adger, Byron Ahn, Carol Neidle, Cathy O'Connor, and Ian Roberts for their feedback on various aspects of my and Daniel's joint work. Thanks also to two anonymous reviewers and to Daniel Erker for their comments on an earlier version of this squib. Errors and infelicities are on my head alone.

References

Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42(3). 503–530. DOI: 10.1017/S002222670600418X.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–373. Amsterdam: John Benjamins. DOI: 10.1075/la.132.16bak.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.

Biberauer, Theresa, Glenda Newton & Michelle Sheehan. 2009. Limiting synchronic and diachronic variation and change: The Final-over-Final Constraint. *Language and Linguistics* 2009(10). 701–743.

Borer, Hagit. 1984. *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris.

Camacho, José. 1999. From SOV to SVO: The grammar of interlanguage word order. *Second Language Research* 15(2). 115–132. DOI: 10.1191/026765899673532714.

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: A cross-linguistic perspective*. New York: Oxford University Press.

Guy, Gregory. 2005. Grammar and usage: A variationist response. *Language* 81(3). 561–563. DOI: 10.1353/lan.2005.0132.

Han, Chung-hye, Jeffrey Lidz & Julien Musolino. 2007. V-raising and grammar competition in Korean: Evidence from negation and quantifier scope. *Linguistic Inquiry* 38(1). 1–47. DOI: 10.1162/ling.2007.38.1.1.

Han, Chung-hye, Julien Musolino & Jeffrey Lidz. 2016. Endogenous sources of variation in language acquisition. *PNAS* 113(4). 942–947. DOI: 10.1073/pnas.1517094113.

Hintz, Daniel J. 2009. Reordenamiento sintáctico en construcciones verbales analíticas del quechua por el contacto con el castellano. *Lingüística* 22. 187–201.

Holmberg, Anders. 2000. Deriving OV order in Finnish. In Peter Svenonius (ed.), *The derivation of VO and OV*, 123–152. Amsterdam: John Benjamins. DOI: 10.1075/la.31.06hol.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Koeneman, Olaf & Hedde Zeijlstra. 2012. One law for the rich and another for the poor: The Rich Agreement Hypothesis rehabilitated. lingbuzz/001462.

Kosmeijer, Wim. 1986. The status of the finite inflection in Icelandic and Swedish. *Working Papers in Scandinavian Syntax* 26. 1–41.

Muntendam, Antje. 2008. Crosslinguistic influence in Andean Spanish: Word order and focus. In Melissa Bowles, Rebecca Foote, Silvia Perpiñán & Rakesh Bhatt (eds.), *Selected proceedings of the 2007 Second Language Research Forum*, 311–322. Somerville, MA: Cascadilla Proceedings Project.

Muysken, Pieter. 1984. The Spanish that Quechua speakers learn: L2 learning as norm-governed behavior. In Roger Andersen (ed.), *Second languages: A crosslinguistic perspective*, 101–119. Rowley, MA: Newbury House.

Newmeyer, Frederick J. 2005. *Possible and probable languages*. Oxford: Oxford University Press.

Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

Sánchez, Liliana. 2003. *Quechua-Spanish bilingualism: Interference and convergence in functional categories*. Amsterdam: John Benjamins. DOI: 10.1075/lald.35.

Schifano, Norma. 2015. *Verb movement: A Pan-Romance investigation*. University of Cambridge. (Doctoral dissertation).

Schifano, Norma. 2018. *Verb movement in Romance: A comparative study*. Oxford: Oxford University Press.

Sheehan, Michelle, Theresa Biberauer, Ian Roberts & Anders Holmberg. 2017. *The Final-over-Final Condition: A syntactic universal*. Cambridge, MA: MIT Press.

Vikner, Sten. 2005. Immobile complex verbs in Germanic. *The Journal of Comparative Germanic Linguistics* 8(1). 85–117. DOI: 10.1007/s10828-004-0726-9.

Walkden, George. 2009. Deriving the Final-over-Final Constraint from third factor considerations. *Cambridge Occasional Papers in Linguistics* 5. 67–72.

Weinreich, Uriel, William Labov & Marvin Herzog. 1968. Empirical foundations for a theory of language change. In Winfred P. Lehmann & Yakov Malkiel (eds.), *Directions for historical linguistics*, 95–195. Austin, TX: University of Texas Press.

Zagona, Karen. 2002. *The syntax of Spanish*. Cambridge: Cambridge University Press.

Chapter 9

Rethinking restructuring

Gereon Müller
Leipzig University

An approach to restructuring with control verbs in German is developed in terms of structure removal, based on an operation Remove that acts as a counterpart to structure-building Merge. The analysis accounts for both monoclausal and biclausal properties.

1 Introduction

Virtually all approaches to restructuring in infinitival constructions developed over the last three decades postulate either uniformly monoclausal structures or uniformly biclausal structures for the phenomenon; i.e., they do not actually rely on a concept of syntactic restructuring. Against this background, the goal of the present paper is to outline an approach to restructuring with control verbs in German that radically departs from standard approaches in that it presupposes that genuine syntactic restructuring does indeed exist, and can be held responsible for conflicting pieces of evidence that suggest both a monoclausal and a biclausal structure. This, in effect, implies a return to earlier transformational approaches according to which an initial biclausal structure is eventually reduced to a monoclausal structure. Arguably, the single main reason why these approaches were at some point generally abandoned is that they depended on reanalysis rules bringing about structure removal that were both unprincipled and unrestricted. I would like to suggest that the situation is different in a derivational minimalist approach where an elementary operation Remove (which removes structure) suggests itself as a complete mirror image of the operation Merge (which builds structure), and can be shown to be empirically motivated in areas unrelated to

Gereon Müller. 2020. Rethinking restructuring. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 149–190. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280643

restructuring. Thus, given that the goal of the present paper is that of "rethinking restructuring", this not only implies a reconsideration of current approaches to restructuring, it also implies thinking of restructuring in terms of genuine restructuring again.

I will proceed as follows. In §2, I present conflicting evidence for restructuring with control verbs in German: there are arguments for a monoclausal analysis, and there are arguments for a biclausal analysis. In §3, I introduce a new approach to structure removal based on the operation Remove, and show what effects Remove can have for heads and phrases. §4 then shows how a Remove-based approach to restructuring captures both the evidence for monoclausality and the evidence for biclausality.

2 Restructuring

Abstracting away from some differences (e.g., with respect to the obligatoriness of extraposition, on which cf. Biberauer et al. 2014), non-restructuring control infinitives in German behave in crucial respects exactly like finite embedded clauses and thus uniformly demand a biclausal analysis in terms of CP embedding. In contrast, restructuring control infinitives in German exhibit both evidence for monoclausality (i.e., for the absence of at least a CP shell, possibly also of a TP or vP shell) and evidence for biclausality. Whether restructuring is possible or not needs to be marked as a lexical property with control verbs; if it is possible, it is always optional with control verbs.[1] In the next two subsections, I will first present some arguments for monoclausality, and then turn to arguments for biclausality of restructuring control infinitives in German.

[1]Two remarks. First, as observed by Fanselow (1989; 1991), there is some variation among speakers as to which (control) verbs count as (non-) restructuring predicates in German. As a tendency, it would seem that there is a correlation with age: the younger the speaker, the more verbs (s)he accepts as a restructuring predicate. Thus, some of the data classified as ungrammatical in what follows because of a wrong lexical choice may actually be acceptable to some speakers. This does not affect the generalization as such.

Second, whereas regular control verbs trigger restructuring optionally throughout, other infinitive-embedding verbs (auxiliaries, modals, causative and perception verbs, and raising verbs) trigger restructuring obligatorily. As a matter of fact, I am not aware of strong arguments for biclausality with these latter classes, and I take it to be a plausible assumption that smaller projections (than CP) are embedded with these non-control verb types to begin with. This leaves open the question of whether they then qualify as purely functional elements (see Wurmbrand 2001; 2004 on functional restructuring vs. lexical restructuring), or whether they have full V status after all, just with complements of a smaller size. In what follows, I will generally disregard restructuring non-control verbs, except for a few cases where their different behavior sheds some light on the analysis of control verbs.

2.1 Arguments for monoclausality

There are several well-known arguments for monoclausality with restructuring control verbs in German (see von Stechow & Sternefeld 1988; Grewendorf 1988; Fanselow 1991; Bayer & Kornfilt 1994; Wurmbrand 2001, and Haider 2010, among others).

2.1.1 Scrambling and unstressed pronoun fronting

First, as first observed by Ross (1967), scrambling is strictly clause-bound in German; as shown in (1a), a CP boundary cannot be crossed by this operation. The same goes for fronting of unstressed pronouns; cf. (1b). Note that embedded *dass* clauses (as in 1a) and embedded verb-second clauses (as in 1b) uniformly block these operations.[2]

(1) German

a. * dass den Fritz$_1$ keiner gesagt hat [$_{CP}$ dass wir t$_1$ einladen
that the Fritz$_{ACC}$ no-one$_{NOM}$ said has that we$_{NOM}$ invite
sollen]
should

b. * dass die Maria es$_1$ meinte [$_{CP}$ solle man t$_1$ lesen]
that the Maria$_{NOM}$ it$_{ACC}$ said should one$_{NOM}$ read

In contrast, control infinitives are transparent for scrambling and unstressed pronoun fronting if they are embedded by a restructuring verb, as in (2a,b) (with the subject control verb *versuchen* 'try' and the object control verb *empfehlen* 'recommend'), but not if they are embedded by a non-restructuring verb, as in (2c,d) (with the object control verb *auffordern* 'request' and the subject control verb *leugnen* 'deny').

(2) German

a. dass den Fritz$_1$ keiner [t$_1$ zu küssen] versuchte
that the Fritz$_{ACC}$ no-one$_{NOM}$ to kiss tried

b. dass die Maria es$_1$ ihm gestern [t$_1$ zu lesen]
that the Maria$_{NOM}$ it$_{ACC}$ him$_{DAT}$ yesterday to read
empfohlen hat
recommended has

[2] Unstressed pronoun fronting is arguably a different movement type from scrambling since it is obligatory (whereas scrambling is optional) and since it shows order-preservation properties (whereas scrambling, almost by definition, does not); see Müller (2001).

c. * dass den Fritz$_1$ keiner die Maria [$_{CP}$ t_1 zu küssen]
that the Fritz$_{ACC}$ no-one$_{NOM}$ the Maria$_{ACC}$ to kiss
aufforderte
requested

d. * dass die Maria es$_1$ gestern [$_{CP}$ t_1 zu kennen] geleugnet hat
that the Maria$_{NOM}$ it$_{ACC}$ yesterday to know denied has

Given that it is the presence of a CP projection that blocks non-clause bound scrambling with finite clauses and non-restructuring infinitives, this suggests that restructuring infinitives lack such a projection.

2.1.2 Extraposition

Extraposition can affect CPs and PPs (plus, somewhat more marginally, DPs) in German; the operation is subject to an upward boundedness constraint (see Ross 1967) according to which a clause boundary must not be crossed in the course of rightward movement. The following examples show how CP extraposition and PP extraposition are impossible across a CP boundary as it shows up with finite clauses (cf. 3a) and infinitival complements of non-restructuring verbs (cf. 3b), respectively (see Müller 1995).[3]

(3) German

a. * [$_{CP_0}$ Er denkt [$_{CP_1}$ dass Antje [$_{DP_2}$ den Versuch t_3]
he$_{NOM}$ thinks that Antje$_{NOM}$ the attempt$_{ACC}$
aufgegeben hat] [$_{CP_4}$ weil er sie nicht mehr sieht] [$_{CP_3}$
given up has because he her not anymore sees
mit fünf Bällen zu jonglieren]]
with five balls to juggle

b. * dass Karl [$_{CP}$ das Buch t_1 zu kennen] geleugnet hat [$_{PP_1}$
that Karl$_{NOM}$ the book$_{ACC}$ to know denied has
über dieses Thema]
about this topic

[3] In (3a), CP_3 undergoes extraposition from within CP_1; CP_4 is an adjunct clause modifying CP_0 (not CP_1). CP_4 thus indicates that CP_3 must have left the domain of CP_1, and this violates the upward boundedness constraint. (The presence of an adjunct in the CP_0 clause is necessary to show that CP_1 has indeed been crossed by extraposition since finite clauses usually follow the verb in German.) This issue does not arise with infinitivals in a pre-verbal position, as in (3b).

Again, infinitival complements of restructuring verbs behave differently in that CP and PP extraposition are possible in these contexts; see (4a,b). This can then be taken to indicate that there is no CP boundary present.

(4) German

a. dass sie [das Buch t_1 zu lesen] versucht hatte [$_{CP_4}$ als sie
that she$_{NOM}$ the book$_{ACC}$ to read tried had when she
dort lebte] [$_{CP_1}$ das alle Preise gewonnen hatte]
there lived that all prizes$_{ACC}$ won had

b. dass ihr keiner [das Buch t_1 zu lesen] empfohlen hat
that her$_{DAT}$ no-one$_{NOM}$ the book$_{ACC}$ to read recommended has
[$_{PP_1}$ über dieses Thema]
about this topic

2.1.3 Multiple sluicing

In multiple sluicing contexts in German, more than one wh-phrase escapes deletion (cf. Merchant 2001). The phenomenon is shown in (5a) (with elided material crossed out); here the two wh-phrases are clause-mates. Next, (5b) shows that simple sluicing can take place across a clause boundary.

(5) German

a. Irgendjemand hat irgendetwas geerbt, aber der Karl weiß nicht
someone has something inherited but the Karl knows not
mehr [$_{CP}$ wer$_1$ was$_2$ ~~t_1 t_2 geerbt hat~~]
more who what inherited has

b. Maria hat behauptet dass sie irgendetwas geerbt hat aber Karl
Maria has claimed that she something inherited has but Karl
weiß nicht mehr [$_{CP}$ was$_1$ ~~Maria t'''_1 behauptet hat~~ [$_{CP}$ ~~t''_1 dass sie~~
knows not more what Maria claimed has that she
~~t'_1 t_1 geerbt hat~~]]
inherited has

However, when the two strategies are combined, ungrammaticality arises: Multiple sluicing is impossible when the two wh-phrases are separated by a clause boundary; see (6).

(6) German
*Irgendjemand hat behauptet, dass Maria irgendetwas geerbt hat, aber
someone has claimed that Maria something inherited has but
Karl weiß nicht mehr [$_{CP}$ wer$_1$ was$_2$ ~~t$_I$~~ ~~behauptet~~ ~~hat~~ [$_{CP}$ ~~dass~~ ~~Maria~~ ~~t$_2$~~
Karl knows not more who what claimed has that Maria
~~geerbt~~ ~~hat~~]]
inherited has

Finally, as noted by Sauerland (1999), whereas non-restructuring verbs do not permit multiple sluicing (with one wh-phrase belonging to the matrix clause, and the other one belonging to the embedded infinitive; see 7b), restructuring verbs permit such multiple sluicing (see 7a).

(7) German
a. Irgendjemand hat irgendetwas zu klauen versucht aber ich weiß
someone has something to steal tried but I know
nicht [$_{CP}$ wer$_1$ was$_2$ t$_1$ [t$_2$ ~~zu~~ ~~klauen~~] ~~versucht~~ ~~hat~~]
not who what to steal tried has

b. ?* Irgendjemand hat irgendetwas zu klauen gezögert aber ich weiß
someone has something to steal hesitated but I know
nicht [$_{CP}$ wer$_1$ was$_2$ t$_1$ [$_{CP}$ t$_2$ ~~zu~~ ~~klauen~~] ~~gezögert~~ ~~hat~~]
not who what to steal hesitated has

As before, this suggests that the complements of non-restructuring verbs involve biclausal structures (with an embedded CP), whereas restructuring verbs optionally involve monoclausal structures (without an embedded CP). Depending on the exact nature of the analysis of multiple sluicing, this argument for monoclausality may or may not be an instance of one of the arguments given above. Thus, Sauerland (1999) assumes that multiple sluicing in German involves a combination of simple wh-movement affecting one wh-phrase, and scrambling affecting the other one(s), which would make the multiple sluicing case an instance of the scrambling case, as discussed in §2.1.1. In contrast, Lasnik (2014) argues that multiple sluicing (in English) involves a combination of simple wh-movement and extraposition; adopting this analysis for German would imply that it is an instance of the extraposition case, as discussed in §2.1.2. Finally, if multiple sluicing in German does in fact indicate an exceptional (recoverability-driven) occurrence of two (or more) genuine instances of wh-movement (cf. Mer-

chant 2001; Heck & Müller 2003), it provides a fully independent argument for selective transparency of embedded infinitivals.[4]

The arguments for monoclausality given so far all involve movement; the final three arguments I want to mention here are somewhat different.

2.1.4 Compactness

Haider (2010) observes that items participating in restructuring are *compact* in the sense that other material cannot linearly intervene. Thus, as shown by the presence of unstressed pronoun fronting from the infinitive, restructuring must have taken place in (8a); and in this configuration, matrix V and embedded V are separated by an intervening adverb, yielding ill-formedness. In contrast, (8b) does not involve restructuring, and the compactness requirement is lifted.

(8) German

a. * dass es_1 keiner [t_1 zu lesen] gestern versucht hat
that it_{ACC} no-one to read yesterday tried has

b. dass der Karl [$_{CP}$ das $Buch_1$ zu kennen] gestern geleugnet hat
that the $Karl_{NOM}$ the $book_{ACC}$ to know yesterday denied has

Haider accounts for compactness by postulating a complex base-generated head analysis for restructuring. However, it looks as though many of the relevant data can be accounted for independently (see Büring & Hartmann 1996; Wurmbrand 2007; Müller 2014: ch. 3; but also Haider 2016 for a critique of PF-based accounts). In addition, the compactness requirement can be circumvented by various kinds of movement operations (verb-second, topicalization), and it does not hold in the third construction (see below; cf. Wurmbrand 2007). Thus, compactness may be an indicator of restructuring, but not without qualifications.

2.1.5 Negation

A well-known argument for monoclausality is that embedded negation can take wide scope over the matrix clause; cf. (9a) (where restructuring can take place in the presence of the restructuring verb *empfehlen* 'recommend') vs. (9b) (where restructuring is not an option with the matrix verb *auffordern* 'request').

[4] In Heck & Müller (2003), the impossibility of (6, 7b) is tied to the presence of a CP phase that precludes long-distance wh-movement of the second wh-phrase via a conspiracy of Chomsky's (2001) (PIC) and a constraint phase balance triggering intermediate movement steps.

(9) German

a. dass Maria ihm [das Buch nicht zu lesen] empfiehlt
that Maria$_{NOM}$ him$_{DAT}$ the book$_{ACC}$ not to read recommends

b. dass Maria ihn [$_{CP}$ das Buch nicht zu lesen] auffordert
that Maria$_{NOM}$ him$_{ACC}$ the book$_{ACC}$ not to read requests

(9a) can have a reading where negation takes embedded scope (and restructuring does not apply: *recommend* ≫ *not*), and a (more salient) reading where negation takes matrix scope (and restructuring has applied: *not* ≫ *recommend*). In contrast, (9b) can only have a reading with embedded scope of negation (*request* ≫ *not*), not one with wide scope of negation (**not* ≫ *request*).

2.1.6 Intonation

Finally, restructuring infinitives typically trigger a different intonational realization from non-restructuring infinitives. Whereas the latter are usually prosodically separated from the matrix clause (by an intonational break, indicated by "|"), the former usually are not. Thus, the restructuring environment in (10a) (signalled by scrambling of the embedded object in front of the matrix subject) is incompatible with an intonational break; the non-restructuring context (signalled by a violation of compactness) favors it.

(10) German

a. dass den Karl$_1$ niemand t$_1$ zu küssen versuchte
that the Karl$_{ACC}$ no-one$_{NOM}$ to kiss tried

b. dass sie | den Karl zu küssen | gar nicht erst versucht hat
that she$_{NOM}$ den Karl$_{ACC}$ to kiss PTCL not PTCL tried has

2.2 Arguments for biclausality

2.2.1 Uniformity of embedding

The first argument for biclausality of restructuring constructions with control verbs in German is a conceptual one (see Koster 1987; von Stechow & Sternefeld 1988): every control verb that permits restructuring can optionally also show up in a non-restructuring context. Thus, there is no control verb like, say, a fictive predicate *entsuchen* 'try' that would permit (11a) (where scrambling to the matrix domain has applied, signalling restructuring) but not (11b) (where compactness is violated, signalling non-restructuring).

(11) German

a. dass den Fritz$_1$ keiner [t$_1$ zu küssen] entsuchte
that the Fritz$_{ACC}$ no-one$_{NOM}$ to kiss tried

b. * dass keiner [$_{CP}$ den Fritz$_1$ zu küssen] gestern entsucht hat
that no-one$_{NOM}$ the Fritz$_{ACC}$ to kiss yesterday tried has

Deriving this implicational generalization requires additional assumptions if restructuring predicates can simply optionally involve TP-embedding, vP-embedding or VP-embedding.[5] However, the generalization follows directly if the only way to end up with such a smaller complement size is via an initial CP embedding that is then subject to some operation bringing about restructuring.

2.2.2 Licensing and interpretation of PRO

A second standard argument for biclausality of restructuring (cf., again, von Stechow & Sternefeld 1988) is that the distribution of the empty pronominal subject of control infinitives (PRO) requires the presence of a CP projection. In its original form, this argument presupposes that every verb must discharge its external θ-role in the syntax, that the external θ-role is represented by PRO, and that PRO must not be governed ("PRO theorem", cf. Chomsky 1981). The PRO theorem is not widely accepted anymore; however, in all approaches that recognize a syntactically represented non-overt external argument like PRO in control infinitives, it needs to be ensured that PRO shows up in these contexts but not in others (finite clauses, exceptional case marking (ECM) environments, raising), and simple accounts would seem to rely on the presence of a C projection.[6] As pointed out by von Stechow & Sternefeld (1988), and Sternefeld (1990), if there is no CP projection, the difference between ECM/raising and control may be blurred.

A related problem arises in approaches that do not recognize PRO for restructuring contexts (because the structure that could introduce the external argument is not present, or because the structure that could license the external argument is not present, or both) but do recognize PRO for non-restructuring contexts with the same predicate (see, e.g., Haider 2010): such a heterogenous analysis invariably requires two radically different approaches to control – e.g.,

[5] Minimally, it would seem that a designated lexical rule would have to be stipulated that derives restructuring versions of verbs from the corresponding non-restructuring versions. Such a way out is in principle unavailable if the lexicon is conceived of as a list of exceptions rather than a place where systematic generalizations can be expressed.

[6] This holds, e.g., for Adger's (2003) approach: on this view, control predicates that embed infinitival clauses (cf. Stiebels 2010 on control into finite clauses in German) select a special type of complementizer which in turn assigns a case-like feature [null] to the embedded subject that requires a non-overt realization not just of the inflectional ending, but of the whole argument DP (as PRO). Also cf. Chomsky & Lasnik (1993); Roberts (1997).

(some operation like) syntactic Agree that determines the interpretation of an embedded PRO via syntactic binding on the one hand (see, e.g., Landau 2000), and (some operation like) functional composition that brings about the identification of an argument of the matrix predicate with the external argument of the embedded predicate on the other hand (see, e.g., Stiebels 2007). None of these two ways to identify argument positions of two verbs can be straightforwardly derived from the other; e.g., minimality may predict object control in the syntax in the unmarked case (see, e.g., Hornstein 2001), whereas simple lexical stipulation determines whether subject or object control takes place in the case of function composition.[7] Crucially, given the independence of the two means to identify argument positions in control, the option of control shift with restructuring is wrongly predicted to be possible. Control shift can take place in various contexts in German (e.g., influenced by passivization of the embedded verb, or in the presence of certain modal verbs; see Růžička 1983; Wurmbrand 2002; Stiebels 2007). However, this phenomenon never shows up with restructuring: there is no matrix verb that triggers object control when it embeds a non-restructuring infinitive, but subject control when it embeds a restructuring infinitive (or vice versa).

2.2.3 Absence of new binding domains

The third argument for biclausal structures is based on the observation that restructuring does not create new binding domains. Thus, an accusative object reflexive in a subject control infinitive (*sich* in 12a,b) can never pick a dative object of the matrix verb (*ihm* in 12a,b) as an antecedent, even if the matrix verb permits restructuring (*versprechen* in 12a,b). This is accounted for if a reflexive pronoun needs to participate in an Agree relation with its antecedent (cf. Reuland 2001; 2011, Fischer 2004, and Hicks 2009, among others), and restructuring environments involve a full clausal CP structure across which Agree is blocked.

(12) German

a.	dass	Karl_1	ihm_2	(PRO_1)	sich_1	zu waschen	versprochen	hat
	that	Karl_{NOM}	him_{DAT}		REFL	to wash	promised	has
b.	* dass	Karl_1	ihm_2	(PRO_1)	sich_2	zu waschen	versprochen	hat
	that	Karl_{NOM}	him_{DAT}		REFL	to wash	promised	has

[7] Thus, an object control verb like *empfehlen* 'recommend' can be assumed to have a simplified entry like λP λy λx recommend(x,y,P(y)), whereas a subject control verb like *versprechen* 'promise' could be specified as λP λy λx promise(x,y,P(x)) – here the only relevant difference is whether the complement predicate applies to the object variable (y) or to the subject variable (x) (after function composition has opened up internal argument position(s) of the embedded predicate via λ conversion plus λ prefixation).

In contrast, if there is no CP present in restructuring environments, it is not obvious how the ill-formedness of (12b) can be derived. The reason is that an accusative object reflexive *can* pick a dative object of the same verb as an antecedent for many speakers of German (see the empirical investigation reported in Sternefeld & Featherston 2003; Featherston & Sternefeld 2003, which contradicts earlier informal judgements reported in Grewendorf 1988); cf. (13).

(13) German
dass Karl$_1$ ihm$_2$ sich$_{1/2}$ im Spiegel gezeigt hat
that Karl$_{\text{NOM}}$ him$_{\text{DAT}}$ REFL in the mirror shown has

In monoclausal approaches to restructuring where the embedded infinitive lacks PRO$_1$ in (12a,b) because it is always either part of a complex verb (as in Haider 2010) or is a bare VP (Sternefeld 2006), the problem is evident: the structural relations between *ihm*$_2$ and *sich*$_2$ in (12b) and in (13) are nearly indistinguishable on this view. However, accounting for the ill-formedness of (12b) also poses a challenge under approaches where the restructuring complement can be a vP or TP containing PRO (Wurmbrand 2001). The reason is that the option of reflexive binding of *sich*$_1$ by the matrix subject *Karl*$_1$ in (13) shows that reflexivization can take place across what one might think should be an intervening potential binder (viz., the indirect object *ihm*$_2$ in 13). The only way out here, it seems, would be to stipulate that external arguments (PRO$_1$ in 12b) intervene for Agree-based reflexive binding in a way that internal arguments (*ihm*$_2$ in 13) do not. However, not even this step would eventually suffice. As shown in (14a), an intervening external argument DP *can* be skipped with PP-internal reflexives in an ECM construction headed by *lassen* 'let' or *sehen* 'see' (see Reis 1976; Grewendorf 1983; Fanselow 1987; Gunkel 2003; Barnickel 2014). This is never possible across a finite clause boundary; see (14b). Crucially, it is also never possible with control infinitives (see 14c), even when restructuring must have taken place (because unstressed pronoun fronting to the matrix domain has occurred; see 14d).

(14) German

a. dass Maria$_1$ [$_{\text{TP}}$ Paul$_2$ [$_{\text{PP}}$ bei sich$_{1/2}$] schlafen] lässt
that Maria$_{\text{NOM}}$ Paul$_{\text{ACC}}$ with REFL sleep lets

b. dass Maria$_1$ sagt [$_{\text{CP}}$ dass Paul$_2$ bei sich$_{*1/2}$ schlafen kann]
that Maria$_{\text{NOM}}$ says that Paul$_{\text{NOM}}$ with REFL sleep can

c. dass Maria$_1$ Paul$_2$ [$_{\text{CP}}$ PRO$_1$ [$_{\text{PP}}$ bei sich$_{1/*2}$] zu schlafen]
that Maria$_{\text{NOM}}$ Paul$_{\text{DAT}}$ with REFL to sleep
verspricht
promises

d. dass Maria$_1$ es$_3$ Paul$_2$ [$_{CP}$ PRO$_1$ t$_3$ [$_{PP}$ bei sich$_{1/*2}$] zu
that Maria$_{NOM}$ it$_{ACC}$ Paul$_{DAT}$ with REFL to
organisieren] verspricht
organize promises

Thus, whatever ultimately accounts for the fact that PP-internal reflexives (in contrast to arguments of the embedded V) can skip over the subject of the infinitive, it is clear that such long-distance reflexivization is blocked by a CP phase boundary. The data then show that a CP is always present with control verbs (restructuring and non-restructuring), and not present with ECM predicates.

2.2.4 Unstressed pronoun fronting

In §2.1.1, unstressed pronoun fronting from a restructuring infinitive was presented as an argument in support of monoclausality, based on the conclusion that the presence of a CP would lead to a violation of locality constraints on movement. Interestingly, unstressed pronoun fronting also provides an argument in support of biclausality, more specifically, the presence of a CP in restructuring environments. Unstressed pronouns must undergo fronting to a position that can only be preceded by a subject DP, which can then be assumed to have undergone optional EPP-driven movement to SpecT; cf. (15a,b) (see Müller 2001; Fanselow 2004). I assume that unstressed pronouns end up in an outer Specv position (more specifically, at the left edge of vP), where they precede DP and PP arguments, including scrambled ones (see 15a–c), adverbials (see 15d), and the base position of subjects (see 15a).

(15) German

a. dass es$_1$ die Maria dem Fritz t$_1$ gegeben hat
that it$_{ACC}$ the Maria$_{NOM}$ the Fritz$_{DAT}$ given has

b. dass die Maria es$_1$ dem Fritz t$_1$ gegeben hat
that the Maria$_{NOM}$ it$_{ACC}$ the Fritz$_{DAT}$ given has

c. * dass die Maria dem Fritz es$_1$ gegeben hat
that the Maria$_{NOM}$ the Fritz$_{DAT}$ it$_{ACC}$ given has

d. * dass die Maria wahrscheinlich es$_1$ dem Fritz t$_1$ gegeben hat
that the Maria$_{NOM}$ probably it$_{ACC}$ the Fritz$_{DAT}$ given has

Complements of non-control (obligatory) restructuring verbs do not have sufficient space for unstressed pronoun fronting. This is shown for auxiliaries in (16a), for raising verbs in (16b), and for ECM verbs in (16c), all of which become well formed if the unstressed pronoun *es* 'it' undergoes longer movement to a position directly after *sie* 'she'.

(16) German

a. * dass sie mir$_1$ schon letzte Woche [t$_1$ es$_2$ gegeben] hat
that she$_{NOM}$ me$_{DAT}$ already last week it$_{ACC}$ given has

b. * dass sie mir schon letzte Woche [es$_2$ zu lesen] schien
that she$_{NOM}$ me$_{DAT}$ already last week it$_{ACC}$ to read seemed

c. * dass sie mich schon letzte Woche [es$_1$ lesen] ließ
that she$_{NOM}$ me$_{ACC}$ already last week it$_{ACC}$ read let

The relevant observation now is that there is a vast improvement with the unstressed pronoun in the embedded domain in the case of control constructions. As shown in (17a,b), restructuring contexts (indicated here by the option of unstressed pronoun fronting of the dative pronoun) seem to provide sufficient space for separate unstressed pronoun fronting (here applying to the accusative pronoun, which of course could also accompany the dative pronoun in the matrix domain). (17b involves the third construction; see the next subsection.)

(17) German

a. dass sie mir$_1$ schon letzte Woche [t$_1$ es$_2$ zu geben]
that she$_{NOM}$ me$_{DAT}$ already last week it$_{ACC}$ to give
versucht hat
tried has

b. dass sie mir$_1$ schon letzte Woche versucht hat [t$_1$ es$_2$ zu
that she$_{NOM}$ me$_{DAT}$ already last week tried has it$_{ACC}$ to
geben]
give

This indicates that there is more structure in control infinitives; assuming raising and ECM environments to involve embedded TPs (Fanselow 1991), the evidence suggests that a CP is required for all cases of unstressed pronoun fronting in German, and that such a CP is therefore present in restructuring contexts with control predicates.[8]

[8]Note that the argument here is indirect since the *actual* landing site of unstressed pronoun fronting, by assumption, is a left-peripheral position in vP. The point is that such movement is evidently only licensed in the presence of a higher CP. There are various possibilities to derive this – including, e.g., postulating an inheritance of the relevant features from C, as suggested in Chomsky (2008); Richards (2007), or postulating that unstressed pronouns must undergo Agree with C. Ultimately, it seems to be a fact about unstressed pronouns (perhaps, more generally, Wackernagel-oriented processes) that they depend on the presence of a CP domain, however this is derived.

2.2.5 The third construction

The fifth and final argument in support of a CP projection for restructuring in German involves the so-called third construction, i.e., constructions involving a combination of leftward scrambling or unstressed pronoun fronting out of a restructuring complement, and rightward extraposition of the restructuring complement itself (see den Besten & Rutten 1989). As noted in §2.1.2, CP, PP, and (to some extent) DP can undergo extraposition in German; however, verbal projections (vP, VP, TP) cannot do so.[9] CP extraposition is shown in (18a,b) (for finite clauses and infinitives, respectively).

(18) German

a. dass er gesagt hat [$_{CP}$ dass es regnet]
that he$_{NOM}$ said has that it$_{NOM}$ rains

b. dass sie versucht hat [$_{CP}$ PRO zu schlafen]
that she$_{NOM}$ tried has to sleep

The impossibility of TP extraposition is illustrated by (19a,b) (based on the assumption that complements of ECM verbs have TP status).

(19) German

a. * dass ich gesehen habe [$_{TP}$ den Mann das Buch lesen]
that I$_{NOM}$ seen have the man$_{ACC}$ the book$_{ACC}$ read

b. * dass sie ließ [$_{TP}$ ihn schlafen]
that she$_{NOM}$ let him$_{ACC}$ sleep

The data in (20a–d) show that vP/VP cannot undergo extraposition either.

(20) German

a. * dass sie t_1 hat [$_{VP}$ gearbeitet]
that she$_{NOM}$ has worked

b. * dass er t_1 hat [$_{VP}$ das Buch gelesen]
that he$_{NOM}$ has the book$_{ACC}$ read

c. * dass er t_1 wird [$_{VP}$ das Buch lesen]
that he$_{NOM}$ will the book$_{ACC}$ read

d. * dass sie hatte [t_1 wollen/gewollt [$_{VP}$ das Buch lesen]]
that she$_{NOM}$ had want/wanted the book$_{ACC}$ read

[9] I hasten to add that this only holds for *Standard German*; see Haegeman & van Riemsdijk (1986); Bader & Schmid (2009); Salzmann (2011; 2013a,b) for variation in other varieties of German, for which the argument to be presented below can therefore not be made.

Against this background, it can be noted that extraposition *is* possible in the third construction, i.e., with scrambling or unstressed pronoun fronting from extraposed restructuring infinitives; see (21a,b) (with *versuchen* as a matrix verb), (21c) (with *versprechen* as a matrix verb), and (21d) (with the object control verb *empfehlen*).[10]

(21) German

a. dass sie ihn$_2$ t$_1$ versucht [$_{CP_1}$ PRO t$_2$ zu küssen]
that she$_{NOM}$ him$_{ACC}$ tries to kiss

b. dass sie das Buch$_2$ t$_1$ versucht hat [$_{CP_1}$ PRO t$_2$ dem Mann$_{DAT}$ zu geben]
that she$_{NOM}$ the book tried has the man to give

c. dass es$_2$ Maria t$_1$ (dem Fritz$_3$) verspricht [$_{CP_1}$ PRO t$_1$ zu lesen]
that it$_{ACC}$ Maria the Fritz$_{DAT}$ promises to read

d. dass es$_2$ Fritz ihr$_3$ t$_1$ empfohlen hat [$_{CP_1}$ PRO t$_1$ zu lesen]
that it$_{ACC}$ Fritz$_{NOM}$ her$_{DAT}$ recommended has to read

This strongly suggests that the extraposed item is a CP. If the third construction were to involve extraposition of a VP (as assumed by Wöllstein-Leisten 2001 and Haider 2010), or of a vP or TP, ungrammaticality would be expected to result throughout in (21).[11]

[10] (21c) and (21d) show that a control verb may take an additional DP argument (DP$_3$) in the third construction. Kiss (1995: 110) claims that examples of this type are impossible; however, I would like to contend that the problem is due to parsing problems: DP$_2$ and DP$_3$ are extremely similar in his examples.

[11] There is in fact one principled exception to the generalization that VP extraposition is impossible in Standard German. In the Ersatzinfinitiv construction, VP extraposition is possible (in fact, obligatory); see (i).

(i) dass sie das Buch hatte lesen wollen
that she$_{NOM}$ the book$_{ACC}$ had read want

I contend that this is the exception that proves the rule. In Ersatzinfinitiv constructions, existing constraints are *violated* in optimal forms so as to satisfy higher-ranked requirements (see Schmid 2005); this holds for morphological selection among verbs (with an infinitive form showing up where a participle would be expected) in the same way that it does for linearization. Note that extraposition in the third construction, unlike what is the case with the Ersatzinfinitiv construction, is strictly optional, and not a repair operation like Ersatzinfinitiv formation.

2.3 Interim conclusion

Summarizing so far, there is evidence both for a truly biclausal (CP) analysis and for a monoclausal analysis of restructuring constructions with control verbs in German. Accordingly, this state of affairs is difficult to account for both in purely monoclausal and purely biclausal approaches. In *monoclausal approaches* (see Geilfuß 1988; Haider 1993; 2010; Kiss 1995; Wurmbrand 2001; 2007; 2015b; Sternefeld 2006, and many others), the evidence for biclausality poses problems that typically require construction-specific assumptions complicating the overall analysis; effects attributable to the presence of a CP projection must be imitated in some other way if a CP projection cannot be present. In *biclausal approaches* (see Baker 1988; Sternefeld 1990; Müller & Sternefeld 1995; Sabel 1996; Roberts 1997; Hinterhölzl 1999, and Koopman & Szabolcsi 2000), the evidence for monoclausality poses problems that typically require extremely abstract interactions of movement operations lacking independent motivation (plus, in many cases, additional stipulations); effects attributable to the absence of a CP projection must be captured by mechanisms that permit selective disregard of the additional structure. What is needed, then, is a way to both have your cake and eat it.

Coanalysis approaches (as pursued in Huybregts 1982; Bennis 1983; Haegeman & van Riemsdijk 1986; Di Sciullo & Williams 1987; Sadock 1991; Pesetsky 1995) are a case in point. Here, both types of evidence can be accommodated because monoclausal and biclausal structures can exist simultaneously. However, these approaches are typically quite unconstrained, and often not fully worked out (especially where restructuring is directly addressed); and it is sometimes not clear why one process would target one kind of structure rather than the other one. That leaves, finally, traditional *reanalysis approaches* (see Ross 1967: Ch. 3, Evers 1975, Rizzi 1982, Aissen & Perlmutter 1983, and von Stechow & Sternefeld 1988): the simple idea underlying these approaches is that a structure that is initially biclausal is reduced to a monoclausal one, via some form of structure removal. The only problem with all the classical reanalysis approaches is that they rely on transformations that are (a) ad hoc, (b) not constrained in interesting ways, and (c) not embedded into a general system of elementary, primitive operations manipulating syntactic structure. The claim that I would like to argue for in what follows is that an analysis based on an elementary, restrictive operation Remove makes it possible to pursue a simple, principled reanalysis approach to restructuring in German.[12]

[12] Thus, I take issue with the claim in Haider (2010: 309) that "radical clause union […] cannot be achieved derivationally since derivations do not destroy or eliminate structures": they do.

3 Structure removal

Suppose that syntactic derivations employ two elementary operations modifying representations: in addition to an operation that *builds* structure – *Merge* (Chomsky 2001; 2008; 2013) –, there is a complementary operation that *removes* structure: *Remove*. In Müller (2016; 2017; 2018), an approach to structure removal based on this operation has been argued to systematically account for cases where there is empirical evidence for conflicting representations (that movement cannot plausibly be invoked to account for). The basic premise is that if Remove exists as the mirror image of Merge, it is expected to show similar properties and obey identical constraints. The assumptions made about Merge are the following. First, Merge is feature-driven.[13] It is triggered by designated [•F•] features, which are ordered on lexical items (see Heck & Müller 2007, Abels 2012, Stabler 2013, Georgi 2014, among others); F here is a variable over categorial features (primarily for external Merge) and movement-related features (like wh, top) that trigger internal Merge. Once a feature has brought about an operation, it is discharged, and disappears. Second, Merge may apply to heads or phrases. This necessitates diacritics on structure-building features: [•F_0•], [•F_2•] for heads and phrases, respectively. Third, Merge obeys the strict cycle condition in (22) (see Chomsky 1973; 1995; 2001; 2008; also cf. Safir 2010; 2015 for this specific version). Based on the concept of domain in (23), the strict cycle condition in (22) blocks operations that exclusively affect positions contained in embedded phrases. Fourth and finally, Merge can be external or internal.

(22) *Strict cycle condition* (SCC):
Within the current XP α, a syntactic operation may not exclusively target some item δ in the domain of another XP β if β is in the domain of α.

(23) *Domain* (Chomsky 1995):
The domain of a head X is the set of nodes dominated by XP that are distinct from and do not contain X.

The assumptions about Remove are identical. First, Remove is feature-driven. It is triggered by designated [–F–] features, which are ordered on lexical items (and can be interspersed with features for structure building). Second, Remove may apply to heads or phrases, so there is a feature [–F_0–] for heads, and a feature [–F_2–] for phrases. If Remove applies to a phrase (via [–F_2–] on a head that triggers the operation), it takes out a whole subtree. Removal of phrases in the

[13] This corresponds to Chomsky's original view but is at variance with his more recent assumption that Merge comes free; see, e.g., Chomsky (2013).

course of the derivation has been argued to take place with external arguments in passive constructions (see Müller 2016), with internal arguments in applicative constructions (see Müller 2017), and with VPs and TPs in various kinds of ellipsis constructions (see Murphy 2015; Murphy & Müller 2016). In what follows, I will exclusively focus on Remove applying to a head (via [–F_0–]) – this is the operation that I assume to take place in restructuring environments. Third, Remove obeys the strict cycle condition in (22). And fourth, Remove can be external or internal. Here I focus on internal Remove, i.e., operations that remove part of the current syntactic structure.[14]

If an [–F_0–] feature on some head X is discharged, it removes the head Y of a projection in the minimal domain of X. Given a bare phrase structure approach, a head's projection does not exist independently of the head. This means that by taking away the head Y, the whole projection line of Y up to YP is removed – but only this: specifiers and complements of Y are not affected by removal. The question then is what happens with the material that was originally included in the removed projection, and that is temporarily split off from the current tree after removal of the head and its projection. In Müller (2018), it is argued that such items are reassociated with the main projection, i.e., with the projection of the head responsible for structure removal, in a way that is maximally structure-preserving, maintaining earlier c-command and linearization relations as much as possible.[15] Predecessors or alternatives of removal of heads by [–F_0–] features (and, consequently, the projections of these heads) include tree pruning (see Ross 1967: Ch. 3); Chomsky's (1981) proposal of S-bar deletion with ECM verbs (and in subject extraction environments – a new version of this latter approach is suggested in Chomsky (2015b: 24) and argued to crucially involve removal of syntactic structure in Hornstein 2014);[16] the approaches to head movement developed in Heycock & Kroch (1994) and Stepanov (2012); the approach to pruning

[14]External Remove may initially look like an unusual concept since such an operation removes items that are not yet part of the current tree; see Müller (2016; 2017) for discussion of some relevant cases.

[15]Note that reassociation is not an instance of Merge: it only applies to phrases (not to heads), the external/internal distinction does not make sense here, and, perhaps most importantly, reassociation is not feature-driven; rather, it is an operation triggered by the need to reintegrate material into the present tree that is temporarily unattached as a consequence of Remove.

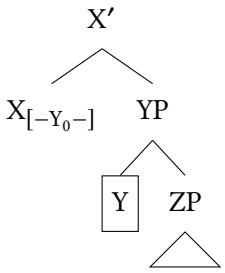
[16]It should be noted, though, that although it is uncontroversial that the approach in Chomsky (2015b) relies on syntactic (rather than, say, phonological) deletion, it is not entirely clear what exactly is subject to removal. Further elaboration in Chomsky (2015a) suggests that Chomsky, despite explicitly proposing a rule "C→ ∅", might have in mind a relativization of the deletion operation to certain kinds of features of C (e.g., the "phase-head feature of C"). However, as argued in Müller (2017), given that syntactic categories are to be viewed as sets of features, this difference would be purely quantitative rather than qualitative.

of ∅-affixes in Embick (2010); the approach to cases of XP movement that can circumvent intervention effects proposed in Heck (2016); and, last but not least, Pesetsky's (2016) exfoliation transformation, which removes embedded CP and TP shells.[17]

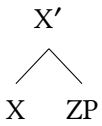
In what follows, I will illustrate the working of head removal by some abstract sample derivations. Consider first the case where the head Y of a complement YP is removed. For now, I assume that Y has a complement ZP but does not have a specifier; I will address this latter scenario momentarily. As shown in (24a), X first combines with YP (triggered by [•Y•] on X); after [•Y•] is discharged and Merge(X,YP) has taken place, $[-Y_0-]$ becomes accessible and triggers removal of the YP shell before being discharged; see (24b). As a consequence, ZP, which is initially split off the tree after YP shell removal, is reassociated with the projection of X in a maximally structure-preserving way: it becomes the new complement of X, which maintains all earlier c-command relations. Note that if X were to be equipped with a removal feature $[-Z_0-]$ instead of $[-Y_0-]$ in (24a), removal of the ZP shell could not take place in the presence of the intervening YP projection, due to the strict cycle condition. However, if X were to be equipped with $[-Z_0-]$ in addition to $[-Y_0-]$ in (24a), and if $[-Z_0-]$ were ranked below $[-Y_0-]$ on the list of operation-triggering features on X, the ZP shell could next be removed on the basis of (24b). In other words: Remove can apply recursively. (This will become relevant in the analysis of restructuring given in the next section.)

(24) Remove and heads: complements w/o specifiers

a. Merge($X_{[\bullet Y\bullet]>[-Y_0-]}$,YP):



b. Remove($X_{[-Y_0-]}$,Y):



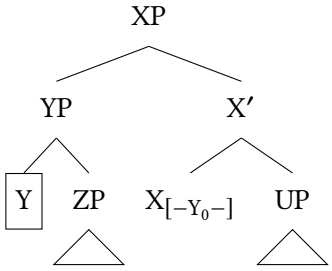
In the same way, Remove applying to heads can also affect a specifier. The operation is shown in (25), where X has first merged with a UP complement; again, an

[17] Exfoliation is similar to Remove applying to heads, but differs from it in some important respects, e.g., by being inherently less local (it takes place across phase boundaries), by not being feature-driven (but instantiating a last resort operation), and by never applying recursively. See Müller (2018) for a more elaborate comparison of the two approaches to shrinking trees.

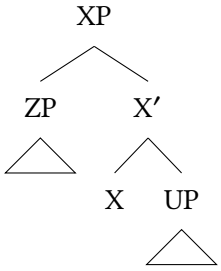
XP included in the specifier (here: ZP) cannot be targeted by the operation, due to the strict cycle condition. ZP reassociates with the X projection as a specifier, in a maximally order-preserving way.[18]

(25) Remove and heads: specifiers w/o specifiers

a. Merge($X'_{[\bullet Y \bullet] > [-Y_0-]}$,YP):



b. Remove($X'_{[-Y_0-]}$,Y):



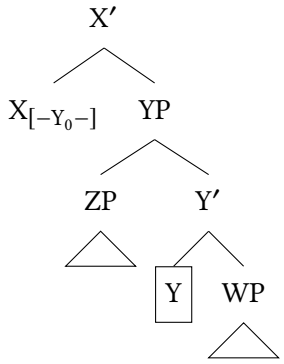
Next consider the situation where a complement projection YP is removed via $[-Y_0-]$ on X, but where the difference to (24) is that Y takes both a complement (WP) and a specifier (ZP). Again, the null hypothesis is that after YP shell removal, WP and ZP reassemble in their original hierarchical and linear order in the XP domain, so that structural changes induced by the operation are minimized – recall that a basic property underlying Remove operations is that they change embedded structures as little as possible. (26) shows how a Remove operation triggered by X and targeting the head of X's complement Y reassociates Y's specifier (ZP) and complement (WP) with the projection of X: ZP becomes a new specifier of X, and WP replaces the original YP in the complement position.[19]

[18] In principle, given an appropriate feature $[-U_0-]$, X could also have removed the UP shell in the presence of a specifier YP, in accordance with the strict cycle condition, in what is essentially a removal analogue to tucking-in derivations with Merge; see Richards (2001).

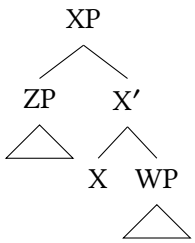
[19] Two remarks. First, it is clear that the earlier c-command relation of X and ZP *is* reversed by reassociation of ZP as X's specifier. Still, this qualifies as the best option since the alternative – reintegrating ZP as a specifier of WP – would (a) change a c-command relation into a dominance relation, and (b) carry out changes in a domain that should not be accessible, given the strict cycle condition. Second, the question arises of what happens if X independently has a feature triggering Merge of a specifier. There are two possibilities: Either this specifier is already in place, or it is merged later. The second case is straightforward; the specifier will be merged on top of the existing structure. As for the first case, ZP will have to be reassociated below the inherent specifier of X, so as to maximize structure preservation. Thus, the outcome is identical.

(26) Remove and heads: complements with specifiers

a. Merge($X_{[\bullet Y \bullet]>[-Y_0-]}$,YP):



b. Remove($X_{[-Y_0-]}$,Y):

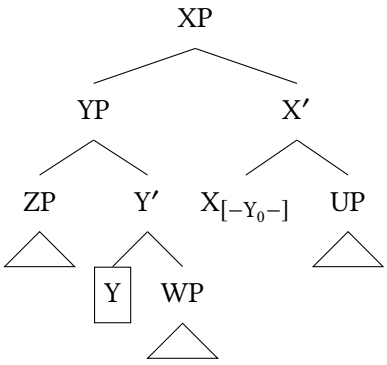


The derivation in (26) illustrates a non-trivial property of Remove operations applying to heads that take a complement and a specifier: ZP undergoes dislocation *without movement* (i.e., without internal Merge of ZP in 26b). This will play a role below.

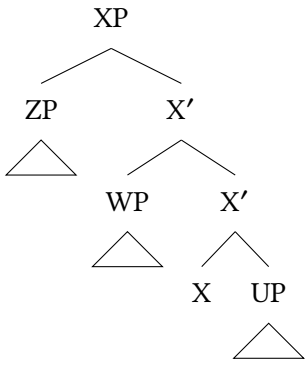
Finally, for the sake of completeness, the scenario where the head (Y) of a specifier (YP) is removed that takes both a complement (WP) and a specifier (ZP) is illustrated in (27). As before, ZP and WP are reassociated with X's projection in a way that maximally maintains earlier c-command and linearization relations, and here this implies that ZP and WP become outer and inner specifiers of X, respectively.

(27) Remove and heads: specifiers with specifiers

a. Merge($X'_{[\bullet Y \bullet]>[-Y_0-]}$,YP):



b. Remove($X'_{[-Y_0-]}$,Y):



Overall, what emerges is a principled approach to reanalysis by structure removal, which is also restrictive, due to the strict cycle condition. The patterns in (24–27) can all be shown to underlie syntactic constructions exhibiting evidence for conflicting structure assignments that are unrelated to restructuring infinitives. For instance, removal of specifier heads with complements and specifiers, as in (27), is argued in Müller (2018) to account for conflicting structure assignments to complex prefield constructions in German (viz., as topicalized headless VPs and as multiple specifiers of C); removal of complement and specifier heads with complements but no specifiers, as in (24) and (25), is argued in Müller (2015) and Puškar (2016) to account for conflicting evidence for nominals as DPs or NPs in Circassian and Serbo-Croatian, respectively, and in Korsah & Murphy (2017) to account for the presence or absence of clausal determiners in Kwa; and removal of complement heads with specifiers, as in (26), is argued in Schwarzer (2016) to account for conflicting evidence concerning the size of *tough*-movement constructions in English and German. (In addition, Dschaak 2017 develops an account of restructuring in Russian along the lines of the present proposal.) In the next section, I develop an approach to restructuring that accounts for the conflicting evidence laid out in §2. I will argue that the evidence for biclausality involves environments before removal of heads, and the evidence for monoclausality involves environments after removal. Removal typically takes place with complements (as in 24 and 26), but in the context of discussing the third construction, I will also argue that it can involve specifiers (as in 25 and 27).

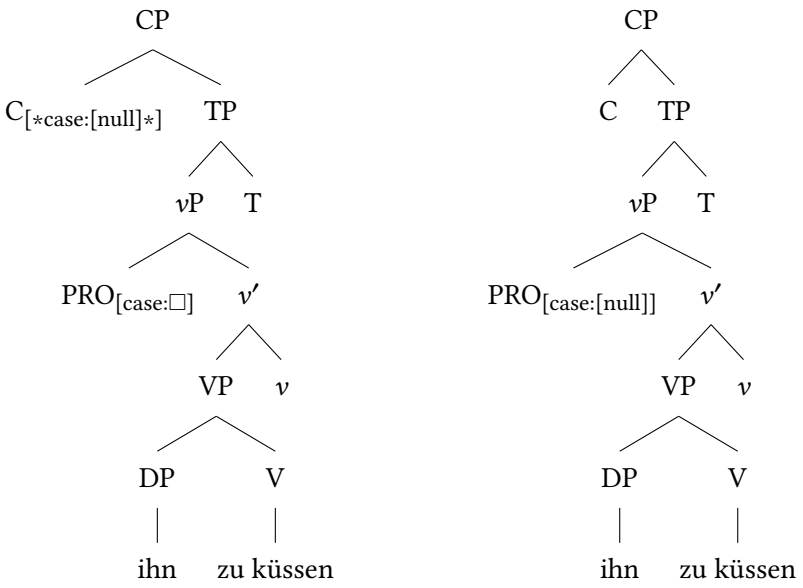
4 Analysis

4.1 Structure removal in infinitival complements

Suppose that all control verbs take CP complements. The special property of restructuring control verbs then is that they can subsequently remove CP and TP layers, yielding derived vP complements.[20] More specifically, I suggest that evidence for biclausality involves a CP structure before removal. Thus, the relevant operations that are indicative of biclausality are counter-bled and counter-fed by Remove. In contrast, evidence for monoclausality involves a vP structure after removal. Consequently, the relevant operations that are indicative of monoclausality are bled and fed by Remove. The derivation of a restructuring control infinitive is shown in Figures 9.1 and 9.2. In Figure 9.1a, infinitival C is merged with

[20] In principle, it is possible to introduce yet more subtle distinctions, with different degrees of removal eventually yielding different final output structures for the infinitival complements; see Fanselow (1991); Wurmbrand (2001; 2015b). Also cf. the remark on long-distance passivization in footnote 28 below.

a TP containing an infinitival V, an object DP that has been assigned accusative case by v, and a PRO subject that does not yet have case. Next, in Figure 9.1b, (cf. §2.2.2), infinitival C for control environments can value the infinitival subject with null case (see footnote 6); I take this to be an instance of Agree.[21]



(a) *Merge* ($C_{[\bullet T \bullet],[*case:[null]*]}$, TP) (b) *Agree* ($C_{[*case:[null]*]}$, $PRO_{case:\square}$)

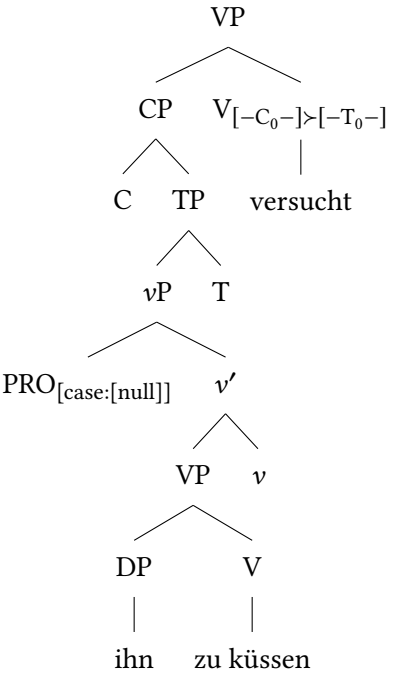
Figure 9.1: Control infinitives

If restructuring does not take place, that is all there is to say. However, if the matrix control predicate has the restructuring property, the derivation proceeds as in Figure 9.2. The lexical property that characterizes a restructuring verb in the present approach is that a $[-C_0-]$ feature and a $[-T_0-]$ feature can be added at the bottom of its stack of operation-triggering features. If this happens, the Merge operation combining V and CP (triggered by a [•C•] feature that uniformly characterizes control verbs) in Figure 9.2a is followed by recursive removal – first of the CP shell (cf. Figure 9.2b), and then of the TP shell (cf. Figure 9.2c).

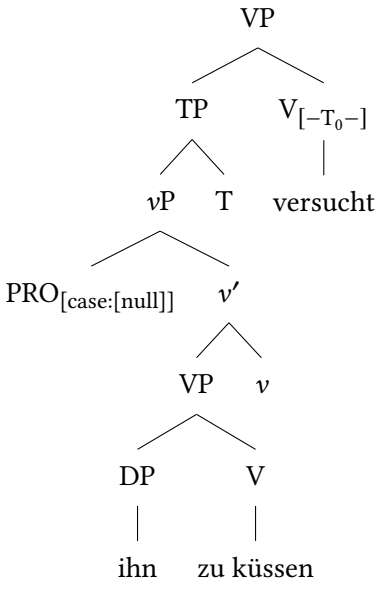
The end result is a proper monoclausal structure.[22]

[21]Here, asterisks indicate that a feature triggers an Agree operation ([∗F∗]). Also, since there is no obligatory EPP feature for German T, there is no reason to assume that PRO must undergo movement to SpecT; it is licensed by C in its in situ (Spec*v*) position.

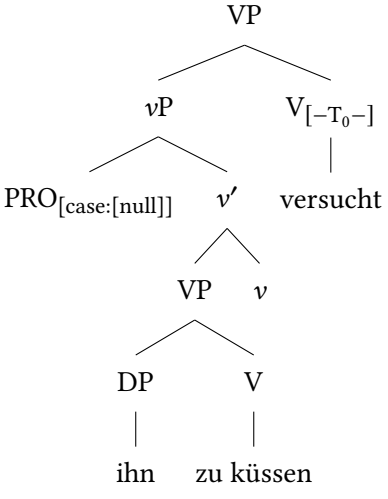
[22]Instantiation of the features for head removal on restructuring control verbs is optional, and it turns out that hardly any restrictions are needed to guarantee only correct outcomes. If the order of the two features on V is reversed ($V_{[\bullet C \bullet]>[-T_0-]>[-C_0-]}$), there can be no removal



(a) *Merge* ($V_{[\bullet C \bullet]>[-C_0-]>[-T_0-]}$, CP)



(b) *Remove* ($V_{[-C_0-]>[-T_0-]}$, CP)



(c) *Remove* ($V_{[-T_0-]}$, TP)

Figure 9.2: Restructuring

4.2 Deriving evidence for biclausality

As noted above, the operations that presuppose the presence of CP are counter-bled and counter-fed by structure removal: removal simply comes too late to bleed or feed operations that are indicative of the CP layer. Let me go through the evidence one by one. First, consider *uniformity of embedding* (§2.2.1). Given that features for removal are optional, the implicational generalization that all control verbs that permit restructuring are also compatible with non-restructuring complements is derived without further ado. The only way to reach vP is via an initial CP: Thus, Remove counter-bleeds feature-driven external Merge.

Second, as for the *licensing and interpretation of PRO* (§2.2.2), PRO is licensed via Agree with an infinitival C that assigns null case to it. Once null case is assigned, it cannot be taken away again. Thus, it does not matter that the context in which PRO can be licensed (viz., a CP) is ultimately destroyed by removal: Remove counter-bleeds PRO licensing.

Let me turn next to the *absence of new binding domains* after restructuring (§2.2.3). Assuming that reflexives are licensed by Agree operations which are blocked by a CP boundary, a reflexive will have its index fixed once the minimal CP is reached. Subsequent structure removal can neither lead to new binding options by adding a binding index on a reflexive if new potential antecedents are around,[23] nor can it undo existing binding indices on a reflexive: Remove counter-feeds new binding of reflexives and counter-bleeds old binding of reflexives.

Fourth, concerning the evidence based on *unstressed pronoun fronting* (§2.2.4), recall that an unstressed pronoun moves to the left edge of vP, but must be licensed in this position by C (perhaps as an instance of Agree, as suggested in footnote 8). Subsequent removal of CP and TP comes too late to block the licensing: Remove counter-bleeds unstressed pronoun fronting.

Fifth, consider the argument based on *the third construction* (§2.2.5): Extraposition of a restructuring infinitive is indicative of its CP status because only CP can undergo extraposition in German; TP, vP, and VP cannot do so. This implies that CP extraposition takes place *before* structure removal; otherwise the

of TP (because of the strict cycle condition), and no removal of CP either (because $[-C_0-]$ is not active before $[-T_0-]$ is discharged). If the matrix verb bears $[-T_0-]$ but not $[-C_0-]$, restructuring also cannot take place (because of the strict cycle condition). Finally, if only $[-C_0-]$ is instantiated, restructuring to TP size would be expected. To avoid such an outcome, it can be assumed that $[-T_0-]$ and $[-C_0-]$ are tied because they are part of the same phase; also see Pesetsky (2016). (That said, most of the evidence for monoclausality would not necessarily be incompatible with a TP status of the complement; the crucial requirement is the absence of CP.)

[23] Also note that unlike English, German does not allow for movement producing new binding options; cf. Barss (1986) vs. Frey (1993) and Büring (2005).

possibility of extraposition would not be explained. For the sake of concreteness, suppose that rightward movement is triggered by an optional designated feature, say [∘X∘] (with X ∈ {C, P, D} in German). A relevant part of the derivation of a sentence like (21a) is shown in Figure 9.3. First, the infinitival CP is merged to the left of V (see Figure 9.3a); then it undergoes extraposition, which I assume to target a right-peripheral specifier position (see Figure 9.3b); but note that assuming extraposition to involve right-adjunction would not substantially change things). In the next two steps, the CP and TP shells are successively removed (see Figure 9.3c,d).

As for the steps in Figure 9.3c,d, recall that there is no problem with Remove affecting specifiers (or adjuncts) rather than complements (cf. 25 and 27). As a matter of fact, there is clear independent evidence for the general possibility of restructuring with specifiers in German. Examples like (28a,b), where scrambling takes place from a subject infinitive, are entirely unproblematic (28b may involve a derived subject, but 28a certainly does not).

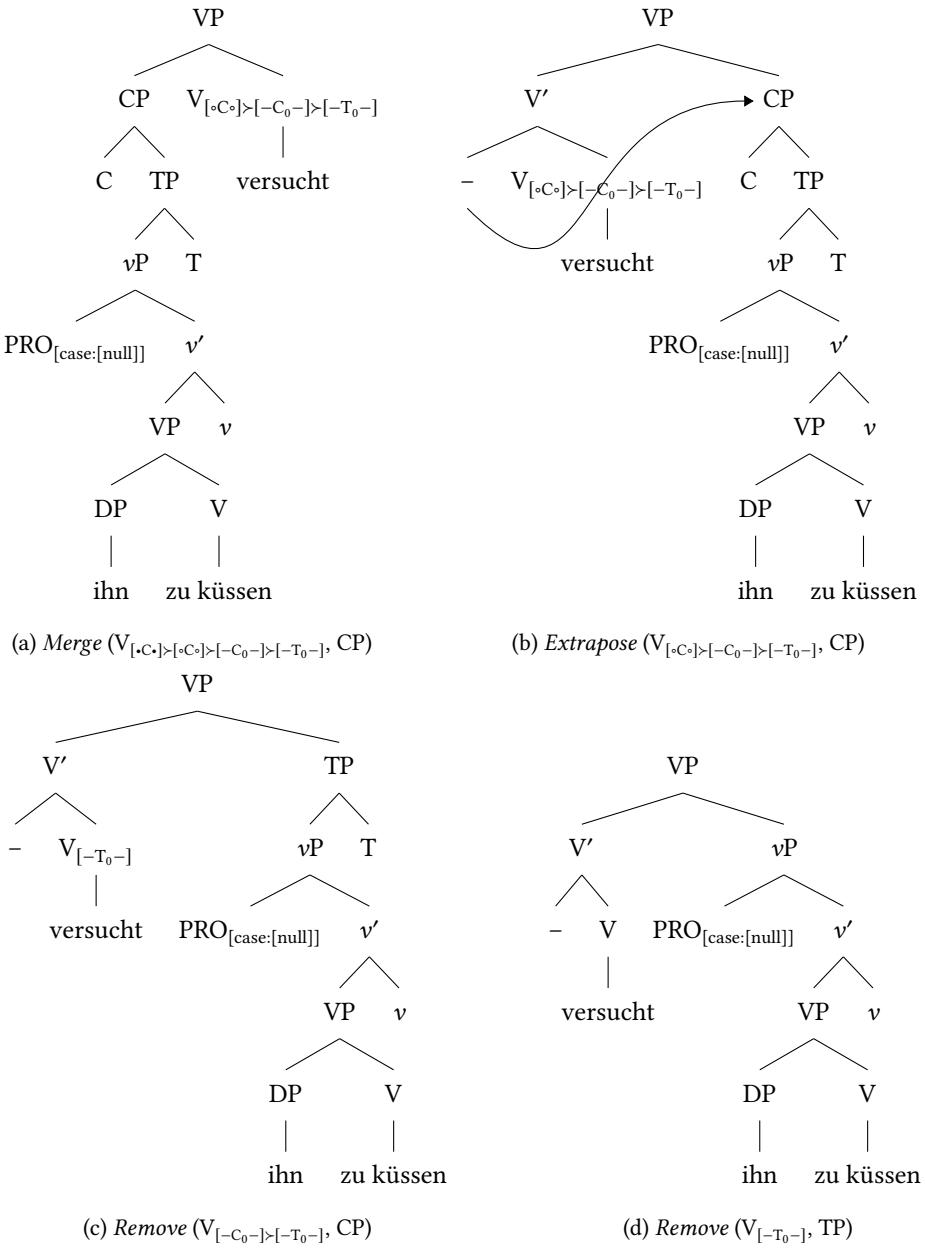
(28) German

a. dass es_1 sich nicht [PRO t_1 zu beanstanden] gehört hat
that it_1 REFL not to object to acceptable is

b. dass $sich_1$ ihm [PRO t_1 zu befreien] gelungen ist
that REFL him_{DAT} to free successful was

The final representation in Figure 9.3d is monoclausal, as required for scrambling and unstressed pronoun fronting to a vP specifier of the matrix V. However, there is a problem: it is not quite clear why a vP in a derived specifier (or adjoined) position does not block extraction via the condition on extraction domains (CED; Huang 1982; Chomsky 1986; Cinque 1990). I will address this issue in the following section. With this proviso, we can conclude that Remove counter-bleeds extraposition: loss of the CP status of the complement in the extraposed position comes too late to block rightward movement (which requires CP status).[24]

[24]The derivation in Figure 9.3 also gives rise to another question: the third construction is possible with periphrastic verb forms; i.e., as an alternative to *versucht* 'tried', as in (21a), there is also the option of *versucht hat* 'tried has', as in (21b). There are (at least) two ways to account for this. First, one might assume that periphrasis comes about by head movement of non-finite lexical V to the auxiliary, followed by discharge of the extraposition feature in the derived position; this would require a minimal modification of the strict cycle condition that incorporates the effect of (this type of) head movement. Second, one might postulate that the two Vs form a single complex head (see, e.g., Zwart 2016 for a recent version of this approach); verb-second movement might then proceed by excorporation.



(a) *Merge* ($V_{[\bullet C \bullet]>[\circ C \circ]>[-C_0-]>[-T_0-]}$, CP)

(b) *Extrapose* ($V_{[\circ C \circ]>[-C_0-]>[-T_0-]}$, CP)

(c) *Remove* ($V_{[-C_0-]>[-T_0-]}$, CP)

(d) *Remove* ($V_{[-T_0-]}$, TP)

Figure 9.3: The third construction

4.3 Deriving evidence for monoclausality

The basic pattern is that operations that presuppose monoclausality are bled and fed by Remove. Let me begin with the simplest cases. First, wide scope of *negation* in restructuring contexts (§2.1.5) follows straightforwardly: scope is an LF-related phenomenon that is determined on the basis of output representations like Figure 9.2c, i.e., after structure removal. Hence, at the stage where the scope of the embedded negation is determined, there is no intermediate clause boundary anymore that might prevent wide scope (or, for that matter, permit embedded scope): Remove feeds scope of negation.[25] Second, similar considerations apply in the case of *intonation* (§2.1.6). The determination of intonational breaks is a phonetic form (PF) process; consequently, it is output representations like Figure 9.2c that are taken into account in order to decide whether intonational breaks can or cannot occur – and after removal, the clause boundary that is indicative of an intonational break is gone: Remove bleeds the generation of smaller intonational phrases.

Next, §2.1.1 (scrambling and unstressed pronoun fronting), §2.1.2 (extraposition), and §2.1.3 (multiple sluicing) all involve evidence for monoclausality based on the a priori unexpected option of extraction (of certain movement types) to take place across a clause boundary with restructuring. An obvious account might therefore rely on the assumption that extraction from the infinitival complement can take place from the in situ position after removal of CP and TP shells, i.e., that Remove directly feeds extraction in the case of movement types that cannot cross a CP boundary. However, there are two problems with this simple view. The first problem concerns successive cyclicity: in general, a phrase that is supposed to undergo extraction from a constituent needs to undergo intermediate movement steps to phase edges, because of the PIC. Accordingly, an item within an infinitival CP that will target a position in the matrix clause (e.g., via scrambling or extraposition) does not know that eventually, there will be no CP (due to removal by the matrix verb); thus, without look-ahead, it will have to undergo movement first to Specv, and then to SpecC.

[25] There is a qualification, though. As observed by Santorini & Kroch (1991), negation is always clause-bound in the third construction; cf. (i) vs. (9a).

(i) German
dass ich seinen neusten Roman beschlossen habe [$_{vP}$ nicht zu lesen]
that I his newest novel$_{Acc}$ decided have not to read
(only narrow scope)

This suggests that, unlike displacement, wide scope is blocked by a vP in a derived (specifier or adjunct) position.

The second problem has already been noted above: recall that a vP in a right-peripheral SpecV position should block scrambling in the third construction, because of the CED (see Figure 9.3d). Taken together, these two problems suggest that the way in which Remove feeds extraction options is somewhat different from the way envisaged under the simple account just sketched.

As a first step to a solution, let us assume that there is some constraint against improper movement that ensures that a CP blocks movement to a clause-external position in the case of scrambling and unstressed pronoun fronting (cf. 1a, 1b, 2c, 2d) and extraposition (cf. 3a, 3b), but not with wh-movement, topicalization or relativization. There are various proposals in the literature as to how the prohibition against movement to low (vP- or TP-internal) positions from a CP can be derived (see, e.g., Müller 2014: Ch. 2; Wurmbrand 2015b; Keine 2016 for three recent attempts); for present purposes, it may suffice to state that such movement (as an instance of Merge) is blocked.

On this basis, consider again the case of scrambling from a restructuring infinitive, as in (2a), repeated here as (29).

(29) German

dass	den Fritz$_1$	keiner	[PRO t_1	zu küssen]	versuchte
that	the Fritz$_{ACC}$	no-one$_{NOM}$		to kiss	tried

Before the infinitival CP is merged with the matrix V, successive-cyclic movement of the embedded object DP *den Fritz* takes place to Spec*v* and SpecC; cf. Figure 9.4.

Next, V combines with CP (see Figure 9.5a); then Remove(V,CP) takes place (see Figure 9.5b). Importantly, DP and TP, as the original specifier and complement of C, are now both reassociated with the matrix V projection in a structure-preserving way, and this means that DP ends up as a specifier of matrix V without having undergone movement to this position. Consequently, there can be no violation of the constraint against improper movement (improper movement can only occur if there is movement in the first place).[26] After this, V removes the TP shell (see Figure 9.5c), which has no further consequences for the moved DP.

As a consequence, DP shows up in the matrix domain without having undergone movement itself, and is now free to move on, yielding, e.g., (29), or, alternatively, to stay in place, with no effects that would be directly discernible since it cannot have crossed matrix VP material (see footnote 19).

[26] See, however, Keine (2016) for evidence that long-distance agreement is subject to the same kinds of restrictions as movement and can also qualify as improper. On this more general view, only operations triggered by features can count as improper; reassociation after structure removal still cannot do so.

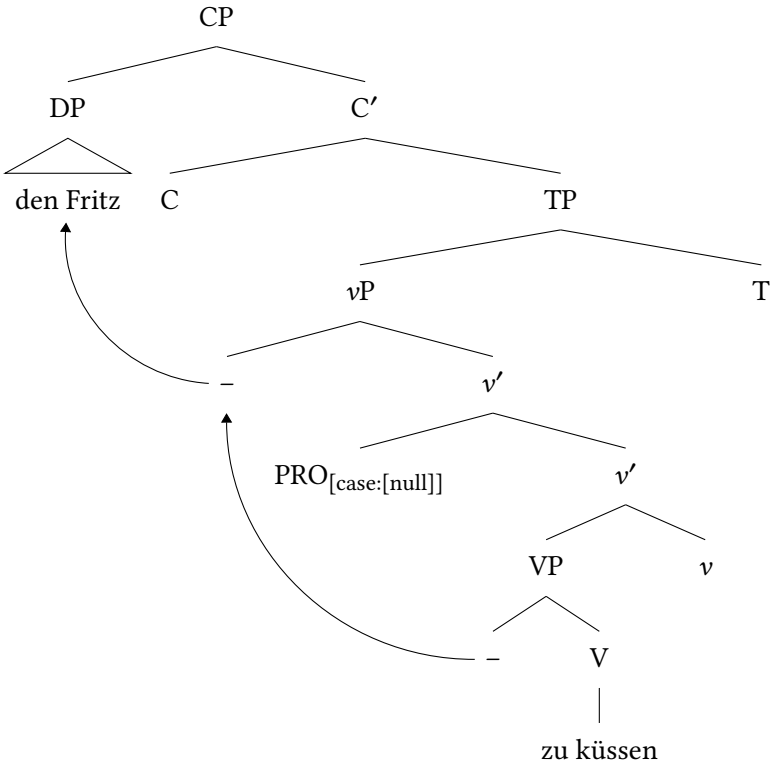
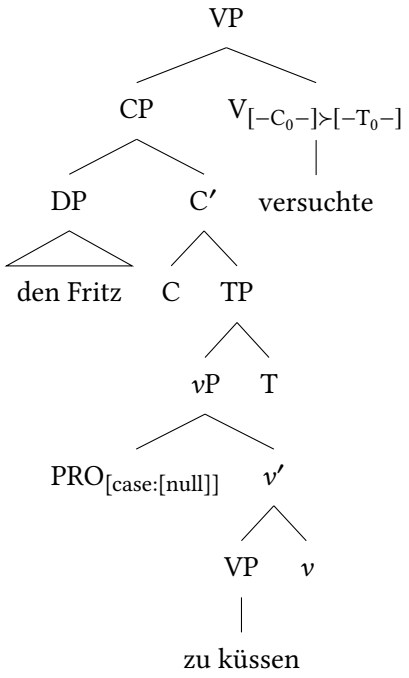
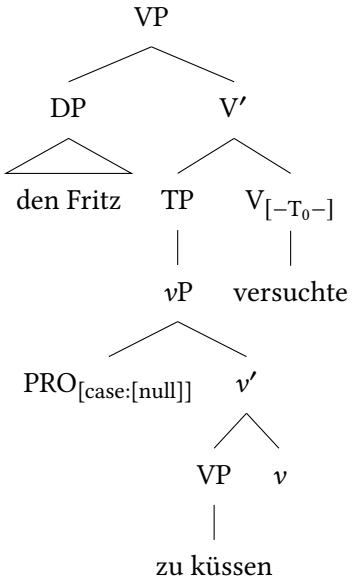


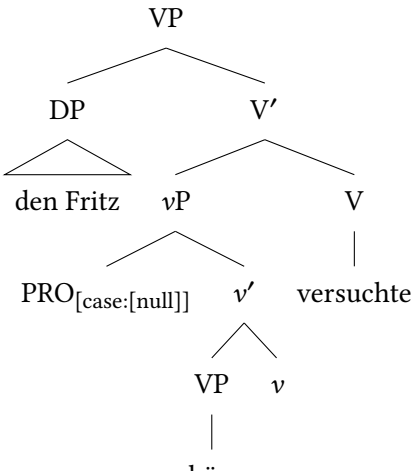
Figure 9.4: Movement in the embedded CP



(a) *Structure before removal*



(b) *Remove* ($V_{[-C_0-]\succ[-T_0-]}$, CP), reassociation of DP



(c) *Remove* ($V_{[-T_0-]}$, TP)

Figure 9.5: Extraction and Restructuring

This explains why scrambling and unstressed pronoun fronting can take place from restructuring infinitives.[27]

The reasoning is basically identical with extraposition: the improper movement effect in the presence of a CP (see 3) can be circumvented after CP removal in restructuring contexts (see 4).

As for recoverability-driven fronting of wh-phrases in multiple sluicing contexts (cf. 7a vs. 6, 7b), recall that there are three competing approaches: the second wh-phrase may have undergone scrambling (Sauerland 1999), extraposition (Lasnik 2014), or wh-movement (Heck & Müller 2003). Assuming that the relevant distinctions in the latter type of approach are due to an *initial* presence or absence of a CP projection, such that the second wh-movement in the embedded domain is blocked in the presence of a CP (as argued in Heck & Müller 2003), we now have a theory-internal argument for the former two approaches (which are both compatible with an initial presence of CP that is subsequently undone by removal).

The final movement-related issue to be addressed concerns scrambling in the third construction; cf. the examples in (21) and the derivation in Figure 9.3. Recall that the problem with the derivation resulting in Figure 9.3d is that scrambling from the vP in the extraposed position should violate the CED. This problem is now solved: almost exactly the same derivation as in Figure 9.5 takes place with

[27]It should be noted that the present analysis does not per se exclude cases like (i-b), where successive-cyclic long-distance movement takes place from a position in CP_3 to the specifier of CP_2 (cf. (i-a)), followed by structure removal induced by the restructuring predicate *versuchen* 'try', subsequent reassociation of DP_0 (plus further scrambling) in the matrix domain, and finally extraposition of CP_3.

(i) a. [$_{VP}$ [$_{CP_2}$ dieses Buch$_0$ [$_{C'}$ C [$_{TP}$ [$_{vP}$ t$''_0$ PRO [$_{VP}$ [$_{CP_3}$ t$'_0$ [$_C$ dass] man t$_0$ lesen
this book$_{ACC}$ that one$_{NOM}$ read
soll] [$_V$ vorzuschlagen]] *v*] T]]] [$_V$ versucht hat]]
should to suggest tried has

b. ?* dass dieses Buch$_0$ keiner [$_{VP}$ [$_{vP}$ t$''_0$ PRO [$_{VP}$ t$_3$ [$_V$ vorzuschlagen]] *v*] [$_V$
that this book$_{ACC}$ no-one$_{NOM}$ to suggest
versucht hat]] [$_{CP_3}$ t$'_0$ [$_C$ dass] man t$_0$ lesen soll]
tried has that one$_{NOM}$ read should

In contrast, if the fronted object *dieses Buch* undergoes topicalization in the same context, there is a marked improvement (but no full acceptability). For the time being, I will leave open the question of whether the ill-formedness of (i-b) can (or should) be made to follow from a general constraint against improper movement, or should be taken to indicate a cumulative effect resulting from the choice of several marked options in the syntax of German (among them extraction from *dass* clauses and complexity of matrix predicate (*vorzuschlagen versucht hat*)).

extraction in the third construction, the only difference being that CP is extraposed prior to removal. Thus, a DP that is in SpecC of the extraposed CP becomes reassociated with VP as a consequence of CP removal in the extraposed position. As before, this means that a DP that has reached SpecC of a restructuring infinitive ends up in the matrix VP domain without having undergone movement to that position; and as before, two possibilities arise: First, DP can undergo further movement in the matrix clause (including scrambling and unstressed pronoun movement). Second, DP may stay in SpecV; since it has not moved there, the position is virtually indistinguishable from a base-merged position at this point. I would like to contend that this second option does indeed have discernible empirical effects: It provides a principled approach to *pseudo-scrambling* phenomena as they have been identified by Geilfuß (1991).

The relevant observation is that items in immediately preverbal positions in the third construction do not exhibit the characteristic properties of *scrambling* in German; they instantiate what has been called *pseudo-scrambling*. Geilfuß (1991) presents evidence from a variety of different phenomena, among them focus projection, wh-scrambling, scope, non-specific indefinites, directional PPs, extraction, idioms, and quantifier floating. Let me just briefly address two of them. First, (30a) shows that maximal focus projection in out-of-the-blue contexts is normally impossible with scrambled items; in contrast, (30b) shows that a pseudo-scrambled DP in the third construction permits focus projection (the effect goes away again if DP_1 were to undergo further displacement to a position in front of the matrix object). In the present approach, this is accounted for straightforwardly: focus projection is incompatible with scrambling, and the pseudo-scrambled DP in (30b) is not moved but transported to matrix SpecV via reassociation after CP removal.

(30) German

a. # Fritz hat das MÄRchen$_1$ einem Kind t$_1$ vorgelesen
Fritz$_{NOM}$ has the fairy tale$_{ACC}$ a child$_{DAT}$ read to

b. Fritz hat einem Kind das MÄRchen$_1$ [$_{VP}$ versucht [t$_1$
Fritz$_{NOM}$ has a child$_{DAT}$ the fairy tale$_{ACC}$ tried
vorzulesen]]
to read to

Second, relative scope illustrates the same effect. Normally, scrambling of one quantified DP across another one leads to scope ambiguities (see 31a). However, extremely local pseudo-scrambling from third construction environments does

not (see 31b). Given the present analysis, DP_1 in (31b) does not exhibit this property indicative of movement for the simple reason that it has reached its position not by movement, but by reassociation after CP removal.

(31) German

a. Er hat mindestens ein Geschenk$_1$ fast jedem Gast t_1
he$_{NOM}$ has at least one present$_{ACC}$ almost every guest$_{DAT}$
überreicht
given
Readings: $\exists > \forall$, $\forall > \exists$

b. Er hat mindestens ein Geschenk$_1$ versucht [fast jedem Gast
he$_{NOM}$ has at least one present$_{ACC}$ tried almost every guest
t_1 zu überreichen]
to give
Readings: $\exists > \forall$, $^{*}\forall > \exists$

To sum up, assuming that the *compactness* property (§2.1.4), to the extent that it holds, can be accounted for in one of the ways suggested in the literature, the empirical evidence for monoclausality highlighted in §2.1 has been derived in toto.

More generally, I would like to conclude that a Remove-based approach to restructuring infinitives embedded under control verbs in German is conceptually viable and empirically motivated; in fact, an analysis in terms of structure removal would seem to be the only kind of principled approach that captures both the evidence for biclausality and the evidence for monoclausality in a straightforward way. Furthermore, the option of deriving local displacement in restructuring contexts as a consequence of reassociation after removal (rather than by movement) offers a new look on pseudo-scrambling in the third construction (and possibly in other contexts as well). All in all, then, it seems to me that there is every reason to return to classical concepts of restructuring as involving a genuine syntactic reduction of clause size; the core problem with these approaches – viz., that the analyses were not sufficiently principled and restricted – can be solved when an elementary operation Remove is identified as the complete mirror image of Merge.[28]

[28]Needless to say, there are many more aspects of restructuring that will ultimately have to be addressed, both in German and, particularly, when it comes to extending the analysis to other languages. Let me just mention two issues that I cannot address here for lack of space. First, *long-distance passivization* has played an important role in the development of restructuring

Abbreviations

ACC	accusative	NOM	nominative
CED	condition on extraction domains	PF	phonetic form
DAT	dative	PIC	phase impenetrability condition
ECM	exceptional case marking	PTCL	particle
EPP	extended projection principle	REFL	reflexive
LF	logical form		

Acknowledgments

This paper is dedicated to Ian Roberts. For comments and discussion, I am grateful to Johanna Benz, Benjamin Bruening, Christina Dschaak, Johannes Englisch, Gisbert Fanselow, Silke Fischer, Kleanthes Grohmann, Fabian Heck, Daniel Hole, Dalina Kallulli, Sampson Korsah, Lanko Marušič, Andrew Murphy, Andrew Nevins, David Pesetsky, Marie-Luise Schwarzer, Volker Struckmeier, Lída Veselovská, Philipp Weisser, Susi Wurmbrand, two anonymous reviewers, and audiences at Universität Leipzig (workshop on shrinking trees), Leucorea Wittenberg (workshop on genus verbi), Universität Stuttgart, Universität Wien (workshop on passives), Tel Aviv University (IATL 33), and SinFonIJA X in Dubrovnik. Research for this article was supported by a DFG Reinhart Koselleck grant (MU 1444/14-1, *Structure removal in syntax*).

References

Abels, Klaus. 2012. *Phases: An essay on cyclicity in syntax*. Berlin: De Gruyter.

Adger, David. 2003. *Core syntax*. Oxford: Oxford University Press.

Aissen, Judith & David Perlmutter. 1983. Clause reduction in Spanish. In David Perlmutter (ed.), *Studies in Relational Grammar 1*, 360–403. Chicago: University of Chicago Press.

Bader, Markus & Tanja Schmid. 2009. Verb clusters in colloquial German. *Journal of Comparative Germanic Syntax* 12(3). 175–228. DOI: 10/b22vds.

theories (see Höhle 1978, Wurmbrand 2001; 2015a,b, Sternefeld 2006, Haider 2010, and Keine & Bhatt 2016, among many others). In Müller (2019), I sketch an analysis in terms of Remove that extends the present analysis.

Second, I have been silent about *status government* (see Bech 1955–1957; Fabb 1984), which is also sometimes viewed as being indicative of restructuring. See Benz (2019) on how the concept of status government interacts with a Remove-based approach to restructuring.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
Barnickel, Katja. 2014. *Opazität in AcI-Konstruktionen*. Universität Leipzig. (MA thesis).
Barss, Andrew. 1986. *Chains and anaphoric dependence*. MIT. (Doctoral dissertation).
Bayer, Josef & Jaklin Kornfilt. 1994. Against scrambling as an instance of Move Alpha. In Norbert Corver & Henk C. van Riemsdijk (eds.), *Studies on scrambling*, 17–60. Berlin: De Gruyter.
Bech, Gunnar. 1955–1957. *Studien über das deutsche Verbum Infinitum*. Reprint 1983. Tübingen: Niemeyer.
Bennis, Hans. 1983. A case of restructuring. In Hans Bennis & W. van Lessen Kloeke (eds.), *Linguistics in the Netherlands*, 9–19. Dordrecht: Foris.
Benz, Johanna. 2019. Structure removal counterfeeds status government. In Andrew Murphy (ed.), *Structure removal*, vol. 94 (Linguistische Arbeits Berichte). Leipzig: Universität Leipzig.
Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2014. A syntactic universal and its consequences. *Linguistic Inquiry* 45(2). 169–225. DOI: 10.1162/LING_a_00153.
Büring, Daniel. 2005. *Binding theory*. Cambridge: Cambridge University Press.
Büring, Daniel & Katharina Hartmann. 1996. The Kayne mutiny. In Uli Lutz & Jürgen Pafel (eds.), *On extraction and extraposition in German*, 59–79. Amsterdam: John Benjamins. DOI: 10.1075/la.17.04bur.
Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Academic Press.
Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.
Chomsky, Noam. 2015a. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies*

and beyond: Studies in honour of Adriana Belletti, 1–16. Amsterdam: John Benjamins. DOI: 10.1075/la.223.01cho.

Chomsky, Noam. 2015b. Simple invisibles. In Ludmila Veselovská & Markéta Vanebová (eds.), *Complex visibles out there*, vol. 4 (Olomouc Modern Language Series), 17–24. Olomouc: Palacky University.

Chomsky, Noam & Howard Lasnik. 1993. Principles and parameters theory. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *Syntax: Ein internationales Handbuch zeitgenössischer Forschung*, vol. 1, 506–569. Berlin: De Gruyter.

Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, MA: MIT Press.

den Besten, Hans & Jean Rutten. 1989. On verb raising, extraposition and free word order in Dutch. In Dany Jaspers, Wim Klooster, Yvan Putseys & Pieter Seuren (eds.), *Sentential complementation and the lexicon: Studies in honor of Wim de Geest*, 41–56. Dordrecht: Foris.

Di Sciullo, Anna Maria & Edwin Williams. 1987. *On the definition of word*. Cambridge, MA: MIT Press.

Dschaak, Christina. 2017. *A structure removal approach to restructuring in Russian*. Universität Leipzig. (BA thesis).

Embick, David. 2010. *Localism vs. globalism in morphology and phonology*. Cambridge, MA: MIT Press.

Evers, Arnold. 1975. *The transformational cycle in Dutch and German*. University of Utrecht. (Doctoral dissertation).

Fabb, Nigel. 1984. *Syntactic affixation*. Cambridge, MA: MIT. (Doctoral dissertation).

Fanselow, Gisbert. 1987. *Konfigurationalität*. Tübingen: Narr.

Fanselow, Gisbert. 1989. Coherent infinitives in German: Restructuring vs. IP-complementation. In Christa Bhatt, Elisabeth Löbel & Claudia Schmidt (eds.), *Syntactic phrase structure phenomena in noun phrases and sentences*, 1–16. Amsterdam: John Benjamins. DOI: 10.1075/la.6.02fan.

Fanselow, Gisbert. 1991. Minimale Syntax. Universität Passau habilitation thesis.

Fanselow, Gisbert. 2004. The MLC and derivational economy. In Arthur Stepanov, Gisbert Fanselow & Ralf Vogel (eds.), *Minimality effects in syntax*, 73–123. Berlin: De Gruyter.

Featherston, Sam & Wolfgang Sternefeld. 2003. The interaction of factors in judgements of reflexive structures: Data from object coreference in German. In Lutz Gunkel, Gereon Müller & Gisela Zifonun (eds.), *Arbeiten zur Reflexivierung*, 25–50. Tübingen: Niemeyer.

Fischer, Silke. 2004. *Towards an optimal theory of reflexivization.* Universität Tübingen. (Doctoral dissertation).
Frey, Werner. 1993. *Syntaktische Bedingungen für die Interpretation.* Berlin: Akademieverlag.
Geilfuß, Jochen. 1988. *Reanalyse?* Universität Tübingen. (MA thesis).
Geilfuß, Jochen. 1991. Scrambling und pseudoscrambling. Ms., Universität Tübingen.
Georgi, Doreen. 2014. *Opaque interactions of Merge and Agree.* Universität Leipzig. (Doctoral dissertation).
Grewendorf, Günther. 1983. Reflexivierung in deutschen AcI-Konstruktionen – Kein transformationsgrammatisches Dilemma mehr. *Groninger Arbeiten zur Germanistischen Linguistik* 23. 120–196.
Grewendorf, Günther. 1988. *Aspekte der deutschen Syntax.* Tübingen: Narr.
Gunkel, Lutz. 2003. *Infinitheit, Passiv und Kausativkonstruktionen im Deutschen.* Tübingen: Stauffenburg Verlag.
Haegeman, Liliane & Henk C. van Riemsdijk. 1986. Verb projection raising, scope, and the typology of rules affecting verbs. *Linguistic Inquiry* 17(3). 417–466.
Haider, Hubert. 1993. *Deutsche Syntax – generativ.* Tübingen: Narr.
Haider, Hubert. 2010. *The syntax of German.* Cambridge: Cambridge University Press.
Haider, Hubert. 2016. Compactness as a PF effect? Ms., Universität Salzburg.
Heck, Fabian. 2016. Non-monotonic derivations. Universität Leipzig habilitation thesis.
Heck, Fabian & Gereon Müller. 2003. Derivational optimization of wh-movement. *Linguistic Analysis* 33. (Volume appeared 2007), 97–148.
Heck, Fabian & Gereon Müller. 2007. Extremely local optimization. In Erin Brainbridge & Brian Agbayani (eds.), *Proceedings of the thirty-fourth Western Conference on Linguistics*, 170–183. Fresno: California State University.
Heycock, Caroline & Anthony Kroch. 1994. Verb movement and coordination in a dynamic theory of licensing. *The Linguistic Review* 11(3–4). 257–283. DOI: 10.1515/tlir.1994.11.3-4.257.
Hicks, Glyn. 2009. *The derivation of anaphoric relations.* Amsterdam: John Benjamins. DOI: 10.1075/la.139.
Hinterhölzl, Roland. 1999. *Restructuring infinitives and the theory of complementation.* UCLA. (Doctoral dissertation).
Höhle, Tilman. 1978. *Lexikalistische Syntax: Die Aktiv-Passiv-Relation und andere Infinitkonstruktionen im Deutschen.* Tübingen: Niemeyer.
Hornstein, Norbert. 2001. *Move! A minimalist theory of construal.* Malden, MA: Blackwell.

Hornstein, Norbert. 2014. Final comments on lecture 4. Faculty of Language Blog, Monday, August 4, 2014.
Huang, C.-T. James. 1982. *Logical relations in Chinese and the theory of grammar*. Cambridge, MA: MIT. (Doctoral dissertation).
Huybregts, Riny. 1982. Class notes. Ms., Tilburg University.
Keine, Stefan. 2016. *Probes and their horizons*. University of Massachusetts, Amherst. (Doctoral dissertation).
Keine, Stefan & Rajesh Bhatt. 2016. Interpreting verb clusters. *Natural Language and Linguistic Theory* Published online. 13.4.2016.
Kiss, Tibor. 1995. *Infinite Komplementation*. Tübingen: Niemeyer.
Koopman, Hilda & Anna Szabolcsi. 2000. *Verbal complexes*. Cambridge, MA: MIT Press.
Korsah, Sampson & Andrew Murphy. 2017. Against CP selection: Evidence from DP shells in Kwa. Ms., Universität Leipzig.
Koster, Jan. 1987. *Domains and dynasties*. Dordrecht: Foris.
Landau, Idan. 2000. *Elements of control*. Dordrecht: Kluwer.
Lasnik, Howard. 2014. Multiple sluicing in English? *Syntax* 17(1). 1–20. DOI: 10.1111/synt.12009.
Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
Müller, Gereon. 1995. On extraposition and successive cyclicity. In Uli Lutz & Jürgen Pafel (eds.), *On extraction & extraposition in German*, 213–243. Amsterdam: John Benjamins. DOI: 10.1075/la.11.08mul.
Müller, Gereon. 2001. Order preservation, parallel movement, and the emergence of the unmarked. In Géraldine Legendre, Jane Grimshaw & Sten Vikner (eds.), *Optimality-theoretic syntax*, 279–313. Cambridge, MA: MIT Press.
Müller, Gereon. 2014. *Syntactic buffers*. Vol. 91. Universität Leipzig: Linguistische Arbeits Berichte.
Müller, Gereon. 2015. Structure removal: A new approach to conflicting representations. Lecture Notes, Universität Leipzig.
Müller, Gereon. 2016. The short life cycle of external arguments in German passive derivations. Ms., Universität Leipzig.
Müller, Gereon. 2017. Structure removal: An argument for feature-driven Merge. *Glossa* 2(1), 28. http://doi.org/10.5334/gjgl.193. DOI: 10.5334/gjgl.193.
Müller, Gereon. 2018. Structure removal in complex prefields. *Natural Language & Linguistic Theory* 36(1). 1–46. DOI: 10.1007/s11049-017-9374-z.
Müller, Gereon. 2019. Long-distance passives by structure removal. Ms., Universität Leipzig.

Müller, Gereon & Wolfgang Sternefeld. 1995. Extraction, lexical variation, and the theory of barriers. In Urs Egli, Peter E. Pause, Christoph Schwarze, Arnim von Stechow & Götz Wienold (eds.), *Lexical knowledge in the organization of language*, 35–80. Amsterdam: John Benjamins. DOI: 10.1075/cilt.114.04mul.

Murphy, Andrew. 2015. Ellipsis as syntactic structure removal. Ms., Universität Leipzig.

Murphy, Andrew & Gereon Müller. 2016. The short life cycle of elided elements: Ellipsis as syntactic structure removal. Ms., Universität Leipzig. (Poster, Workshop *Ellipsis Across Borders*, Sarajevo 2016).

Pesetsky, David. 1995. *Zero syntax*. Cambridge, MA: MIT Press.

Pesetsky, David. 2016. Exfoliation: Towards a derivational theory of clause size. Ms., MIT, Cambridge, MA.

Puškar, Zorica. 2016. Nominal expressions in languages without articles: Towards resolving the NP/DP debate by structure removal. Ms., Universität Leipzig.

Reis, Marga. 1976. Reflexivierung in deutschen A.c.I.-Konstruktionen: Ein transformationsgrammatisches Dilemma. *Papiere zur Linguistik* 9. 5–82.

Reuland, Eric. 2001. Primitives of binding. *Linguistic Inquiry* 32(3). 439–492. DOI: 10.1162/002438901750372522.

Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.

Richards, Marc. 2007. On feature inheritance. *Linguistic Inquiry* 38(3). 563–572. DOI: 10.1162/ling.2007.38.3.563.

Richards, Norvin. 2001. *Movement in language*. Oxford: Oxford University Press.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Roberts, Ian. 1997. Restructuring, head movement, and locality. *Linguistic Inquiry* 28(3). 423–460.

Ross, John Robert. 1967. *Constraints on variables in syntax*. Cambridge, MA: MIT. (Doctoral dissertation).

Růžička, Rudolf. 1983. Remarks on control. *Linguistic Inquiry* 14. 309–324.

Sabel, Joachim. 1996. *Restrukturierung und Lokalität*. Berlin: Akademie-Verlag.

Sadock, Jerry. 1991. *Autolexical syntax*. Chicago: Chicago University Press.

Safir, Ken. 2010. Viable syntax: Rethinking minimalist architecture. *Biolinguistics* 4. 35–107.

Safir, Ken. 2015. It's a conspiracy! The A/A′-distinction as an epiphenomenon. Ms., Rutgers University.

Salzmann, Martin. 2011. Resolving the movement paradox in verb projection raising. In *Empirical issues in syntax and semantics 8*, 453–485. Paris: CNRS.

Salzmann, Martin. 2013a. New arguments for verb cluster formation at PF and a right-branching VP: Evidence from verb doubling and cluster penetrability. *Linguistic Variation* 13. 81–132. DOI: 10.1075/lv.13.1.03sal.

Salzmann, Martin. 2013b. Rule ordering in verb cluster formation. In *Rule interaction in grammar*, vol. 90 (Linguistische Arbeits Berichte), 65–121. Universität Leipzig: Institut für Linguistik.

Santorini, Beatrice & Anthony Kroch. 1991. Remnant extraposition in German. Ms., University of Pennsylvania.

Sauerland, Uli. 1999. Locality in ellipsis: Sluicing vs. multiple sluicing. Ms., Universität Tübingen.

Schmid, Tanja. 2005. *Infinitival syntax: Infinitivus pro participio as a repair strategy*. Amsterdam: John Benjamins. DOI: 10.1075/la.79.

Schwarzer, Marie-Luise. 2016. *Tough: Easy*. Universität Leipzig. (MA thesis).

Stabler, Edward. 2013. Two models of minimalist, incremental syntactic analysis. *Topics in Cognitive Science* 5(3). 611–633. DOI: 10.1111/tops.12031.

Stepanov, Arthur. 2012. Voiding island effects via head movement. *Linguistic Inquiry* 43(4). 680–693. DOI: 10.1162/ling_a_00111.

Sternefeld, Wolfgang. 1990. Scrambling and minimality. In Günther Grewendorf & Wolfgang Sternefeld (eds.), *Scrambling and barriers*, 239–257. Amsterdam: John Benjamins.

Sternefeld, Wolfgang. 2006. *Syntax*. Two volumes. Tübingen: Stauffenburg.

Sternefeld, Wolfgang & Sam Featherston. 2003. The German reciprocal *einander* in double object constructions. In Lutz Gunkel, Gereon Müller & Gisela Zifonun (eds.), *Arbeiten zur Reflexivierung*, 239–265. Tübingen: Niemeyer.

Stiebels, Barbara. 2007. Towards a typology of complement control. In Barbara Stiebels (ed.), *Studies in complement control*, vol. 47 (ZAS Papers in Linguistics), 1–80. Berlin: ZAS.

Stiebels, Barbara. 2010. Inhärente Kontrollprädikate im Deutschen. *Linguistische Berichte* 224. 391–440.

von Stechow, Arnim & Wolfgang Sternefeld. 1988. *Bausteine syntaktischen Wissens*. Opladen: Westdeutscher Verlag.

Wöllstein-Leisten, Angelika. 2001. *Die Syntax der dritten Konstruktion*. Tübingen: Stauffenburg Verlag.

Wurmbrand, Susi. 2001. *Infinitives: Restructuring and clause structure*. Berlin: De Gruyter.

Wurmbrand, Susi. 2002. Syntactic vs. semantic control: Proceedings of the 15th Workshop on Comparative Germanic Syntax. In Jan-Wouter Zwart & Werner Abraham (eds.), *Studies in comparative Germanic syntax*, 93–127. Amsterdam: John Benjamins. DOI: 10.1075/la.53.07wur.

Wurmbrand, Susi. 2004. Two types of restructuring—Lexical vs. functional. *Lingua* 114(8). 991–1014. DOI: 10.1016/S0024-3841(03)00102-5.

Wurmbrand, Susi. 2007. How complex are complex predicates? *Syntax* 10(3). 243–288. DOI: 10.1111/j.1467-9612.2007.00110.x.

Wurmbrand, Susi. 2015a. Complex predicate formation via Voice incorporation. In Léa Nash & Pollet Samvelian (eds.), *Approaches to complex predicates*, 248–290. Leiden: Brill.

Wurmbrand, Susi. 2015b. Restructuring cross-linguistically. In Thuy Bui & Deniz Özyıldız (eds.), *NELS 45: Proceedings of the Forty-Fifth Annual Meeting of the North East Linguistics Society*, 227–240. Amherst, MA: GLSA Publications.

Zwart, Jan-Wouter. 2016. Periphrastic morphology, verb clusters, and verb movement. Handout of a talk at CRISSP/KU Leuven, Brussels.

Chapter 10

Rethinking phrase structure

Howard Lasnik
University of Maryland at College Park

Zach Stone
University of Maryland at College Park

We investigate structural properties of two set-theoretic models of phrase structure, namely the phrase markers of LSLT and bare phrase structure. We demonstrate that neither set-theoretic model has a nice notion of "substructure" which is well-behaved with respect to the extension condition. We compare these with graph- and order-theoretic representations which have well-behaved structure-preserving maps for characterizing both the extension condition and the operation Agree.

1 Introduction

We review two models of phrase structure in Generative Grammar and survey their structural properties with respect to substructures and isomorphism. We especially look at how these structural notions bear on the extension condition. Specifically, we show that neither formal representation captures a sufficiently general form of the extension condition, while the correct properties are captured straightforwardly both by graph- and order-theoretic representations.

We use standard set-theoretic notation: we sometimes indicate a set by writing its elements in braces $A = \{a_i\}_{i \in I}$; we use the symbol $A \subset B$ to represent that every element of A is an element of B, called a (potentially improper) subset; we use $A \cong B$ to indicate that there is some bijection between the sets; we use $A \cup B$ to represent the union of two sets; we use A^* to represent the set of all *words*, or

Howard Lasnik & Zach Stone. 2020. Rethinking phrase structure. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 191–206. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280645

strings of finite length spelled from symbols of A; we represent a set-function $f : A \rightarrow B$, or sometimes just $A \rightarrow B$.

We discuss substructures and isomorphism somewhat informally, though all forms of them discussed can be made precise in the language of model theory or category theory.

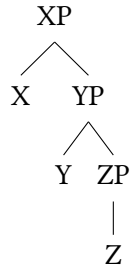
2 Phrase markers and reduced phrase markers

Lasnik (2006) briefly points out an issue that arises with respect to the extension condition (EC), the Minimalist version of the principle of the cycle proposed by Chomsky (1993), or, more precisely, the deduction of it by Chomsky (2000). Chomsky (1993: 22) formulated EC as follows:

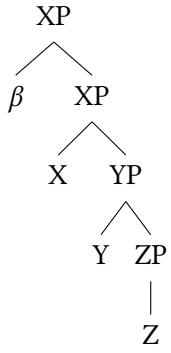
(1) GT [generalized transformation] and Move α extend K to K′ which includes K as a proper part.

The Chomsky (2000) rationale for EC is that derivations conform to a condition demanding that there be no tampering by a transformation with already existing structure. If an item is newly attached at the "top" of a tree, the former tree is assumed to be completely preserved as a sub-tree by external merge, and also by internal merge on the copy theory of movement. Here's a simplified toy illustration. Start with the tree in (2).

(2)



(3)



Now suppose β is adjoined to XP in accord with (1). The resulting tree is (3), which clearly includes (2) as a sub-tree, the intended consequence.

But now consider these structures in terms of their set-theoretic representations, for example, as in LSLT (Chomsky 1975 [1955]). The picture in (2) stands for the actual object in (4), a set of strings:

(4) {XP, X YP, X Y ZP, X Y Z}

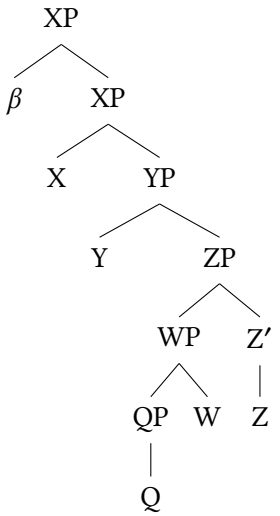
And the picture in (3) stands for the actual derived object in (5):

(5) {XP, *β* XP, *β* X YP, *β* X Y ZP, *β* X Y Z}

Notice that (4) is in no respect a sub-object, i.e., a subset, of (5). And this is not because of any special property of the example chosen. It is invariably true that if we adjoin something to the "top" of an LSLT-style phrase marker (PM), the resulting set is never a superset of the original. That is, we have dramatically "tampered" with the original set: It is gone.

It is important to realize that the same conclusion follows on any "purely" set theoretic implementation of syntactic theory. One other such implementation is that of Lasnik & Kupin (1977). In that framework as in that of LSLT, a PM is a set of strings. The difference is that for L&K the PM consists entirely of the terminal string and "monostrings" (strings comprised of exactly one non-terminal symbol surrounded by any number of terminal symbols). L&K called their PMs reduced phrase markers (RPMs). To see that the same conclusion outlined above happens with RPMs, we need to slightly complicate the example discussed, since there, it turns out that the PM and RPM are the same. So consider the slightly more complex tree in (6):

(6)



The initial RPM is (7):

(7) {XP, X YP, X Y ZP, X Y WP Z, X Y QP W Z, X Y Q W Z′, X Y Q Z}

And the derived RPM is (8):

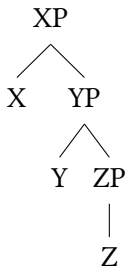
(8) {XP, β XP, β X YP, β X Y ZP, β X Y WP Z, β X Y QP W Z, β X Y Q W Z′, β X Y Q W Z}

Once again, the initial set is not a subset of the derived set. In fact, as with the LSLT PMs, there is no obvious simple set-theoretic relation at all between them.

This is a special case of a more pervasive limitation of such purely set-theoretic formalizations: constituents are never sub-structures (subsets in this instance), nor are many core syntactic configurations, such as the template for a specifier.

Surprisingly, attaching at the very "bottom" does yield a superset of the initial set, the exact opposite of the evidently desired result. We illustrate this beginning with the simple structure in (2), repeated here, followed by the RPM (which, as noted earlier, is identical to the LSLT PM in this case):

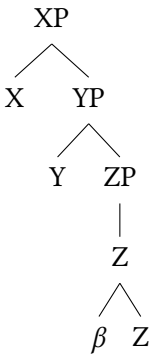
(9)



(10) {XP, X YP, X Y ZP, X Y Z}

This time, adjoin β at the bottom, in extreme violation of EC:

(11)



The new set is (12):

(12) {XP, X YP, X Y ZP, X Y Z, X Y β Z}

But surprisingly this time the original object is not tampered with as (10) ⊂ (12). It is safe to conclude, then, that if Chomsky's deduction of EC is to be maintained, neither classic set-theoretic formalization of phrase structure is appropriate.

In summary, while producing the "wrong" result, RPMs have a well-defined notion of substructure. For example, (10) ⊂ (12) is a subset relation, and the defining relations of an RPM – precedence and dominance – are "preserved" by this inclusion (for example, the monostring X YP dominates the monostring X Y ZP in (10), as does the corresponding monostring in (12)).

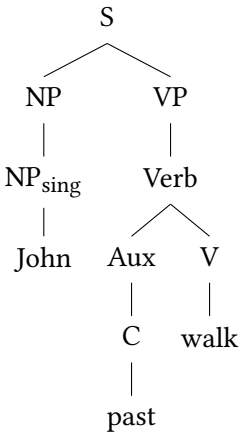
There is also a clear notion of *isomorphism* between RPMs, which will be important in §3.2. Roughly, if N and M are two sets of nonterminals and T and S sets of terminals, a pair of bijections $f : N \to M$ and $g : T \to S$ extends to a bijection between sets of strings $(f + g)^* : (N \cup T)^* \to (M \cup S)^*$ (replacing each nonterminal symbol A in a string from $(N \cup T)^*$ with $f(A)$ and each terminal symbol t with $g(t)$) and hence between monostrings. Given such bijections, we can compare RPMs F and G consisting of monostrings from $(N \cup T)^*$ and $(M \cup S)^*$, respectively, by using the bijection $(f + g)^*$ restricted to $F \subset (N \cup T)^*$ and $G \subset (M \cup S)^*$ (if possible). We could say that two RPMs F and G over (N, T) and (M, S) respectively are isomorphic if we can rename monostrings from F as monostrings in G along the bijection, and vice-versa (using the inverse of $(f+g)^*$ restricted to $G \to F$), extending to a bijection $F \cong G$, such that two monostrings ϕ and ψ in F are in a precedence or dominance relation exactly when the corresponding monostrings in G are.

Before proceeding, we note in passing that it is not only the case that in the LSLT model, attachment at the top does not "preserve structure", but also that attachment at the top is literally impossible, at least for a transformation. Transformations in that framework consist of a structural analysis (SA) and a SC (structural change). The former determines whether the T is applicable to a particular PM, while the latter indicates the operation to be performed. An SA is a sequence of "terms", each term a (string) variable, a constant (i.e., a syntactic symbol), or a linear combination of any of the preceding. Consider Chomsky's auxiliary transformation "affix hopping" as presented by Chomsky (1957). The following is one of a family of 20 SAs embodied by the T:

(13) X – *past* – V – Y

Applicability is determined by comparing the SA with the members of the set to establish satisfaction. Any string satisfies a variable, while a constant is satisfied only by that very symbol. The T with SA in (13) is applicable to the PM pictorially represented in (14).

(14)



The PM is in (15).

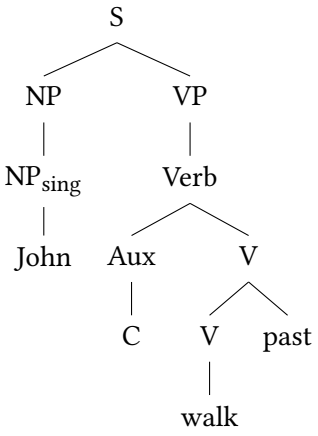
(15) {S, NP VP, NP Verb, NP Aux V, NP Aux walk, NP C V, NP C walk, NP past V, NP past walk, NP_{sing} VP, NP_{sing} Verb, NP_{sing} Aux V, NP_{sing} Aux walk, NP_{sing} C V, NP_{sing} C walk, NP_{sing} past V, NP_{sing} past walk, John VP, John Verb, John Aux V, John Aux walk, John C V, John C walk, John past V, John past walk}

In this case, applicability of the transformation is established by any of 3 members of the set:

(16) NP past V NP_{sing} past V John past V

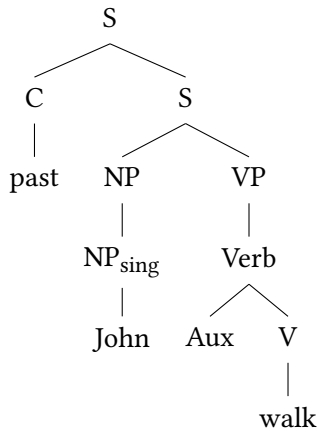
Notice that every member of any PM has symbols in a linear order; every pair of symbols in a member are in the precedence relation. Thus, the symbols in any SA are likewise necessarily in a linear order. Thus, a symbol can adjoin to one that follows it (as in affix hopping, where past will adjoin to V), or to one that precedes it. The result of the operation is in (17).

(17)



An operation that would adjoin a symbol to a dominating symbol is literally unstatable. But any singulary movement T satisfying the extension condition would have to do exactly this. Suppose, for example, we wanted to apply a C fronting type operation (something like Chomsky's T_q) to (14), but which would left-adjoin C to S (in accord with EC), as pictured in (18):

(18)



In the LSLT formalism C and S would both have to be mentioned in the SA. So perhaps the SA could be (19a) or (19b):

(19) a. X–S–C–Y
b. X–C–S–Y

But now look again at the PM (15) to which we would want to apply (19). There is no member of that set that contains both S and C, so the transformation could never apply. This example is completely representative. No movement transformation in the LSLT framework would ever be able to apply in accord with EC.[1]

Interestingly, the L&K framework also forbids EC-satisfying operations, but only by stipulation. Within that model, as noted above, the phrase markers are RPMs, sets consisting of the terminal string and monostrings. Determination of transformational applicability then has to be somewhat different. In particular, it is small sets of monostrings, rather than single ones, that are relevant. L&K provide a definition of precedence between monostrings, and then simply stipulate in their definition of "basic analyzability" that any qualifying set of monostrings

[1]As a reviewer observes, older formulations of the cyclic constraint, as in Chomsky (1965) or the strict cycle condition of Chomsky (1973), do not run into this difficulty, since they only required operations to target topmost domains, and not the root per se.

must be pairwise in the precedence relation. That line of their definition can be eliminated leaving the remainder intact. The effect of this simplification would be to allow a set of monostrings not in the precedence relation, and hence in the dominance relation, to qualify. And this, of course, would allow EC-satisfying operations.

3 Bare phrase structure

Bare phrase structure (BPS, Collins & Stabler 2016 (C&S); Chomsky 2000; Fukui 2011) takes an alternative approach to phrase-markers. BPS uses the set-theoretic $\in$-relation to describe constituency. We fix the instantiation of BPS described in Chomsky (2000; 2008) and formalized in C&S.

In these models, MERGE is a structure-building operations which takes two objects A and B and forms $\{A, B\}$.[2] From this definition, we can recover an "immediately contains" relation between the objects A and B and $\{A, B\}$ by using the elementhood relation. Explicitly, we say that X is *immediately contained* in Y if and only if $X \in Y$.[3] General *containment* is defined as the transitive closure of this relation. Explicitly, we can inductively define containment by saying that X is contained in Y if $X \in Y$ or $X \in Z$ for some Z contained in Y.

Strictly speaking, this is a relation which is defined on the entire model of the ambient set theory, not on a single set X which represents a single syntactic object, as in the case of the precedence and dominance relations between elements of an RPM. That is, containment is a relation between sets in the entire class of sets, not between elements ("nodes") of a single syntactic object. Accordingly, a substructure with respect to the $\in$ relation refers not to a subset of any object in the model, but rather to a *submodel* of the model of set theory.[4]

It is straightforward to show that constituents are not in general subsets of a BPS syntactic object X.

(20) Let A, B, C, and D be lexical items or complex syntactic objects. Construct $X = \text{MERGE}(A, \text{MERGE}(B, \text{MERGE}(C, D))) = \{A, \{B, \{C, D\}\}\}$. Then, $\{C, D\}$ is contained in X, but $\{C, D\} \not\subset X$.

As syntactic objects X are also not models of set theory, but rather the elements of such a model, the submodel relationship which preserves the $\in$ relation also cannot be the correct notion of substructure for syntactic objects.

[2] C&S, Def. 13.
[3] C&S, Def. 8.
[4] Chang & Keisler (1990).

We now present arguments that the $\in$ relation, and its transitive closure, while providing an accurate characterization of the *containment* relation,[5] do not provide a *substructure* relation between syntactic objects. Unfortunately, constituency cannot be used to determine the appropriate notion of substructure, since, in trees, "A contains B" is coextensive with "the constituent dominated by B is a substructure of the constituent dominated by A". In other words, we cannot tell the containment relation apart from substructure inclusions between constituents. However, in slightly relaxed notions of substructures, $\in$ is clearly behaving as a primitive containment relation between nodes, and not a substructure inclusion. We turn to some motivating examples.

In C&S, lexical items are treated as a triple of sets of features (SEM, SYN, and PHON). The features of a syntactic object X are formalized externally with a TRIGGERS function. C&S keep track of which features have been satisfied by removing elements from the sets of features associated to X via TRIGGER. Chomsky suggests in *Categories and transformations* (CT, 1995: Chapter 4) that certain formal features may be *erased* upon satisfaction, or at the interfaces.[6] We first look at how C&S formalize their calculus of features. C&S's feature calculus is meant to capture this intuition.

(21) (C&S Def. 26) TRIGGERS is any function from each syntactic object A to a subset of the trigger features of A, meeting the following conditions:
   i. If A is a lexical item with n trigger features, then TRIGGERS(A) returns all of those n trigger features. (So when $n = 0$, TRIGGERS(A) = {}.)
   ii. If A is a set, then $A = \{B, C\}$ where TRIGGERS(B) is nonempty, and TRIGGERS(C) = {}, and TRIGGERS(A) = TRIGGERS(B) − {TF}, for some trigger feature TF $\in$ TRIGGERS(B). Otherwise, TRIGGERS(A) is undefined.
   iii. Otherwise, TRIGGERS(A) is undefined.

This goes hand in hand with their definition of triggered merge.

(22) (C&S Def. 27) Given any syntactic objects A, B, where TRIGGERS(A) $\neq$ {} and TRIGGERS(B) = {}, MERGE(A, B) = $\{A, B\}$.

The idea is that two items may only merge when one has remaining trigger features, and the other does not. If defined, the trigger features of $\{A, B\}$ are

[5] Ignoring issues relating to "occurrences" of lexical items – i.e. non-tree structures resulting from the elementhood graphs of sets.

[6] Chomsky (1995: 280): "Erasure is a 'stronger form' of deletion, eliminating the element entirely so that it is inaccessible to any operation, not just to interpretability at logical form (LF)."

just those of the triggering object A with the triggering feature removed. Notice, however, that TRIGGER keeps track of the feature changes externally, in that no features of heads contained in A or B are changed. Under such a method, the set-theoretic structure of syntactic objects alone does not encode the featural changes. We want to "internalize" the feature calculus so that MERGE actually results in changes in the structure of the objects it combines.

We have at least two reasonable options for formally realizing these notions of erasure/deletion within a syntactic object itself: by removing the element in question from the syntactic object, or by changing the element in some way which marks it as inoperative. We will show that either method results in an object which the $\in$ relation and its transitive closure both fail to treat as related to the original object in any straightforward way. We will extend the argument to cases of AGREE.

3.1 Method one: Removal of the feature

For any sets A and B, we can construct a set $A - B = \{a \in A : a \notin B\}$, their *difference*, which removes B-elements from A.

Let A be a lexical item and X and Y be syntactic objects (lexical items or otherwise). We treat lexical items as in CT, where A is literally a *set* of features. Take the syntactic object $\text{MERGE}(A, Y) = \{A, Y\}$.[7] Suppose that when this object is merged with X, a feature of the head A is checked, removing $f \in A$, resulting in the object $\{X, \{A - \{f\}, Y\}\}$. Alternatively, if features are not deleted in syntax, we may say that some interface only sees the structure $\{X, \{A - \{f\}, Y\}\}$, which should be a substructure of $\{X, \{A, Y\}\}$.

In the first case, we should like to describe in what sense $\{A - \{f\}, Y\}$ is a substructure of $\{A, Y\}$ in that they have the same phrase structure, with the former simply missing a feature of the latter, so that we can state a form of the EC. In the second case, we should like to describe how $\{X, \{A - \{f\}, Y\}\}$ is a substructure of $\{X, \{A, Y\}\}$.

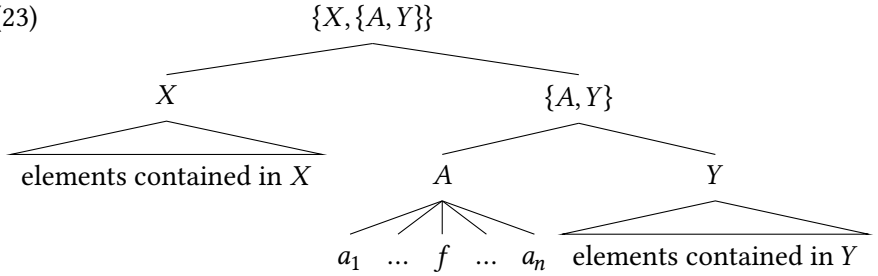
As expected, a subset relation fails to hold in both cases: $\{X, \{A - \{f\}, Y\}\} \not\subset \{X, \{A, Y\}\}$, and $\{A - \{f\}, Y\} \not\subset \{A, Y\}$. However, there is also no containment relation between the syntactic objects. In fact, there is no straightforward set-theoretic relation between these objects. While a subset relation $\{A - \{f\} \subset A\}$ does hold, $\{A - \{f\}, Y\} \not\subset \{A, Y\}$. More generally, for any constituent M containing a head A from which we remove a feature, the resulting constituent M' will simply be a distinct set from M (often with the same number of elements as M). In

[7] For simplicity, we delete no features in the first step, though the argument still holds if we do remove a feature of A (or Y) during this first step.

this example, $\{X, \{A - \{f\}, Y\}\}$ and $\{X, \{A, Y\}\}$ have the same number of elements, though A and $A - \{f\}$ do not, assuming A is finite.

On the other hand, there are canonical ways to draw graph-theoretic objects from well-founded sets. One method produces trees: draw a set X as a root, and write all of its elements as immediate daughters. We repeat the process at each child, writing the same element multiple times if necessary. This process is described in Aczel (1988).

(23)



A *graph* can be defined as a set X together with a relation $R \subset X \times X$. For syntactic objects K, we can define a set X of occurrences of contained elements, with $R \subset X \times X$ being the immediate containment relation between the appropriate occurrences; see C&S (§4, Def. 18) for a formal treatment.

We can define a *subgraph* relation between two graphs (X, R) and (Y, S) if $X \subset Y$ and we have a relation xRx' for $x, x' \in X$ if and only if xSx' in Y. We can then form the graph-theoretic tree associated to $\{X, \{A - \{f\}, Y\}\}$, which is clearly a subgraph of the graph in (23). We could similarly use the containment relation in place of the immediate containment relation, which would describe the syntactic objects as partially ordered sets, with the substructure relation being a subspace inclusion of finite partial orders.

3.2 Method two: Changing (the value of) a feature

Changing the "value" or otherwise adding diacritical marks to an element is another way to formally represent the status of a feature in a syntactic object.

In this case, suppose that we have again constructed $\{A, Y\}$ which we intend to merge with X in a way which will alter a feature $f \in A$. This alteration could be realized as a bijection $m : A \to A'$, where A' is the same set as A, except the feature f has been replaced by $\overline{f}$, the "inoperative" form of f.

However, $\{A, Y\}$ is not a subset of $\{X, \{A', Y\}\}$, nor do we have a containment relation between the two sets. Much like subsets are not the relevant notion of

substructure for BPS sets, neither will bijection be the appropriate notion of isomorphism. For, depending on whether we allow MERGE to combine identical sets or not, every BPS set will have cardinality 1 or 2, and hence be in a bijection with the set $1 = \{0\}$ or $2 = \{0, 1\}$. So while $\{A', Y\}$ and $\{A, Y\}$ are "isomorphic" in that there is a bijection between them, so are they both isomorphic to $\{X, \{A, Y\}\}$, showing that this is not the correct notion of "isomorphism" between the objects, in that it totally ignores constituency.

Again, we may convert $\{A, Y\}$ and $\{X, \{A', Y\}\}$ into graph- or order-theoretic trees. We can define an isomorphism between graphs (X, R) and (Y, S) as a bijection $m : X \to Y$ such that xRx' in X if and only if $(mx)S(mx')$ in Y (or similarly, an isomorphism of partial orders as a bijection $m : P \to Q$ such that $x \leq x'$ in P if and only if $m(x) \preceq m(x')$ in Q). Using these definitions, two graph- or order-theoretic trees (X, R) and (Y, S) will be isomorphic if and only if they have the same number of nodes with the same constituency relations.[8] Using this definition, the graphs associated to $\{A, Y\}$ and $\{A', Y\}$ will be isomorphic, such that $\{A, Y\}$ is isomorphic to a subgraph of $\{X, \{A', Y\}\}$ in the appropriate way.

Alternatively, we might think of this "value" or "activity" as a property of a feature which is explicitly part of its structure. This again has a straightforward formalization when the syntactic objects are graphs: we define a graph-with-value as a graph (X, R) together with a function $v : X \to \{\top, \bot\}$ where we interpret $v(x) = \top$ as meaning "x is inactive". We define a homomorphism between graphs-with-values $f : (X, R, v) \to (X', R', v')$ as a graph homomorphism such that if $v(x) = \top$, then $v'(f(x)) = \top$, i.e. inactive features stay inactive, but active features may be deactivated. Using this structure, the inclusion of an operand A into larger object X, while deactivating a feature in A, would be a homomorphism.

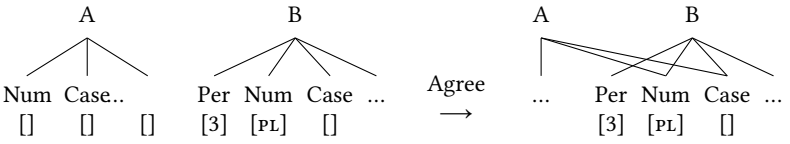
3.3 Agree

The above examples showed that the feature-deletion and feature-valuation methods of modeling MERGE do not lead to substructure embeddings or homomorphisms between BPS sets in any obvious sense. In contrast, relations between derived syntactic objects are straightforward when represented as graphs (possibly with extra structure). Chomsky (1999) has a "valuation" version of agreement, which is subject to similar analysis as the valuation case for selection above. We look now at a feature-sharing approach to agreement, and similarly show that the structural relation between the input structures and output structures is given

[8] Though, this ignores the "occurrence" relations which indicate which nodes are "copies" of others. On the other hand, the multidominant picture of a tree, called the *canonical picture* in Aczel (1988), and given in Fig. 3 in C&S, would not have this issue, and could be used instead.

straightforwardly by graph homomorphisms, while there is no clear associated notion for sets.

Frampton & Gutmann (2000) give an explicit architecture for agreement as feature-sharing using the set-theoretic structure of BPS:

> Consider [(24)] and suppose that Agree applies to the pair of nodes.
>
> (24) $\{Num_1, Case_2, \ldots\}, \{Per_3, Num_4, Case_5, \ldots\}$
>
> [...] suppose that Agree induces feature sharing, so that matching features coalesce into a single shared feature, which is valued if either of the coalescing features is valued. So [(24)] produces:
>
> (25) $\{Num_6, Case_7, \ldots\}, \{Per_3, Num_6, Case_7, \ldots\}$
>
> The value of Num_6 is the coalescence of the values of Num_1 and Num_4. The value of $Case_7$ is the coalescence of the values of $Case_2$ and $Case_5$. New indices were chosen, but index 6, for example, could just as well have been 1 or 4. The choice of index is not a substantive question, assuming that it is suitably distinguished.
>
> If the two coalescing features are both valued to start with, it is not clear that the result is coherent. But this will never arise, because Agree is driven by an unvalued feature. A picture will make the idea clearer. Agree takes [(26a)] into [(26b)], assuming that none of the features indicated by the ellipsis marks match.
>
> (26) a. b.
>
> 
>
>
> (Frampton & Gutmann 2000)

The arrow "Agree" in Frampton & Gutmann's figure can clearly be viewed as a pair of graph homomorphisms from each graph on the lefthand side to the graph on the righthand side, or as single graph homomorphism from the "structured

disjoint union"[9] of the graphs on the lefthand side to the graph on the righthand side. If we view the valuations as properties attached to the nodes of the graph, then we can additionally view this map Agree as a graph homomorphism which preserves those properties (e.g. a PL node gets taken to a PL node).

However, it is again difficult to describe the relationship above when we view the objects as BPS sets. Usually at least one of A or B above will be in a phrase when agreement is applied. Suppose it is B, and we have $B \in \ldots \in X$. We intend to construct from A and X an object $\{A', X'\}$, where A' and X' are exactly A and X, but where the number and case features have been replaced accordingly. Again, we will have no subset, containment, or other obvious set-theoretic relation between A or X and $\{A', X'\}$.

Another application of isomorphism appears implicitly here. Frampton & Gutmann note that the specific index for the element representing the shared feature does not matter, so long as it is suitably distinguished. Again, while the set-theoretic statement of this is somewhat complex (and relies on knowing the specific indices used elsewhere in syntactic objects contained in the current one), the graph-theoretic notion is quite elegant: the righthand side above is determined up to isomorphism of graphs (possibly with values assigned to nodes).

4 The extension condition in the theory of phrase structure

Using LSLT phrase markers, constituents do not arise as substructures in any straightforward way. Accordingly, even if we allow operations which have the effect of the EC, it will not be strictly true that the inputs to the operation are substructures of the output.

In BPS, if we represent feature-changes at all in syntactic objects, either by means of deletion or alteration, it is no longer straightforward in what sense the inputs to MERGE are substructures of or are contained in the output. C&S and Chomsky (2000) only avoid this problem by not annotating the "feature-updates" in the syntactic objects themselves, the former by keeping track of the features in "scoreboard" sets external to the syntactic object (though relevant to determining properties of it, such as labeling), where the latter does not address the treatment of features in syntactic objects formally at all.

If we choose the second method which "alters" features, and implement it in the syntax, then the input $\{A, Y\}$ to MERGE will not be a substructure of the output

[9] Formally, this is the *coproduct* of graphs in the category of directed graphs.

$\{X, \{A', X\}\}$ or immediately contained in it. Similarly, no substructure of $\{A, Y\}$ will be contained in $\{X, \{A - \{f\}, Y\}\}$ if the first method is used in syntax. Both lead to complications in stating the extension condition for BPS.

However, BPS sets can be viewed as an "encoding" of graphs or partial orders using some canonical translation of them. These graphs essentially arise from constructing a set of elements contained in a syntactic object X (possibly with occurrences), and restricting the $\in$ relation between sets to this set. In C&S, many of the important structural properties of syntactic objects – e.g. c-command, relative minimality and maximality (of projections), and specifiers – are similarly defined not on a syntactic object X itself but the associated graph of occurrences of elements contained in X, using a relation based on $\in$ as its "edge relation". Accordingly, the graph- and order-theoretic representation of BPS objects provides a coherent notion of substructure and isomorphism, which makes statement of the EC straightforward using either method described above.

Pure set-theoretic representations limit the distinctions that can be made. To the extent that human language does not rely on the encoding of the mathematically unavailable distinctions, we should favor a theory based on such representations, as we want to limit the descriptive power of the theory as much as is empirically possible, in line with the general Chomskian program. But where we do need to make such distinctions in a full account of human language, we must move to a richer theory of representations, as we have explored here. Studying substructures and isomorphism as they can be used to state the EC provide just one example of how understanding formal properties of the representation of syntactic objects can clarify the relationship between structure-building operations and the properties of the syntactic objects themselves.

Abbreviations

3	third person	PM	phrase marker
BPS	bare phrase structure	RPM	reduced phrase marker
EC	extension condition	SA	structural analysis
LF	logical form	SC	structural change
PL	plural		

Acknowledgements

We are very pleased to help honor Ian Roberts, who has encouraged the field to rethink so many topics in syntax. We are indebted to the two reviewers, whose comments helped us substantially improve the presentation.

References

Aczel, Peter. 1988. *Non-well-founded sets* (CSLI Lecture Notes 14). Stanford, CA: CSLI Publications.

Chang, Chen Chung & H. Jerome Keisler. 1990. *Model theory*. Elsevier.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1973. Conditions on transformations. In Stephen Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, 232–286. New York: Academic Press.

Chomsky, Noam. 1975 [1955]. *The logical structure of linguistic theory*. New York: Plenum.

Chomsky, Noam. 1993. A minimalist program for linguistic theory. In Kenneth L. Hale & Samuel Jay Keyser (eds.), *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1999. Derivation by phase. *MIT Occasional Papers in Linguistics* 18.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Collins, Chris & Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19(1). 43–78. DOI: 10.1111/synt.12117.

Frampton, John & Sam Gutmann. 2000. Agreement is feature sharing. Ms., Northeastern University.

Fukui, Naoki. 2011. Merge and bare phrase structure. In Cedric Boeckx (ed.), *The Oxford handbook of linguistic minimalism*, 73–95. Oxford: Oxford University Press.

Lasnik, Howard. 2006. Conceptions of the cycle. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Wh-movement: Moving on*, 197–216. MIT Press.

Lasnik, Howard & Joseph J. Kupin. 1977. A restrictive theory of transformational grammar. *Theoretical Linguistics* 4(1–3). Reprinted in Lasnik, Howard. 1990. *Essays on restrictiveness and learnablity*, 17-41. Dordrecht: Kluwer, 173–196. DOI: 10.1515/thli.1977.4.1-3.173.

Chapter 11

Strong and weak "strict cyclicity" in phase theory

Ángel J. Gallego
Universitat Autònoma de Barcelona

This paper explores the possibility that the no tampering condition (NTC) is eliminated in favor of a strong version of the phase impenetrability condition (PIC). This possibility is welcome on theoretical grounds, given the redundant nature of the NTC and the PIC. I review empirical evidence indicating that the (original formulation of the) NTC is violated phase-internally, a possibility that does not extend to the PIC. In so doing, I also consider the weak version of the PIC discussed in Chomsky (2016).

1 Efficient computation

Generative Grammar has endorsed various economy principles (from Chomsky's 1975 [1955] *traffic convention* to Chomsky's (1995) *minimal link condition*, going through many others). All such proposals adhere to a "least effort" desideratum attributed to the syntactic computation of the faculty of language. Within the Minimalist program (MP), the basic structure-building operation is Merge – the only one that "comes free," without justification (Chomsky 2001: 3; 2008: 137).

Assuming it operates without bounds, Merge takes two objects, α and β, to construct a new object, γ. Additional applications of Merge target γ, which is the only object left in the derivation (Chomsky 1995: 243), to yield γ′, and then γ″, and so on and so forth – again, without bounds:[1]

[1]In Chomsky (2007: 11; 2008: 139) it is assumed that the free nature of Merge follows from LIs having an EDGE FEATURE (EF) that is undeletable and can thus give rise to an unbounded application of Merge. I will not assume EFs. Apart from the empirical advantage of dispensing with EFs (they have no realization in any language, so they are a purely theory-internal device),

Ángel J. Gallego. 2020. Strong and weak "strict cyclicity" in phase theory. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 207–226. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280647

(1) a. Merge(α, β) = {α, β}
b. Merge(λ, γ) = {λ, γ}
c. Merge(ψ, γ′) = {ψ, γ′}

That α and β are no longer available was expressed in the following passage:

> Applied to two objects α and β, Merge forms the new object K, *eliminating α and β*. (Chomsky 1995: 243, my emphasis)

A Merge-based system is enough to capture the property of CYCLICITY, that is, "in essence, the intuition that the properties of larger linguistic units depend on the properties of their parts" (Chomsky 2012: 1).[2] It is easy to see that a cyclic system will be largely compositional (Chomsky 2007: 5; 2012: 2): if computation is meaningful in an efficient manner, the interpretation of a given linguistic object will not be changed later on, which corresponds with "the general property of STRICT CYCLICITY" (Chomsky 2007: 5).[3] Therefore, whereas cyclicity follows from Merge alone, strict cyclicity requires something else – the mere existence of such an operation does not in and of itself guarantee the conservation of the already assembled structure. This is the natural scenario where MP invokes so-called third factor conditions, which fall into two broad categories (Chomsky 2005):

(2) Third-factor conditions
a. Principles of data analysis that might be used in language acquisition and other domains;
b. Principles of structural architecture and developmental constraints that enter into canalization, organic form, and action over a wide range, including *principles of efficient computation*, which would be expected to be of particular significance for computational systems such as language. It is the second of these subcategories that should be of particular significance in determining the nature of attainable languages. (Chomsky 2005: 6, my emphasis)

this allows us to dispense with the technical problems discussed in Narita (2014), related to the lack of EF percolation.

[2] As an anonymous reviewer observes, this is not the case if Merge allows, e.g., countercyclic infixing of SPEC-T after C has already been merged (see Chomsky 2008), or Parallel, Sidewards, Late, etc. Merge. Cf. Chomsky et al. (2019) and references therein for discussion.

[3] Of course, the interpretation of "Mary" is different in *Someone called Mary* and *Mary called someone*. That the interpretation of a given SO cannot be changed should thus be restricted to a post-Merge scenario, a possibility that is not entertained in feature-based approaches to theta-roles.

Different conditions have been put forward in order to capture the idea that linguistic objects generated by the syntactic computation cannot be changed (where *change* covers a wide range of possibilities: deletion, feature-valuation, late-insertion, tucking-in, etc.), especially by adding ad hoc symbols or performing operations that depart from least effort metrics. This is precisely the role played by the INCLUSIVENESS CONDITION (IC, Chomsky 1995: 228), the NO TAMPERING CONDITION (NTC, Chomsky 2008: 138), and the PHASE IMPENETRABILITY CONDITION (PIC, Chomsky 2000). Putting details aside, IC, NTC and PIC all play a similar role in the current model, which was already noted by Juan Uriagereka in his annotated version of Chomsky (2001):

> So the Extension Condition [still holds]. This is somewhat surprising, given the [adoption of] "tucking-in" in Chomsky (2000). In effect, *we have several things ensuring the cycle*. The EC, in a radical way for the upward boundary of the phrase marker; the PIC for a kind of downward boundary, beyond which the system doesn't see any further operations; the idea of interpretation/evaluation at the strong phase in addition to both of these, as the derivation unfolds; and, finally, the phase-like access to the Numeration. *Much room for improvement and unification …*
>
> (Uriagereka 1999a, my emphasis)

Such a redundant scenario is not expected, if only at a purely methodological level. This note argues that (the strong version of) the NTC can be subsumed under the PIC, given that local (phase-internal) modification is possible.[4] Discussion is divided as follows: §2 reviews the different conceptions of the NTC that have been entertained within MP and the empirical problems that have been observed for it; §3 turns its attention to the PIC, focusing on the recent possibility that the complement of a phase does not leave the computation (Chomsky 2008; 2016); in §4, I argue that (the strong) NTC can be eliminated adopting a strong version of the PIC, whereby transferred computation is forgotten (literally expunged), yielding a straight version of strict cyclicity; §5 summarizes the main conclusions.

2 Merge and the NTC

There is a very close relationship between Merge and the NTC on the one hand, and between Transfer and the PIC on the other (as we will see in more detail

[4] Probably, the same can be said of the IC, by simply observing that labels, indices, traces, and similar devices are not part of any I-language.

in §3). In fact, I would like to underscore the fact that, whereas Transfer and the PIC (as well as the operations of Feature Inheritance (FI) and Agree)[5] apply at the phase level, Merge and the NTC do not invariably so (Chomsky 2007: 17; 2008: 143; 2013: 40, 42). I state this correlation as follows, which I would like to build on to argue that there is a deep connection between the phase-based architecture and the (mildly) context-sensitive nature of the Faculty of Language (cf. Chomsky 1956; Uriagereka 2008):[6]

(3) a. EM = context-free
b. IM/Agree/Transfer = (mildly) context-sensitive

In what follows I would like to briefly review the different formulations of the NTC. As the reader will see, the conclusion will be that there are various situations where a weak version of the NTC must be assumed, not only for operations like FI or Agree (Chomsky 2007: 19, fn. 26),[7] but also for Merge.

In Chomsky (2000; 2001; 2004; 2005), no explicit mention to the NTC is made. Instead, the EXTENSION CONDITION (EC) is responsible for capturing the idea that Merge always applies to the edge of an SO. Thus, EC makes sure that, given {α, β}, a new element δ can only be merged as in (4a), not (4b), which would be counter-cyclic.

(4) a. {δ, {α, β}}
b. {{α, δ}, β}}

Chomsky (2000: 136) discusses these options, noting that (4a) satisfies the EC whereas (4b) satisfies Local Merge. In the same breath, he notes that

> weaker assumptions suffice to bar [(4a)] but still allow Local Merge under other conditions. Suppose that operations do not tamper with the basic relations involving the label that projects: the relations provided by Merge and composition, the relevant ones here being sisterhood and c-command. (Chomsky 2000: 136)

[5] I assume that Agree actually implies a complex set of operations: Feature Inheritance, Match, Valuation and Deletion. Deletion is meant to cover erasure of uninterpretable φ-features, but it can also be applied to heads, as in Chomsky's (2015) analysis of *that*-deletion. Cf. Epstein et al. (2016) alternative in terms of phase-cancellation. Cf. Gallego (2014) for an alternative approach to FI, with interesting consequences for Chomsky's (2015) analysis of the EPP, discussed in Gallego (2017).

[6] It is typically assumed that all operations but EM apply at the phase level, simultaneously (Chomsky 2004: 116; 2005: 19; 2007: 17; 2008: 155). This raises questions for derivational systems, where the application of rules is ordered, as in Chomsky (2015).

[7] FI is reinterpreted as copying in Chomsky (2013: 47). This also departs from the strong NTC (unless we adopt the formulation in Gallego 2014).

Chomsky (2000: 137) goes on to argue that "derivations then observe the condition [(5)], a kind of economy condition, where R is a relevant basic relation".[8]

(5) Given a choice of operations applying to α and projecting its label L, select one that preserves R(L, γ)

(5) holds in general, except for head adjunction. In the case of XP merger, Chomsky (2000) observes that EC must be satisfied for second-Merge, but not for subsequent applications or Merge – the creation of specifiers, which amounts to accepting tucking-in (Richards 1997).

In Chomsky (2004), it is explicitly noted that the EC can come in a strong and a weak version, the latter accepting deviations from (5):

> Cyclicity of derivation requires that Merge to α always be at the edge of α, satisfying *an extension condition, strong or weak* ("tucking in") […] There appears to be one significant counterexample to cyclic Merge: late insertion of adjuncts […] Elementary considerations of efficient computation require that Merge of α to β involves minimal search of β to determine where α is introduced, as well as least tampering with β: search therefore satisfies [Local Merge], and Merge satisfies an EC, with zero search. One possibility is that β is completely unchanged (the *strong EC*); another natural possibility is that α is as close as possible to the head that is the label of β, so that any Spec of β now becomes a higher Spec ("tucking in," in Norvin Richards's sense). Further questions arise under Merge with multiple Specs. Assume some version of the EC to hold, in accord with SMT. (Chomsky 2004: 109, my emphasis)

The NTC is first introduced in Chomsky (2005), when discussing conditions of efficient computation. What I would like to capitalize on from the following quote is how similar NTC and PIC are, in the sense that the former appears to be related to the fact that what has been constructed in the course of a derivation *can be forgotten*; this is relevant, since this is typically the hallmark of the PIC.

> One natural property of efficient computation, with a claim to extralinguistic generality, is that operations forming complex expressions should consist of no more than a rearrangement of the objects to which they apply, not modifying them internally by deletion or insertion of new elements. If

[8] This is what Lasnik & Uriagereka (2005: Ch. 2) and Epstein et al. (2012: 256) refer to as Law of Conservation of Relations.

> tenable, that sharply reduces computational load: *what has once been constructed can be "forgotten" in later computations, in that it will no longer be changed*. That is one of the basic intuitions behind the notion of cyclic computation. The EST/Y-model and other approaches violate this condition extensively, resorting to bar levels, traces, indices, and other devices, which both modify given objects and add new elements. A second question, then, is whether all of this technology is eliminable, and the empirical facts susceptible to principled explanation in accord with the "no-tampering" condition of efficient computation […] Assuming the NTC that minimizes computational load, both kinds of Merge to A will leave A intact. That entails merging to the edge, the EC, which can be understood in different ways, including the "tucking-in" theory of Richards (1997), which is natural within the probe-goal framework of recent work, and which can also be interpreted to accommodate head adjunction. (Chomsky 2005: 11, 13, my emphasis)

Notice that what this says is that the NTC is a third-factor condition on the way Merge operates.[9] More precisely, the NTC guarantees that when Merge applies to α and β, we obtain a new SO, γ, which can then be merged with further objects. So, for instance, if γ is merged with δ, given that α and β themselves are gone from the computation, the only way for this to happen is by forming {γ, δ}. This way, Merge must be to the edge as it cannot tamper with the objects it applies to – in the case at hand, Merge cannot break up γ or tamper with it.

What is relevant about Chomsky (2008) is the discussion of certain situations that threaten the strong NTC: FI and the analysis of subject raising to SPEC-T.

> A natural requirement for efficient computation is a "no-tampering condition" (NTC): Merge of X and Y leaves the two SOs unchanged. If so, then Merge of X and Y can be taken to yield the set {X,Y}, the simplest possibility worth considering. Merge cannot break up X or Y, or add new features to them. Therefore Merge is invariably "to the edge" and we also try to establish the [IC] dispensing with bar levels, traces, indices, and similar descriptive technology introduced in the course of derivation of an expression […] Note that SMT might be satisfied even where NTC is violated – if the violation has a principled explanation in terms of interface conditions (or perhaps some other factor, not considered here). The logic is the same as in the case of the phonological component, already mentioned […] *The device of inheritance* […] *is a narrow violation of NTC*. The usual question therefore

[9] This formulation states that the NTC is Merge-sensitive alone, which opens the door for conditions being sensitive to independent operations.

> arises: does it violate SMT? If it does, then the device belongs to UG (perhaps parametrized), lacking a principled explanation. But the crucial role it plays at the C-I interface suggests the usual direction to determine whether it is consistent with SMT though violating NTC. If the C-I interface requires this distinction, then SMT will be satisfied by an optimal device to establish it that violates NTC, and inheritance of features of C by the LI selected by C (namely T) may meet that condition. If so, the violation of NTC still satisfies SMT. (Chomsky 2008: 138, 144, my emphasis)

Chomsky (2007; 2008) assumes that φ-features are generated in phase heads, from which they are downloaded (downward percolation) to non-phase heads. Following Richards (2007), the process is taken to be mandatory under the PIC: Since these features must be deleted, they must end up in the Transfer domain.[10] FI has consequences for the analysis of raising-to-subject, as discussed by Epstein et al. (2012). In particular, suppose the derivation of *Don Quixote fought the windmills* is as depicted in (6):

(6) a. {Don Quixote, {v*{fought, {the,windmills}}}} = v*P
b. Merge (T,v*P) = {T, {Don Quixote, {v*{fought, {the,windmills}}}}}
c. Merge (C,TP) = {Cφ, {T, {Don Quixote, {v*{fought, {the,windmills}}}}}}
d. FI (C,T) = {C, {Tφ, {Don Quixote, {v*{fought, {the,windmills}}}}}}
e. IM (*DQ*,TP) = {C, {Don Quixote, {Tφ, {*t*, {v*{fought, {the,windmills}}}}}}}

The problematic steps in (6) are (d) and (e), but (e) more clearly so. As Epstein et al. (2012) discuss, the original (SPEC-less) TP must be disconnected from C so that the external argument (EA) *Don Quixote* undergoes IM with it; when this new (SPEC-ful) TP is created, and it is then reconnected to C. The operation is thus ternary, in that Merge must target the EA, TP, and C. Noam Chomsky (p.c.) notes that this is a narrow extension of Merge, but does not depart from it in the way head movement does, since the EA is merged with TP, which it is a term of.

So far, as we can see, a key trait of NTC/IC-constrained Merge (α, β) is that α and β cannot be modified: they are left unchanged, no features, indices, etc. can be added to them by Merge. Chomsky (2007) gives another twist by noting that while Merge cannot modify α or β, some subsequent operation might:

[10] As pointed out in footnote 7, Chomsky (2013) suggests that FI is actually a form of copying. If correct, FI could simply be reduced under the copy theory of movement, as argued in Gallego (2014).

> Merge (X1,…,Xn) = Z, some new object. In the simplest case, n = 2, and there is evidence that this may be the only case (Richard Kayne's "unambiguous paths"). Let us assume so. Suppose X and Y are merged. Evidently, efficient computation will leave X and Y unchanged (the no tampering condition NTC). We therefore assume that NTC holds unless empirical evidence requires a departure from SMT in this regard, hence increasing the complexity of UG. Accordingly, we can take Merge (X,Y) = {X,Y}. *Notice that NTC entails nothing about whether X and Y can be modified after Merge* […] Under NTC, merge will always be to the edge of Z, so we can call this an edge feature EF of W. (Chomsky 2007: 8, my emphasis)

This observation can probably be related to Chomsky's (2015: 10–11) analysis of phase-head deletion (de-phasing), which triggers a process that makes a non-phase head inherit all the properties of a phase head. De-phasing is put forward in order to account for the fact that subjects can be extracted from *that*-less clauses (an empty category principle (ECP) violation in earlier terminology). So, as is well-known, subject extraction across a CP is ruled out if that is spelled out (cf. Chomsky 1986; Rizzi 1990):

(7) $[_{CP}$ Who does the book say $[_{CP}$ (*that) $[_{TP}$ t_{Who} stabbed Caesar]]]?

Chomsky (2015) reinterprets this phenomenon in order to argue that C can undergo deletion. This makes T inherit phasehood, which makes it strong, with no need for a DP to occupy SPEC-T for labeling reasons (cf. Gallego 2017). More to the point, Chomsky (2015: 11) argues that "The natural assumption is that phasehood is inherited by T […] along with all other inflectional/functional properties of C (φ-features, tense, Q), and is activated on T when C is deleted".[11]

Let us take stock. NTC is the formalization of the idea that computation applies in an efficient way, so that Merge (α, β) cannot modify α and β themselves. This strong formulation of the NTC, which bars tucking in and derives the copy theory of movement (CTM), captures more than mere cyclicity. In particular, what I would like to emphasize is that by not letting Merge modify what it applies to, the NTC further captures some form of strict cyclicity too. To see this, let us go back to (1), repeated as (8) below:

(8) Merge (α, β) = {α, β} = γ

After (8), the workspace contains γ and nothing else, so α and β are no longer available (Chomsky 1995: 243). At this point, we may want to merge γ and a new object, δ:

[11]Noam Chomsky (p.c.) elaborates on this by noting that the NTC states that an SO should not be modified by Merge, which doesn't literally imply that it cannot be deleted.

(9) Merge(δ, γ) = {δ, γ}

δ is either internal or external to γ. If external, δ is drawn from the lexicon. This is External Merge (EM). If internal (e.g., δ = α), then δ is a term of γ. Assuming the NTC, γ cannot be modified, so it must remain {α, β}, which yields {α, {α, β}}, and thus two copies (occurrences) of α. More importantly for our purposes, the strong NTC entails that {α, β} must be left as it is, so merger of α will not tamper with γ by removing α. There is no need for an extra operation (Copy) for IM, just like it is not needed for EM – if α were taken from the lexicon, it would not be copied.[12]

This said, there are two potentially problematic aspects about the NTC. The first one follows from the very fact that the strong NTC runs into the empirical problems in (10):[13]

(10) Violations of strong NTC

a. Feature Inheritance (Chomsky 2008)
b. IM to SPEC-T (after EM (C,TP)) (Chomsky 2008)
c. Tucking-in (Richards 1997)
d. Head movement (Chomsky 2001)
e. De-phasing (Chomsky 2015)
f. Phase-cancellation (Epstein et al. 2016)

Apart from these *local* (phase-bounded) violations of the NTC, there is another important observation to be made about the strong NTC, namely the redundancy between it and the PIC, as I discuss in the following section.

[12]The problem is more general if α and β remained in the workspace, along with γ. As Noam Chomsky (p.c.) points out, it has always been assumed that they do not, for the generative procedure constructs a single object, not a multiplicity of objects. Changing that convention would mean that instead of a generative process for expressions, we would be designing a generative process for an arbitrarily large collection of expressions. For instance, suppose that we hold that after EM(α, β) = γ = {α, β}, the workspace contains α, β, γ. We then have a new question: what is the relation between α in the workspace (call it α1) and α in γ = {α, β} (call it α2)? They are either copies or repetitions. If they are copies, everything goes haywire. Thus, if we continue to Merge to α1 finally yielding the finite clause FC, and to γ yielding the finite clause FC′, then the two clauses would contain the two copies α1 and α2, so one should be deleted, and if one enters into some relation (say anaphora) then the other does, etc. Things get much worse if, as this proposal allows, we construct simultaneously indefinitely many finite clauses. This is not only dubious, and in fact makes the notion of "copy" collapse.

[13]If the NTC is restricted to Merge, as Noam Chomsky (p.c.) notes, then only (10b) and (10c) are problematic.

3 Transfer and the PIC

We have seen that the NTC has two formulations, strong and weak. Let me express this as follows:

(11) a. Strong NTC (NTC_S) = SOs cannot be changed by Merge
b. Weak NTC (NTC_W) = SOs can be changed locally, but not by Merge

What I would like to discuss is the fact that NTC_S is virtually analogous to the PIC. The PIC was proposed in order to capture strict cyclicity, so that "operations cannot 'look into' a phase below" (Chomsky 2000: 108). Chomsky (2004) relates the PIC to the operation Transfer (a wider version of Spell-out, capturing the interaction between NS and both interfaces), which is defined in (12):

(12) Transfer hands D-NS over to Φ and to Σ. (Chomsky 2004: 107)

In Chomsky (2004), Transfer makes it impossible for the externalization systems to access what has been cashed out at previous phases. The possibility that the same happens in the case of the narrow computation is not so clear:

> When a phase is transferred to Φ, it is converted to PHON. Φ proceeds in parallel with the NS derivation. Φ is greatly simplified if it can "forget about" what has been transferred to it at earlier phases; otherwise, the advantages of cyclic computation are lost […] PIC sharply restricts search and memory for Φ, and thus plausibly falls within the range of principled explanation […] *It could be that PIC extends to NS as well, restricting search in computation to the next lower phase.* (Chomsky 2004: 107, my emphasis)

That the PIC does not carry over to the computation is connected to the existence of structures, in Icelandic or Spanish, like those in (13), where T can agree with the in-situ internal argument (IA):

(13) {T, {v*, {V, IA}}}
Agree

Empirically, (13) requires the φ-probe to override the PIC and access the complement domain of v* (see Richards 2012). In order to tackle this, Chomsky (2001; 2004) adopts a weak version of the PIC, which led to a scenario analogous to that of the NTC, with both strong and weak versions:

(14) a. Strong PIC (PIC1 or PIC_S)
In phase α with head H, the domain of H is not accessible to operations outside α; only H and its edge are accessible to such operations. (Chomsky 2000: 108)

b. Weak PIC (PIC2 or PIC_W)
[Given structure $[_{ZP}$ Z … $[_{HP}$ α [H YP]]], with H and Z the heads of phases]: The domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations. (Chomsky 2001: 14)

PIC2 is incompatible with FI, so in Chomsky (2008) it is discarded. Consider the following discussion, which suggests that phases that have been transferred can in principle be accessed (modulo intervention effects). Chomsky concludes that the effects of the PIC hold for the interfaces, but not necessarily NS:

> For minimal computation, as soon as the information is transferred it will be forgotten, not accessed in subsequent stages of derivation: the computation will not have to look back at earlier phases as it proceeds, and cyclicity is preserved in a very strong sense. Working that out, we try to formulate a PIC, conforming as closely as possible to SMT […] *Note that for narrow syntax, probe into an earlier phase will almost always be blocked by intervention effects*. One illustration to the contrary is agreement into a lower phase without intervention in experiencer constructions in which the subject is raised (voiding the intervention effect) and agreement holds with the nominative object of the lower phase (Icelandic). *It may be, then, that PIC holds only for the mappings to the interface, with the effects for narrow syntax automatic*. (Chomsky 2008: 143, my emphasis)

Chomsky (2016) in fact argues that Transfer should not eliminate anything from the NS. Otherwise, it would not be possible to explain how the structures in (15) are formed:[14]

(15) a. $[_{\alpha}$ The idea $[_{\beta}$ that the Earth is round]] was rejected t_{α}

b. $[_{\alpha}$ That $[_{\beta}$ I kept my job]] seems to t_{α} bother Mary

The problem here is as follows: in both cases, β is a phase, so it should be transferred before α is raised to matrix SPEC-T. But how can β be pronounced along with α if it is gone from the computation? Chomsky (2016) claims β is never gone from the workspace, but rendered inaccessible by Transfer. There

[14] I put aside another situation where the PIC is strongly violated: covert movement. This matter is pointed out (not addressed) in Chomsky (2004: 111; 2005: 13).

are two ways to interpret this version of the PIC, which I will call PIC3: what's been processed is either (i) totally inaccessible or (ii) cannot be changed.[15] Given the data in (15), (i) must be dismissed. We therefore expect that violations of the PIC do not change whatever is inside the transferred phase. This crucially allows us to change what is outside it, including the φ-probe of matrix T in (16), taken from Fernández-Serrano (2016):

(16) Spanish
Me encantan [$_{CP}$ PRO escuchar [$_{v*P}$ t_{PRO} t_{v*} [$_{VP}$ V truenos]]]
to.me love-3.pl listen thunder
'I love to listen to thunder.'

Let us therefore assume the PIC3 allows access into a lower phase, as long as it is not modified. This makes it difficult to keep the copy/repetition distinction. Take (17), call it K, where the lower phase complement containing β, that is {α, β}, has already been transferred:

(17) K = {…{P, {α, β}}

Imagine we now merge β with K. β could be taken from the lexicon, so it would be a repetition. Can it be a copy? Given that {α, β} is not expunged from the derivation, the question is whether NS can tell whether β is taken from the lexicon or it is interpreted as an occurrence of the β contained within P's complement. If {α, β} can be accessed, the system cannot tell the difference. But we want to exclude this, or successive cyclic movement would go away. Island conditions would be affected too. Notice that the logic here is clear: the copy/repetition distinction does not require changing anything within the already passed phase. So, it should be possible to do that, given Chomsky's (2016) PIC3.

A way out would be to assume, as Noam Chomsky (p.c.) suggests, that if β raises from {α, β}, then both {α, β} and β itself have been modified: {α, β} by now containing a copy that is part of chain, and β by the mere fact of becoming a discontinuous object. Now, if this is correct, even the application of IM to *Who* changes the v*P and *Who* in (18).

(18) {Who, {Samson, {v*, {defeated, *t*}}}}

[15] A reviewer points out that what I call PIC3 is actually a conception of Transfer and its effect on transferred material, not the PIC, which "describes the timing of Transfer and the size of the transferred object". For the purposes of this paper, I will not dwell on this (to me, largely terminological) issue. The PIC was meant to state what is accessible and what is not after Transfer (a mapping operation) applies. All I am assuming is that the PIC3 says that everything is actually accessible after Transfer as long as it is not changed.

Presumably, this has not been considered problematic, for it does not violate the PIC, but it does the NTC_S. Now, we have seen that NTC_S and PIC are remarkably similar in that they both capture strict cyclicity. If nothing else, (18) shows another scenario where I depart from the NTC_S. I take this to indicate that the NTC_S is to be dispensed with entirely. More controversially, I also argue that the NTC_W is dispensable, *if* the PIC can play its role. Under PIC1, which I repeat here as (19), this replacement is possible:

(19) Strong PIC (PIC 1 or PIC_S) In phase α with head H, the domain of H is not accessible to operations outside α; only H and its edge are accessible to such operations. (Chomsky 2000: 108)

What (19) says is enough to capture the effects of the NTC. In particular, the fact that the objects generated in the course of the derivation cannot be tampered with. Notice that this *does* allow tampering *before Transfer applies*, but we have seen that this is empirically sustained. To the cases listed in (10), we can add a sixth one, which follows from the PIC3:

(20) Violations of NTC_S

a. Feature Inheritance (Chomsky 2008)
b. IM to SPEC-T (Chomsky 2008)
c. Tucking-in (Richards 1997)
d. Head movement (Chomsky 2001)
e. De-phasing (Chomsky 2015)
f. Phase-cancellation (Epstein et al. 2016)
g. IM (chain creation)

In the next section, I would like to summarize the main ideas of the previous pages and, at the same time, argue that the PIC3 can be eliminated in favor of the PIC1. In so doing, I also discuss how the data mentioned in Chomsky (2016) can be handled under such proposal. The proposal entails that Transfer eliminates material from the workspace, yielding a more effective reduction of computational load – the original motivation behind phase theory (cf. Chomsky 2000).

4 NTC eliminated: Some consequences

Let me spell out the interim conclusions so far. I will phrase them as questions:

(21) a. Do we need both NTC and the PIC?
b. If we need the PIC, do we need the PIC3?

Both NTC and PIC express an efficiency desideratum, namely that a given SO should not be changed (manipulated, tampered with, altered, etc.) once it has been created. This creates a redundancy, as I have pointed out.[16] At the same time, we have seen different phenomena indicating that the strong version of the NTC cannot be maintained. Should the weak version be? I think it should not, just like the weak PIC (the one in Chomsky 2001). This raises the more general question whether the strong PIC could be the only cyclic principle. If so, then the derivation can allow tampering up to the phase level, when Transfer applies. Suppose the derivation has assembled α and β to yield this:

(22) {α, β}

Suppose next that we apply IM to β. If the NTC does not hold, this could yield (23), potentially affecting the CTM.

(23) {β, {α}}

Note that this derivation is not forced (thus, the CTM does not go away), but the question is whether the step in (23) creates a problem. It is not clear that it does, at least if something like (23) is at stake for de-phasing (cf. Chomsky 2015).

If the only cyclic condition is the PIC, the next question is (21b). Recall that there are two empirical arguments to sustain it. The agreement facts (cf. 16) could be tackled if Agree takes place at the border of NS-externalization, not in NS. This would have two welcome consequences. On the one hand, we could explain the parametric nature of Agree, which I would like to relate to Chomsky's (2014) *thesis T*:

(24) Language is optimized relative to the conceptual-intentional (CI) interface alone, with externalization a secondary phenomenon. (Chomsky 2014: 7)

[16] A reviewer does not see the redundancy, as (s)he takes the NTC to be a third-factor condition on Merge (defining a Merge-cycle that adds stuff to the derivation) and the PIC to be a natural result of Transfer (which removes stuff from the workspace). Given the (empirical) arguments given below (and in Chomsky et al. 2019) it is unlikely that the PIC actually removes stuff from the workspace.

The *thesis T* tells us that efficiency of operations should be found in the NS → SEM channel, not in the NS → PHON one, which is further consistent with the claim that "language is primarily an instrument of thought, with other uses secondary" (Chomsky 2014: 7). If Agree is pushed to NS → PHON, then the fact that its effects are subject to parametrization (as appears to be the case), would fall into place, and would also be compatible with the idea that language variation and parametrization are to be found only there (Chomsky's 2001 uniformity principle; cf. Chomsky 2010; Berwick & Chomsky 2011).

Another consequence of this concerns the very nature of Agree, which is a complex operation, consisting of Match, Valuation, Transfer and Deletion. Chomsky (2004 et seq.) takes these operations to somehow apply simultaneously (at the phase level), but this is hardly consistent with a derivational system, for operations must be ordered (as in Chomsky 2015).[17] Plausibly, the operations should be ordered as follows:

(25)
1. Match (NS)
2. Valuation (NS)
3. Transfer (NS → SEM/PHON)
4. Deletion (PHON)

As noted in Epstein & Seely (2002), this timing is problematic, since it entails that uninterpretable features will be valued before Transfer, becoming undistinguishable from interpretable ones. Unless Deletion could apply at SEM too somehow deleting the uninterpretable, but valued, φ-features of v* and C, operations would have to apply simultaneously, which, as noted, is odd within a derivational system. A way out is at hand if the derivation can somehow remember that φ-features were introduced as unvalued. This should be possible, given the relevance of phase-level memory to distinguish trivial/non-trivial chains, which in its most direct interpretation would entail revamping the long-abandoned idea or feature chains (Chomsky 1995: 262, 270–271, 383, fn. 27, abandoned in Chomsky 2000 due to the intricacies of head movement). So, if Merge could apply not only to LIs, but also to features – more precisely, to their values, which is what seems to be copied from one LI to another, then this would assimilate Valuation to Merge, making it possible for the system to remember that a valued feature was introduced as unvalued, which would signal it as uninterpretable. The technical solution I am sketching would not be too different from FI itself. In brief,

[17] If Transfer is part of externalization, then it can be subject to parametrization (for the same reasons Agree would be). This opens the possibility that the effects of Transfer vary from language to language (cf. Uriagereka's 1999b radical or conservative Spell-out).

we could dispense with the simultaneity of operations and perhaps the need for Agree to apply in NS alone if Merge could apply to LIs, features and values.

Obata's (2010) data are different. Consider (26):

(26) [α That [β Judas left the dinner]] seemed [to t_{α} worry everyone]]

Here β is transferred before α is raised to matrix SPEC-T, which makes it impossible for it to be spelled-out where we see it. However, even if we assumed that the PIC leaves β accessible (through the PIC3), this does not cover IM. That is, it is only α (presumably its head, *that*) that can raise to matrix SPEC-T, so how can β be pied-piped along with α? If we allowed that, then we would also be changing the already transferred object, as noted for (18) above. A possible way out for these cases is that what is transferred is turned into a pair ⟨X,Y⟩. I would like to connect this to Chomsky's (2004) analysis of adjuncts, which adopted (27):[18]

(27) In ⟨α, β⟩, α is spelled out where β is. (Chomsky 2004: 199)

If Transfer converts the structure into some kind of pair, then when IM targets α, the actual pronunciation of β (or some part of it) could be possible. This would have the effect of placing β in a "secondary plane" (Chomsky 2004), but we want α (the phase edge), and α alone, to remain in the primary plane. This is what the PIC1 bought us, which brings back the possibility that Transfer can yield (28), removing the complement domain from NS (cf. Ott 2011):

(28) a. {Edge, {P, {β}}}
b. Transfer (β) = {Edge, {P}} or {Edge, P}

If Transfer applies this way, there would be tampering, but locally. (28) would make it possible for P to be the head of the entire phase, with consequences for the v*-EA relation (cf. Epstein & Shim 2015).

5 Conclusions

This paper has discussed the nature of different conditions put forward to capture computational efficiency within minimalism, most importantly, the NTC and the PIC. Given their redundant nature (they both aim at capturing the idea behind the strict cycle, namely that SOs formed in the course of a derivation

[18] Cf. Chomsky (2008: 139) for similar ideas in the case of Merge.

cannot be changed at subsequent stages), one of them should be dispensed with. I have argued that strict cyclicity effects follow from the PIC alone. The decision is justified on methodological and empirical grounds. The former have to do with the multiplicity of conditions favoring strict cyclicity. The latter concern the empirical evidence showing that the strong version of the NTC cannot be maintained.

The strong PIC (or PIC1 cf. Chomsky 2000), which is the one that should be adopted, forces successive cyclic movement (SCM). Since nothing is left in the (primary plane of) computation after Transfer, that's the only way for a chain to be created. It also follows that the SO that has been cashed out cannot be modified: it is gone from the workspace. Interestingly, there are no violations of the PIC analogous to those of the strong NTC, which is another argument to stick to the former. Interestingly, it seems that only CP and vP give rise to SCM – NPs, PPs and other categories lack it (cf. Gallego 2012; van Urk 2016), which may provide yet another reason to defend that only CP and vP are phases.

Abbreviations

CI	conceptual-intentional	IC	inclusiveness condition
CTM	copy theory of movement	IM	Internal Merge
EA	external argument	MP	Minimalist program
EC	extension condition	NS	narrow syntax
ECP	empty category principle	NTC	no tampering condition
EF	edge feature	PIC	phase impenetrability condition
EM	External Merge		
EPP	extended projection principle	SCM	successive cyclic movement
EST	extended standard theory	SMT	strong Minimalist thesis
FI	Feature Inheritance	SO	syntactic object
IA	internal argument	UG	Universal Grammar

Acknowledgements

I am very happy to contribute to this volume in honor of Ian Roberts, a key figure in the field of Generative Grammar. I had the opportunity to work with Ian back in 2008, when he supervised a British Academy visiting fellowship I was awarded with, right after I became a doctor. I remember that experience (with long appointments at Ian's office) as a very important one in my career and in my personal growth too.

A previous version of this paper was presented at the University of Michigan, in a talk organized within the UMich linguistics colloquium series on 10 March 2017. I would like to thank the audience of that talk for questions and comments. For discussing these matters with me, I am also indebted to Noam Chomsky, Sam Epstein, Hisa Kitahara, Dennis Ott, and Daniel Seely. This research has been partially supported by grants from the Ministerio de Economía y Competitividad (FFI2017-87140-C4-1-P), the Generalitat de Catalunya (2017SGR634), and the Institució Catalana de Recerca i Estudis Avançats (ICREA Acadèmia 2015). Usual disclaimers apply.

References

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its evolution. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Chomsky, Noam. 1956. Three models for the description of language. *IRE Transactions on Information Theory* 2. 113–124.

Chomsky, Noam. 1975 [1955]. *The logical structure of linguistic theory*. New York: Plenum.

Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2004. Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and beyond* (The cartography of syntactic structures 3), 104–131. Oxford: Oxford University Press.

Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36(1). 1–22. DOI: 10.1162/0024389052993655.

Chomsky, Noam. 2007. Approaching UG from below. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax–semantics*, 1–30. Berlin: De Gruyter.

Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.

Chomsky, Noam. 2010. Some simple evo-devo theses: How true might they be for language? Alice V. and David H. Morris Symposium on Language and Communication, Stony Brook University, New York, USA.

Chomsky, Noam. 2012. Foreword. In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 1–7. Berlin: De Gruyter.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.

Chomsky, Noam. 2014. Minimal recursion: Exploring the prospects. In Thomas Roeper & Margaret Speas (eds.), *Recursion: Complexity in cognition*, 1–15. Dordrecht: Springer.

Chomsky, Noam. 2015. Problems of projection: Extensions. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Studies in honour of Adriana Belletti*, 1–16. Amsterdam: John Benjamins. DOI: 10.1075/la.223.01cho.

Chomsky, Noam. 2016. A dialogue with Noam Chomsky. Barcelona, 6 November 2016.

Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions and challenges. *Catalan Journal of Linguistics*. DOI: 10.5565/rev/catjl.288.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2012. Structure building that can't be. In Myriam Uribe-Etxebarria & Vidal Valmala (eds.), *Ways of structure building*, 366–384. Oxford: Oxford University Press.

Epstein, Samuel David, Hisatsugu Kitahara & T. Daniel Seely. 2016. Phase cancellation by external pair-merge of heads. *The Linguistic Review* 33. 87–102.

Epstein, Samuel David & T. Daniel Seely. 2002. Rule applications as cycles in a level-free syntax. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 65–89.

Epstein, Samuel David & Jae-Young Shim. 2015. Two notes on possible approaches to the unification of theta relations. *Linguistic Analysis* 40(1–2). 3–20.

Fernández-Serrano, Irene. 2016. Parameters of hyper-agreement in spanish dialects. Talk at Workshop on Formal Approaches to Romance Microvariation, University of Bucharest, Bucharest, 24–26 November 2016.

Gallego, Ángel J. (ed.). 2012. *Phases: Developing the framework*. Berlin: De Gruyter.

Gallego, Ángel J. 2014. Deriving feature inheritance from the copy theory of movement. *The Linguistic Review* 31(1). 41–71.

Gallego, Ángel J. 2017. Remark on the EPP in labeling theory: Evidence from romance. *Syntax* 20(4). 384–399. DOI: 10.1111/synt.12139.

Lasnik, Howard & Juan Uriagereka. 2005. *A course in minimalist syntax: Foundations and prospects*. Malden, MA: Wiley-Blackwell.
Narita, Hiroki. 2014. *Endocentric structuring of projection-free syntax*. Amsterdam: John Benjamins.
Obata, Miki. 2010. *Root, successive-cyclic and feature-splitting internal merge: Implications for feature-inheritance and transfer*. University of Michigan. (Doctoral dissertation).
Ott, Dennis. 2011. A note on free relative clauses in the theory of phases. *Linguistic Inquiry* 42(1). 183–192.
Richards, Marc. 2007. On feature inheritance: An argument from the phase impenetrability condition. *Linguistic Inquiry* 38(3). 563–572.
Richards, Marc. 2012. On feature inheritance, defective phases, and the movement–morphology connection. In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 195–232. Berlin: De Gruyter.
Richards, Norvin. 1997. *What moves where in which language*. MIT. (Doctoral dissertation).
Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
Uriagereka, Juan. 1999a. Comments on "Derivation by phase". Ms., University of Maryland.
Uriagereka, Juan. 1999b. Multiple spell-out. In Norbert Hornstein & Samuel David Epstein (eds.), *Working minimalism*, 251–282. Cambridge, MA: MIT Press.
Uriagereka, Juan. 2008. *Syntactic anchors: On semantic structuring*. Cambridge: Cambridge University Press.
van Urk, Coppe. 2016. On the distribution of reflexes of successive cyclicity. Talk at BCGL 9, Brussels (Belgium).

Chapter 12

On the coordinate structure constraint and the adjunct condition

Željko Bošković
University of Connecticut

The paper argues for a unification of the ban on extraction out of conjuncts and the ban on extraction out of adjuncts based on the semantics of traditional adjunction modification on which such modification actually involves coordination, with ConjP present in the syntax of traditional adjunct modification. It is shown that there are a number of similarities in the islandhood of conjuncts and the islandhood of adjuncts. Thus, extraction out of conjuncts and extraction out of adjuncts are shown to be exceptionally possible in exactly the same environments, which can be captured if the two involve the same syntactic configuration. The proposed analysis is also shown to capture in a principled way a number of differences in the strength of the violation with extraction out of conjuncts and adjuncts in various languages/contexts, the emphasis regarding the former being on Galician, English, Japanese, and Serbo-Croatian.

1 Introduction

The goal of this paper is to explore the possibility of a unification of two rather ill-understood islands, namely the coordinate structure constraint (CSC) and the adjunct condition (AC). The CSC is standardly assumed to have two parts, given in (1) and (2) below. However, recent research has shown that the two parts of the traditional CSC need to be separated, since there are languages which are sensitive to only one of the constraints in (1–2). Oda (2017) in fact explicitly argues for their separation, providing strong arguments to this effect based on a number of languages. Thus, he notes that Japanese observes (1), but not (2), allowing extraction of conjuncts but not extraction out of conjuncts. The same holds

Željko Bošković. 2020. On the coordinate structure constraint and the adjunct condition. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 227–258. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280649

for Serbo-Croatian (SC), as discussed in Stjepanović (2014) (see Oda 2017 for a list of languages that obey (1) but not (2)). In light of their arguments, I will also separate the two parts of the traditional CSC,[1] focusing on (1) (though I will also make some remarks regarding (2) below). As a result, for ease of exposition I will use the term CSC to refer only to (1). (Where it is necessary to make a distinction between (1) and (2) I will use the terms CSC-1 and CSC-2 respectively.)

(1) The coordinate structure constraint – extraction out of conjuncts (CSC-1)
Extraction out of conjuncts is disallowed.

(2) The coordinate structure constraint – extraction of conjuncts (CSC-2)
Extraction of conjuncts is disallowed.

Turning to adjuncts, the traditional ban on extraction out of adjuncts is given in (3).

(3) The adjunct condition (AC)
Extraction out of adjuncts is disallowed.

The paper will explore the possibility of a unification of (1) and (3), which are illustrated by (4) and (5) respectively.[2]

(4) * What$_i$ did you see [a picture of t_i] and a painting of Storrs?

(5) ?* What$_i$ did you fall asleep [after John had fixed t_i]?

Before getting into the issue of islandhood of conjuncts and adjuncts, a brief note is in order regarding extraction of conjuncts and adjuncts. It is standardly assumed that conjuncts and adjuncts differ in this respect, conjuncts being unmovable and adjuncts movable. It is actually not clear that this is indeed the case. Thus, as noted above, many languages allow extraction of conjuncts. Furthermore, a number of authors have argued that what looks like adjunct extraction actually involves base-generation of adjuncts in their surface position (e.g. Uriagereka 1988; Law 1993; Stepanov 2001b). The standard assumptions in this respect are thus incorrect, at least with respect to conjuncts. At any rate, as noted above, the goal of this paper is not to examine extraction of conjuncts and adjuncts, but islandhood of conjuncts and adjuncts themselves (i.e. extraction out of conjuncts and adjuncts), though some remarks regarding extraction of conjuncts

[1] On separating the two parts of the CSC, see also Grosu (1973) and Postal (1998).

[2] The slight difference in the grammaticality status of (4) and (5) will be accounted for under the unified analysis proposed below.

and adjuncts will be made below from the perspective of a unified analysis of (1) and (3) (more precisely, it will be shown that (2) is not an impediment to such an analysis).

The starting point in the discussion will be the semantics for adjuncts given in Higginbotham (1985). Higginbotham argues that traditional adjunction modification (henceforth traditional adjuncts) actually involves coordination semantically.[3] For example, the rough semantics of (6a) is something like (6b), which can be paraphrased as *There is an event which is walking by John and it is slow.*

(6) a. John walked slowly.
b. $\exists e$[Walk(John, e) and Slow(e)]

Takahashi (1994) made an important observation that under Higginbotham's semantics of adjuncts, where adjuncts essentially involve coordination, it may be possible to unify the ban on extraction out of conjuncts and the ban on extraction out of adjuncts by reducing the latter to the former.[4] Under Higginbotham's semantics, where adjuncts are in fact conjuncts, extraction out of an adjunct does involve extraction out of a conjunct, which makes the unification plausible and appealing. The unification, however, raises an issue. In Takahashi's analysis, while conjuncts and adjuncts are treated in the same way semantically (following Higginbotham), they are treated very differently syntactically, since Takahashi follows standard assumptions in the syntactic literature where coordination involves the presence of a conjunction phrase (ConjP), while adjuncts involve adjunction, with no ConjP present. Thus, the direct object in (4) is a ConjP, with the conjuncts located in the Spec and the complement position of ConjP ((7); the issue of where exactly the conjuncts are located within ConjP is debated in the literature (see e.g. Munn 1993; Progovac 1999), the details of their placement will not matter for our purposes). On the other hand, there is no ConjP in (5). Semantically, the VP and the traditional adjunct are conjoined here. However, this is not reflected in the structure, since Takahashi assumes, following standard assumptions, that the adjunct is adjoined to VP, as in (8).

[3]There is a long line of research in this tradition, see e.g. Davidson (1967); Parsons (1980; 1990); Dowty (1989); Takahashi (1994); Progovac (1998; 1999); Hunter (2011). I refer to Higginbotham (1985) as the representative of this line of research because Takahashi (1994) bases his account of the adjunct condition on it, as discussed below (following Takahashi, I also generalize this approach to adjunct modification in general).

[4]It is worth noting here that Ross (1974) suggested a unification of the CSC with the complex NP constraint (clausal complements of nouns are also sometimes treated as adjuncts, see e.g. Stowell 1981; Takahashi 1994).

(7) * Who$_i$ did you see [$_{ConjP}$ [a picture of t_i] and [a painting of Storrs]]?

(8) ?* What$_i$ did you [$_{VP}$ [$_{VP}$ fall asleep] [after John had fixed t_i]]?

A serious issue then arises: locality of movement is standardly assumed to be a syntactic effect. However, under the above analysis, conjuncts and adjuncts are unified only semantically, they are not unified syntactically in that they involve very different syntactic configurations. It is then not clear that Higginbotham's conjunction semantics of adjuncts can help us here.

While this paper will also take the conjunct semantics of adjuncts seriously, taking it in fact as the point of departure, it will also take seriously the issue of the syntax-semantics mapping here. An obvious question arises in this respect: What would be the syntax that would most straightforwardly correspond to the conjunct semantics of adjuncts? The answer is quite obvious in fact. It is a syntax that involves a ConjP, where e.g. VP and the adjunct in (6) are conjoined. The only difference with true coordination would then be that the conjunction head is phonologically null.[5]

This paper will then take the conjunct semantics of adjuncts seriously, assuming that it is also reflected in the syntax. From this perspective, it is easy to see how (1) and (3) can be unified. Since they involve the same configuration, whatever rules out extraction out of conjuncts will also rule out extraction out of adjuncts.[6]

An important remark is, however, in order here. It seems fair to say that the CSC and the adjunct condition (AC) are the least understood of the traditional

[5]This is in fact what Progovac (1998; 1999) argues for. Thus, Progovac (1998) adopts the structure in (i), where VP is the Spec of ConjP and the adverbial is a complement of a null conjunction (the structure is slightly richer in Progovac 1999).

(i) [$_{ConjP}$ VP [$_{Conj'}$ Conj AdvP]]

In this respect, Progovac (1998; 1999) is an important predecessor of the current work.

It should also be noted that the discussion in this paper raises an issue of whether phrases are ever generated as adjuncts (in the traditional understanding of the term). While the discussion in this paper falls in line with attempts to abandon adjunction as a distinct structure-building mechanism, showing that adjunction can indeed be eliminated goes beyond the scope of this paper.

[6]There is an important issue that arises here. Under the analysis outlined above, not just the adjunct, but also the VP is a conjunct in constructions that involve traditional VP-adjunction. It appears that extraction out of the VP should then also be ruled out here. This is a serious issue that any unification of the CSC and the adjunct condition based on Higginbotham's semantics of adjuncts needs to address. I will provide an account of this issue in §4 below (see Takahashi 1994 for an alternative account which is however based on the assumption that conjuncts and adjuncts have a different syntax).

islands. The suggestion made above reduces two mysteries to one. Resolving this mystery, which would involve providing an actual account of the CSC, however, goes beyond the scope of this paper. Any attempt to do that would involve a detailed discussion of the structure of coordination, as well as the theories of the locality of movement, which is currently based on the theory of phases. A number of issues would arise in this respect: the precise definition of phases, the precise statement of the phase impenetrability condition (PIC) and the notion of *edge*, the issue of the generalized extended projection principle (EPP) effect as it applies to successive-cyclic movement, the theory of labeling, which has been argued to interact with the theory of phases in the locality of movement effects (see Bošković 2015; 2018), etc; the list certainly does not end here. Addressing all of this would go way beyond the scope of this paper.[7] The scope of the paper is more modest: to point out a number of similarities between extraction out of conjuncts and extraction out of adjuncts which can be taken to justify unifying the two. Higginbotham's semantics of adjuncts, when taken seriously from the syntactic point of view, provides a basis for such a unification since the two then have essentially the same structure. Determining the precise source of islandhood of that structure is beyond the scope of this paper (as a result, a number of phenomena noted below will only be discussed at a descriptive level). I will therefore simply use the term islandhood informally below. In several places, the discussion will become more detailed structurally and theoretically when it comes to islandhood – in fact, the paper will provide a principled account of a number of differences in the strength of the violation with extraction out of various conjuncts and adjuncts (as well as the voiding of their islandhood in certain cases); however, the exact reason for the islandhood of conjuncts will not be provided below. In this respect, the paper can be considered to be programmatic, providing a foundation for future work that will account for the islandhood of the syntactic configuration under consideration here (see Bošković 2020).

Having laid down the necessary background, the general line of argumentation, and the limits of the current work, I now turn to making a case for unifying (1) and (3). In that vein, in §§2 and 3 I note a number of similarities between the CSC and the adjunct condition. §4 discusses and resolves some potential impediments to the unification of the islandhood of conjuncts and adjuncts. §5 discusses extraction of conjuncts and adjuncts. §6 concludes the paper.

[7] See, however, Bošković (2017; 2020).

2 The stubbornness of the CSC and the AC

As discussed above, a unification of the traditional coordination and the traditional adjunction has plausible semantic grounds, which can be taken to be reflected in the syntax. From that perspective, it is not surprising that the traditional coordination and the traditional adjunction share some syntactic properties, in particular islandhood. The unification reduces two islands to one, which is already conceptually appealing, especially in light of the fact that we are dealing here with a rather mysterious issue. (Admittedly, we still have a mystery, but reducing two mysteries to one does leave us in a less mysterious state).

One point that has generally been overlooked in the literature on islandhood is worth emphasizing here. For pretty much all islands, it has been noted that there are languages that do not obey them. Thus, there are languages that do not obey the subject condition (e.g. Japanese; see Stepanov 2001a for a more exhaustive list), there are languages that do not obey the *wh*-island constraint (e.g. Swedish, see Engdahl 1986), there are languages that do not obey the complex NP constraint (e.g. Bantu languages, see Bošković 2015). The CSC and the AC stand out rather prominently in this respect. I am not aware of any language that does not obey the CSC and the AC.[8] From the current perspective, that the CSC and the AC behave in the same way in this respect is not surprising: we are after all dealing with one and the same constraint here – that the two behave in the same way in the relevant respect is then expected.

3 Some exceptions to the CSC and the AC

3.1 A semantically-based exception

It is well-known that there are exceptions to both the AC and the CSC (see Truswell 2011 and references therein for the former and Postal 1998 and references therein for the latter). Interestingly, some of these exceptions are rather similar in nature. Thus, extraction from an adjunct is possible in some cases where there is a contingent relationship between the relevant events. Importantly, the same kind of exception is found with the CSC. The former is illustrated by (9) and the latter by (10).

(9) a. What$_i$ did you come around [to work on t$_i$]?
b. What$_i$ did Christ die [to save us from t$_i$]? (Truswell 2011: 131)

[8] As is well-known and as we will see below, there are particular coordinations and adjunctions that allow extraction (in fact likely universally). What I am referring to here is different, namely I am not aware of any language that would allow extraction out of all coordinations and all adjuncts, where conjuncts and adjuncts simply would not be islands at all.

(10) a. This is the drug which$_i$ athletes [take t_i] and become quite strong.
b. the stuff which$_i$ Arthur sneaked in and [stole t_i] (Postal 1998: 53)

There are no good explanations for why under the semantic condition noted above the adjunct condition effect and the CSC effect are voided, and I will not provide one in this work. What is important for our purposes is that the two behave in the same way here. A unified approach to the two in this respect has not been attempted before even at a descriptive level; what complicates the situation even further when it comes to providing an actual account is that only argument (both DP and PP) extraction is allowed in the exceptional context in question, non-argument extraction is still unacceptable, as illustrated below.

(11) * How$_i$ did you come around [to work on that car t_i]?

(12) * How$_i$ should athletes [take that drug t_i] and become strong?

This, however, further confirms that the CSC and the AC behave in the same way here, which can be interpreted as calling for a unified analysis of the two. The suggestion made here achieves this trivially, by treating the CSC and the AC as one and the same phenomenon.

3.2 Across-the-board movement and parasitic gaps

There is another well-known exception to the CSC which is not semantically based (i.e. it is not semantically restricted like the one noted directly above). The exception, noted already in Ross (1967), concerns across-the-board (ATB) movement. As is well-known, an unacceptable extraction out of a conjunct can be made acceptable if the extraction takes place out of each conjunct in the coordination.

(13) Who did you see enemies of and friends of?

(14) cf. *Who did you see John and enemies of?

There is an obvious counterpart of this with the AC, which is the traditional parasitic gap construction (see also Haïk 1985; Huybregts & van Riemsdijk 1985; Williams 1990; Franks 1993; Progovac 1998; Nunes 2004).

(15) What did you file without reading?

(16) cf. *What did you file the book without reading?

From the current perspective, (15–16) can be looked at on a par with (13–14). Just like the unacceptable case of extraction out of a conjunct in (14) becomes acceptable if extraction takes place out of both conjuncts, as in (13), so does the unacceptable case of extraction out of a conjunct in (16) (the traditional adjunct being a conjunct under the current analysis) become acceptable if extraction takes place out of both conjuncts, as in (15) (VP being a conjunct under the current analysis; see below for extraction out of the VP here).

There have in fact been many attempts to unify the ATB and the parasitic gap construction (see the references cited above); the current perspective can be taken to provide motivation for those attempts (Takahashi 1994 in fact also argues for a unification of the two from the perspective of Higginbotham's semantic treatment of adjuncts (recall, however, that Takahashi treats conjuncts and adjuncts differently syntactically).

3.3 The edge exception

Bošković (2018) notes another exception to the AC. Bošković (2018) shows that the AC effect is quite generally voided for elements that are base-generated at the adjunct edge, also providing an account of this state of affairs where the problem with extraction out of adjuncts arises with movement to the adjunct edge (which is required by the PIC); elements that are base-generated at the adjunct edge can then extract. The details of the account are not important for our purposes; what is important is that elements base-generated at the edge of an adjunct can extract out of it.

One illustration of this effect is provided by the different behavior of agreeing possessors and adnominal genitive complements with respect to extraction out of adjuncts in Serbo-Croatian (SC). Consider first the former. Agreeing possessors in SC have been argued to be base-generated at the edge of the TNP.[9] As one argument to that effect, consider the following binding contrast between English and SC, noted in Despić (2011; 2013).

(17) a. His$_i$ latest movie really disappointed Kusturica$_i$.
b. Kusturica$_i$'s latest movie really disappointed him$_i$.
c. Serbo-Croatian (Despić 2011: 31; 2013: 245)
*Kusturicin$_i$ najnoviji film ga$_i$ je zaista razočarao.
Kusturica's latest movie him is really disappointed
d. *Njegov$_i$ najnoviji film je zaista razočarao Kusturicu$_i$.
his latest movie is really disappointed Kusturica

[9]The term TNP is used neutrally, for whatever the categorial status of the relevant element is.

Under the assumption that traditional Specs c-command out of the phrase where they are located, Kayne (1994) takes the acceptability of (17a,b) to indicate that English possessors are not located in SpecDP, but in the Spec of a lower phrase, SpecPossP, with the DP confining the c-command domain of the possessor. Despić (2011; 2013) observes that in SC, a language without articles which has been argued by a number of authors to lack DP (e.g. Corver 1992; Zlatić 1997; Trenkić 2004; Bošković 2005; 2012; 2014; Marelj 2011; Despić 2011; 2013; Runić 2014a,b; Takahashi 2012; Talić 2014; 2015), possessors do c-command out, as indicated by the binding violations in (17c,d) (condition B is at issue in 17c and condition C in 17d), which contrast with English (17a,b). Despić takes the contrast in question as indicating that DP is missing in SC, with the possessor located in the highest projection of the traditional NP.

Turning now to adjuncts, SC is rather productive regarding the possibility of traditional NPs (TNPs) functioning as adjuncts. One such case is given below, where an instrumental nominal functions as an adjunct (see Bošković 2018 for discussion of such adjuncts).

(18) Serbo-Croatian
Trčao je šumom.
run is forest.INS
'He ran through a/the forest.'

That the instrumental nominal in (18) is indeed an adjunct is confirmed by extraction. First, its extraction out of islands yields an ECP-strength, not a subjacency-strength violation (compare 19a,b).

(19) Serbo-Croatian

a. * Šumom_i se pitaš [kad je trčao t_i].
forest.INS REFL wonder when is run
'You wonder when he ran through a/the forest.'

b. ?? Šumu_i se pitaš [kad je posjekao t_i].
forest.ACC REFL wonder when is cut-down
'You wonder when he cut down a/the forest.'

In addition to agreeing possessors, which roughly correspond to English *'s*-genitives, nominal arguments in SC can be expressed through adnominal genitive, which roughly corresponds to English *of*-genitives; the element bearing adnominal genitive occurs in the complement position of the noun. Returning now

to the instrumental adjunct under discussion, notice that while extraction of genitive complements of nouns is in general somewhat degraded in SC, (20a), which involves extraction out of the nominal under consideration, is clearly worse than (20b), which involves extraction out of an object. This confirms the adjunct status of the instrumental TNP (20a is worse than 20b because it involves extraction out of an adjunct).

(20) Serbo-Croatian

a. * Moga djeda$_i$ je trčao [šumom t_i].
my.GEN grandfather.GEN is run forest.INSTR
'He ran through the forest of my grandfather.'

b. ?? Moga djeda$_i$ je volio [šumu t_i].
my.GEN grandfather.GEN is loved forest.ACC
'He loved the forest of my grandfather.'

As noted above, Bošković (2018) shows that in contrast to elements that are not base-generated at an adjunct edge, elements that are base-generated at an adjunct edge can be moved out of adjuncts. The adnominal genitive 'my grandfather' in (20a) is base-generated in the N-complement position. Recall, however, that an agreeing possessor that precedes the nominal is generated at the TNP edge. Importantly, such possessors can move out of the adjunct under consideration.

(21) Serbo-Croatian
Ivanovom$_i$ je on trčao [t_i šumom].
Ivan's.INS is he run forest.INS
'He ran through Ivan's forest.'

Bošković (2018) provides a number of additional cases which also show that elements that are base-generated at an adjunct edge can move out of adjuncts, in contrast to those that are not generated at an adjunct edge.[10]

What is important for our purposes is that the CSC behaves just like the AC in this respect. Recall that an agreeing possessor can extract out of a TNP adjunct,

[10] One such case is given in (i) (see Bošković 2018 for an account why (i) is unacceptable in English).

(i) Izuzetno$_i$ se on [t_i loše] ponašao?
extremely is he badly behaved
'He behaved extremely badly.'

while an adnominal genitive cannot. Coordinations behave in exactly the same way: an agreeing possessor can extract out of a conjunct (22), but an adnominal genitive cannot (23).[11]

(22) Serbo-Croatian
Markovog$_i$ je on [t$_i$ prijatelja] i [Ivanovu sestru] vidio.
Marko's.ACC is he friend.ACC and Ivan's.ACC sister.ACC seen
'He saw Marko's friend and Ivan's sister.'

(23) Serbo-Croatian
*Fizike$_i$ je on [studenta t$_i$] i [Ivana] vidio.
physics.GEN is he student.ACC and Ivan.ACC seen
'He saw a student of physics and Ivan.'

What is important for our purposes is that both traditional adjuncts and traditional conjuncts exceptionally allow extraction of elements that are base-generated at their edge.

To sum up the discussion in this section, we have seen that in a number of environments extraction is exceptionally possible out of conjuncts and adjuncts. Significantly, the enviroments where extraction is exceptionally possible out of conjuncts and adjuncts are the same – all the contexts discussed in this section exceptionally allow extraction out of both conjuncts and adjuncts (see below for an additional case). That the two behave in the same way in this respect then provides an argument that they should be unified, which is straightforwardly accomplished if they involve the same syntactic configuration.

4 Some differences between the CSC and the AC and rescue by PF deletion

Above, I have discussed a number of similarities between CSC effects and AC effects which can be captured under the analysis on which traditional adjunction actually involves coordination, which is motivated by Higginbotham's semantics of adjunction. There are, however, also some differences between the two, which

[11]Left-branch extractions in SC are best when the remnant precedes the verb, but the relevant contrast is also there when the coordination follows the verb. Notice that there is an interfering factor when such extraction is attempted out of the second conjunct. As noted in Stjepanović (2014) and discussed below, *i* 'and' is a proclitic, which procliticizes to the element following it. A problem then arises if the element following it is a trace.

will be discussed in this section, starting with an obvious difference.[12] Consider (24–25), which are intended to represent a case of traditional coordination (24) and a case of traditional adjunction (25), which is also treated as involving coordination under the current analysis.

(24) DP & DP

(25) VP & Adjunct

The conjuncts in the traditional coordination in (24) are symmetric regarding islandhood in that extraction is banned out of each conjunct (putting aside the ATB case).

(26) a. * Who$_i$ did you see [a friend of t$_i$] and John?
b. * Who$_i$ did you see John and [a friend of t$_i$]?

However, this is not the case with (25), where extraction is not banned out of the first conjunct, i.e. VP.

(27) What$_i$ did you [buy t$_i$] slowly?

A question then arises under the current analysis regarding the source of this difference. In particular, what raises the issue here is the grammaticality of (27), which appears to be unexpected.

As noted above, providing an account of the unacceptability of extraction out of conjuncts goes beyond the scope of this paper. I simply assume here that conjuncts are islands (as explicitly also argued in Oda 2017). The islandhood of conjuncts is apparently voided for the VP conjunct in (27). The question is why. There is actually a rather straightforward answer to this question.

Bošković (2011; 2013b) discusses a variety of islands from a number of languages and observes that movement of the head of an island voids islandhood (for additional arguments to that effect, see Bošković 2015). Based on this, Bošković establishes the generalization in (28).

(28) Traces do not head islands.

[12] A reviewer notes that coordination and traditional adjunction differ regarding gapping, compare *John ate an apple and Mary a pear* with **John ate an apple after Mary a pear*. The difference can be accounted for under Johnson's (2009) analysis of gapping (gapping is actually quite generally disallowed in embedded clauses, even with coordination).

Bošković (2013b) provides a number of arguments for (28). As an illustration, consider the saving effect of article incorporation on islandhood in Galician, also discussed in Uriagereka (1988; 1996). Galician has a rather interesting phenomenon of D-to-V incorporation, which quite generally voids islandhood of the DP from which the incorporation takes place (see Uriagereka 1988; 1996; Bošković 2013b). Thus, Galician disallows movement from definite DPs, as in (29). However, the violation is voided when D incorporates into the verb, as shown by (30).[13] Further confirmation of the islandhood-voiding effect of article incorporation is provided by (31). Extraction from adjuncts is banned in Galician, as in (31). However, the ban is voided under D-incorporation, as in (32) (the same holds for the subject condition effect, which is also voided under article incorporation).

(29) Galician (Uriagereka 1988: 81)
* e de quén$_i$ viches [$_{DP}$ o [$_{NP}$ retrato t$_i$]]?
and of who saw(you) the portrait

(30) Galician (Uriagereka 1988: 81)
e de quén$_j$ viche-lo$_i$ [$_{DP}$ [$_{D'}$ t$_i$ [$_{NP}$ retrato t$_j$]]]?
and of whom saw(you)-the portrait
'So, who have you seen the portrait of?'

(31) Galician (Bošković 2016: 58)
?? de que semana$_j$ traballastedes [$_{DP}$ o Luns t$_j$]
of which week worked(you) the Monday
'Of which week did you guys work the Monday?'

(32) Galician (Bošković 2016: 58)
de que semana$_j$ traballastede-lo$_i$ [$_{DP}$ [$_{D'}$ t$_i$ Luns t$_j$]]
of which week worked(you)-the Monday

These cases illustrate the generalization in (28). The islandhood of the DPs from (29) and (31) is voided in (30) and (32), where the relevant DPs are headed by a trace, due to the movement of the head of the DP in question. Bošković (2013b; 2015) provides a number of other cases from a wide range of languages that illustrate the same effect (thus, Bošković 2013b shows that, among other things, Baker's (1988) government transparency corollary effects are also subsumed under (28); i.e. they also involve islands that are headed by a trace.) Under (28), if the head of an island α undergoes movement, the islandhood of α is voided, making movement out of α possible.

[13] As discussed in Uriagereka (1988), when the article incorporates the final *s* of the verb is truncated.

Bošković (2011; 2013b) also provides an account of the effect in question, which unifies it with the rescuing effect that ellipsis has on islandhood, noted by Ross (1969) and illustrated by (33).[14]

(33) a. * She kissed a man who bit one of my friends, but Tom does not realize [which one of my friends]$_i$ she kissed [**a man who bit t$_i$**].

b. She kissed a man who bit one of my friends, but Tom does not realize [which one of my friends]$_i$ ~~she kissed [a man who bit t$_i$]~~.

(Ross 1969: 276)

The effect from (33) is standardly treated in terms of rescue by PF deletion (Chomsky 1972; Merchant 2001; Lasnik 2001; Fox & Lasnik 2003; Hornstein et al. 2003; Boeckx & Lasnik 2006; Bošković 2011 among others): a * is assigned to an island when movement crosses it. If the * remains in the final PF representation, a violation incurs. If a later operation like ellipsis deletes the category that contains the *-marked element, the derivation is rescued. Under the standard analysis, then, when *wh*-movement crosses the island in (33) the island is *-marked in both (33a) and (33b). Since the *-marked element is deleted in (33b) the islandhood effect disappears in this example.

Bošković (2011; 2013b) also provides a rescue-by-PF deletion account of the generalization in (28), unifying (28) with the rescuing effect of ellipsis on islandhood. Bošković argues that what is *-marked is not the whole island, but the head of the island. This means that in e.g. (29), what is *-marked is the head of the object DP. The reason for the rescuing effect of head movement in (30) is that the *-marked element in the head position of the object DP is actually a copy that is deleted under copy deletion in PF. The offending *-marked element is thus deleted in PF in (30), just as it is in (33). The analysis quite generally captures the generalization in (28).[15] (Bošković 2011 also extends the analysis to the generalization that traces do not count as interveners (Chomsky 1995). In the relevant cases, the *-marked intervener is also removed under PF copy deletion, see the discussion below).

[14] See, however, Abels (2011); Barros et al. (2014).

[15] The analysis predicts that head movement is not sensitive to (non-relativized minimality) islands, more precisely, that the head of an island can move out of the island since the locality violation will be rescued by deleting the copy of the moved head (the prediction holds only for the head of the island and does not hold for relativized minimality – i.e head-movement constraint – violations; see Bošković 2013b). Bošković (2013b) provides a number of cases from a variety of languages that this is indeed the case (in fact, Galician article incorporation – cf. (32) –, which is also acceptable without *wh*-movement, is one such case; see also Bošković 2013b on noun incorporation in Kinyarwanda, Chichewa, and Southern Tiwa).

At any rate, what is important for our purposes is that head movement voids islandhood: if the head of an island undergoes movement, the islandhood effect disappears, making movement out of the island possible.

Returning to the potentially problematic case in (27), we now have a straightforward explanation why movement out of the VP, which is a conjunct hence an island under the current analysis, is allowed in this case. The reason is V-to-v movement.[16] Being a conjunct, the VP (i.e. the bracketed element) in (27) is an island. However, V-to-v movement, i.e. movement of the head of the VP, voids the islandhood of the VP, allowing movement out of this VP, as in (27). The grammaticality of (27) is then just another instance of the general rescuing effect of head movement on islandhood, given in (28). The potential obstacle to the unification of the CSC and the AC that was raised by (27) is thus rather straightforwardly resolved; the reason for the grammaticality of (27) is an independent and more general effect regarding locality of movement.

The analysis does not only remove a potential problem for the unification of the CSC and the AC raised by (27) but it also makes a prediction. Consider again (24–25). Just like in (25) movement of the head of the VP conjunct makes movement out of the VP possible so should movement of the head of the corresponding conjunct in (24) make movement out of this conjunct possible. The prediction can in fact be tested with respect to Galician. The issue here is whether article incorporation in Galician also improves extraction out of a conjunct. It turns out that it does. Consider (34–35) (the Galician data below are due to Juan Uriagereka, p.c.; *a* in (34–35) is a differential object marker).

(34) Galician
* De quén$_i$ vistedes [o amigo t$_i$] e-mais [a Xan] onte?
of who (you)saw the friend and DOM Xan yesterday
intended: 'You saw [[the friend of who] and [Juan]] yesterday?'

(35) Galician
?? De quén$_i$ vistede-lo$_j$ [t$_j$ amigo t$_i$] e-mais [a Xan] onte?
of who (you)saw-the friend and DOM Xan yesterday

(34) shows that extraction out of a conjunct is not possible in Galician, i.e. conjuncts are islands. Importantly, (35), which involves article incorporation from the conjunct from which *wh*-movement takes place, is clearly better than (34),

[16] There are various proposals in the literature regarding the exact identity of the relevant head and the height of V-movement (e.g. we could be dealing here with a vP conjunct, with the verb moving to VoiceP above vP, see Collins 2005); I simply use v for ease of exposition.

which does not involve article incorporation. Article incorporation thus also improves extraction out of conjuncts.

Putting for the moment the residual awkwardness of (35) aside, and focusing on the fact that (35) is better than (34), the current analysis unifies the grammaticality of (27) with the improvements that article incorporation causes for *wh*-movement in (31–32) and (34–35). All the relevant cases involve extraction out of a conjunct where the head of the conjunct undergoes movement.

Consider now why, in contrast to (27) and (32), (35) is still degraded (although better than (34), which is what is crucial here for our purposes).[17] Oda (2017) captures the two parts of the CSC, i.e. (1–2), by proposing that both individual conjuncts and ConjP are islands. What this entails for our purposes is that with extraction out of a conjunct, what is *-marked is the head of the conjunct itself, as well as the head of ConjP (given that what is *-marked is the head of an island). In (34), both *-marked heads survive into PF, hence the strong unacceptability of the construction. On the other hand, in (35), the *-marked head of the conjunct is removed in PF through copy-deletion. However, the *-marked head of ConjP is still present in PF. I suggest that this is the reason for the residual awkwardness of (35). Article-incorporation voids the islandhood of the conjunct itself, by turning its head into a trace (i.e. a copy that is deleted in PF). However, it does not affect the islandhood of ConjP. The analysis thus captures the contrast between (34) and (35), as well as the fact that (35) itself is still degraded.

What about (27) and (32), which involve traditional adjunction? I suggest that what is important here is that the ConjP head in these examples is phonologically null. In this respect, the head of ConjP in (27) and (32) in fact does not differ from the head of the first conjunct in (27) and the second conjunct in (32) – in all these cases the relevant head is phonologically null. Now, it is standardly assumed that intervening heads block head movement (see e.g. Roberts 2010). There is an additional implicit assumption here: in all the cases that are traditionally given as an illustration of this effect the blocking head is overt. This is in fact reminiscent of another standard assumption, noted briefly above, that traces do not count as interveners.[18] What traces and null heads have in common is that they are both

[17](32) is actually slightly awkward (meriting at most ?). The proposal below will not explain the residual awkwardness of (32), which I leave open here (also putting it aside below), merely noting that there may be a weak intervention effect associated with phrasal movement from the second conjunct crossing the first conjunct, also a phrase (32 is in fact fully acceptable if it involves only head-movement/article incorporation, see Bošković 2013b); in this respect compare also (35) with (39) below and note that (26b) is worse than (26a); for discussion of the effect in question, which I put aside here, see Bošković (2020), who also shows that the effect is selective in that it depends on labeling (so it does not arise in all relevant contexts).

[18]Notice that there is no conflict between the assumption that traces do not count as interveners

phonologically null; this means that null elements do not count as interveners. Bošković (2011) in fact provides a rescue by PF deletion account of the trace case that can be generalized to the null head case. Bošković (2011) argues that with intervention effects, what is *-marked is the intervener itself. With traces, the intervener is deleted in PF, which voids the intervention effect. Another way to look at this is that the locality effect is voided if the *-marked element is not realized (i.e. pronounced) in PF, i.e. a * induces a violation in PF only if it is PF realized, i.e. if it is present on a PF-realized element.[19]

There is independent evidence for the above account of (27), where the reason why (27) does not display the CSC effect, although adjunction is treated as coordination, is that the ConjP head is phonologically null here. Progovac (1998; 1999), who also argues for a unified analysis of coordination and traditional adjunction based on the coordination analysis of the latter, observes that in some cases the ConjP head can in fact be overt with traditional adjunction based on examples like (36). Importantly, extraction out of the VP conjunct is degraded in such cases: (37a,b) are worse than (27). This is exactly what is expected: since the *-marked head of ConjP is phonologically realized in (37a,b), in contrast to (27), examples (37a,b) are degraded, in contrast to (27).

(36) a. Mary read his paper, and quickly.
b. John read the book, and avidly.

(37) a. ?? What did Mary read, and quickly?
b. ?? What did John read, and avidly?

We now have all we need to account for the full paradigm under consideration. In (27) and (32), both the islandhood of the relevant individual conjuncts and the islandhood of ConjP is voided since both the head of the relevant conjuncts and the head of ConjP are phonologically null. On the other hand, in (35), only the head of the conjunct is null, which means that the islandhood of the conjunct, but not the islandhood of ConjP, is voided here. Notice also that (34) is worse than (31), which is also captured under the current analysis. (34) in a sense involves two violations, since the heads of both islands, the relevant conjunct and ConjP,

for extraction and the blocking effect of *wh*-traces on *wanna*-contraction. Under multiple spell-out (see Uriagereka 1999; Epstein 1999; Chomsky 2000; 2001 among many others), it is not a *wh*-trace but the *wh*-phrase itself that blocks *wanna*-contraction (see Bošković 2013a, where it is shown that this kind of approach also captures the traditional claim that NP-traces do not block contraction; traces actually never block contraction, only heads of chains do under a multiple spell-out analysis).

[19]Though see below for a potential alternative.

are phonologically overt. On the other hand, in (31) only the former is phonologically overt: the islandhood of ConjP is voided here since the head of ConjP itself is phonologically null. Furthermore, notice that standard CSC violations like (26a) are worse than traditional adjunction cases with an overt conjunction like (37). This is also expected and can be accounted for on a par with the contrast between (31) and (34): (26a) involves two island violations since both the head of the conjunct island and the head of ConjP are overt while in (37) only the head of ConjP is overt. The proposed analysis thus captures the full paradigm in (26–27, 31–32, 34–35, and 37): it captures the fact that (27) and (32) are better than the rest of this paradigm, the contrast between (34) and (35) as well as the fact that (35) is still degraded, and the fact that (34) is more strongly degraded than (31) and that (26) is more strongly degraded than (37).[20]

What is particularly important for our purposes is that the current analysis unifies the grammaticality of (27) and the improvement that article incorporation causes in (34–35). In both cases we are dealing with extraction out of a conjunct where the head of the conjunct undergoes movement, voiding the islandhood of the conjunct. The grammaticality of (27) then turns out not only not to be a problem for the unified CSC/AC analysis, but it in fact has its counterpart with the traditional CSC, thus providing an argument for the unified analysis. In other words, we are dealing here with another case where movement out of a conjunct is exceptionally allowed, which also extends to traditional adjunction. In fact, the effect holds not only for what under the traditional view would be considered to be the "host" of adjunction, i.e. the VP in (25), but also for the traditional adjunct itself. As shown in (31–32), the islandhood of extraction out of adjuncts is also voided under movement of the adjunct head. I conclude therefore that what appeared here to be a difference between the CSC and the AC is in fact another case where the two behave in the same way, which can be added to the cases discussed in §3: both the CSC and the AC effect are voided under head movement of the head of the conjunct/adjunct.

There is still one missing piece needed to complete the paradigm regarding the rescuing effect of head movement on extraction from conjuncts. Returning to (24–25), we have seen that head movement rescues extraction out of both conjuncts in the traditional adjunction case in (25), i.e. it makes extraction out of both VP and the traditional adjunct possible. Regarding (24), we have seen

[20]One issue that I will put aside here is whether extraction out of all conjuncts can be saved by movement of the conjunct head. What is important for us is that this is in principle possible, hence needs to be allowed. Whether there are factors that constrain the effect in question will be left for future research (see Bošković 2017, where it is argued that the status of a conjunct with respect to phasehood matters here; for relevant discussion see also Bošković 2020).

that head movement of the head of the conjunct makes extraction out of the first conjunct possible. The remaining piece of the puzzle concerns extraction out of the second conjunct in (24). Does head movement of the head of that conjunct make extraction out of it possible? We have confirmed the rescuing effect of head movement on extraction out of a conjunct regarding the first conjunct in (24) with article incorporation in Galician. Does the effect also hold for extraction from the second conjunct? In fact, it does. Conjunction *e mais* in Galician can host article incorporation. Crucially, extraction out of the second conjunct is worse in (38) than in (39), the difference here being that the article head of the second conjunct, from which *wh*-extraction takes place, undergoes incorporation only in (39). (Not surprisingly given the above discussion, while better than (38), (39) is still degraded.)

(38) Galician
* De qué cidade$_i$ vistedes um retrato de Diego e mais [a
of what city (you)saw a portrait of Diego and the
paisaxe t_i]?
landscape

(39) Galician
??? De qué cidade$_i$ vistedes um retrato de Diego e-mai-la$_j$ [t_j
of what city (you)saw a portrait of Diego and-the
paisaxe t_i]?
landscape

I will conclude the discussion in this section with an example which can be analyzed in several ways within the approach argued for here. The example is given in (40).

(40) * What$_i$ did you see [pictures of t_i] and paintings of Storrs?

The conjunct from which extraction takes place in (40) is most often assumed to be a DP, headed by a null D. Given the grammaticality status of (40), here we do want the *-marking on the head of the conjunct to contribute to the ungrammaticality of the example.

There are several possibilities here. One possibility is that the conjunct is actually smaller than DP, with the noun located in (possibly moving to) the head position of the conjunct. Nothing special would then need to be said about such cases.

If the conjunct is a DP, with the noun located lower than D, we could assume that this is actually a D that is deleted in PF, with PF D-deletion either not yet

having taken place at the point when *-marking is checked, or with *-marking interfering with the required D deletion here. However, what may be relevant here is that DP is a phase, in contrast to ConjP (see Bošković 2017 for relevant discussion). In light of this, it is possible that, as suggested above, *-marking on null heads never matters (i.e. it does not induce a PF violation) but that *-marked heads are unable to send their complement to spell-out. The standard assumption is that phasal heads send their complement to spell-out *after* all their uninterpretable features are checked; under the suggestion made here *-marking has a similar effect to uninterpretable features in that it prevents spell-out. As a result, the *-marked null D in (40) would not be able to send its complement to spell-out.[21]

There is another possibility here. Assume a framework like Distributed Morphology, where phonological features are inserted in PF to essentially lexicalize appropriate feature matrices. As argued in Progovac (1998; 1999) and discussed briefly in §6 (see footnote 27), the reason why $Conj^0$ is typically not lexicalized with traditional adjunction is the *avoid overt conjunction principle*, which works in a similar way as Chomsky's (1981) avoid pronoun principle. We can then assume that in the relevant situations (see §6 for why this happens with traditional adjunction), the feature matrix of the conjunction head (or the pronoun in the cases where the avoid pronoun principle is relevant, see Holmberg 2005) is deleted, as a result of which phonological features cannot be inserted. This is not the case with the null D in (40). The feature matrix of this null D simply does not correspond to any phonological features (in contrast to the conjunction head, where, unless the relevant feature matrix is deleted, phonological features would be inserted): there is no deletion of the feature matrix here that would prevent phonological feature insertion. Under this analysis, the difference between the null Conj head in examples like (27) and the null D in examples like (40) with respect to *-marking is treated in the same way as the difference between the article and its trace in Galician examples like (29–30): In all these cases the relevant

[21]I assume that spell-out must take place for each phasal level, which means that we do have a violation here. Notice also that there is still a difference here with the Galician case in (30), where the *-marked element in D is deleted under copy deletion. Under the analysis under consideration, the spell-out for the DP phase in (30) would be triggered only after D-incorporation (with copy deletion appropriately ordered), which is in fact in line with Chomsky's (2001) proposal that the spell-out for phase XP is triggered by a higher phase head. (Note also that, as argued in Bošković 2015, D-incorporation is driven by an uninterpretable feature of D, which means that D anyway could not trigger spell-out before it moves.) It should, however, be noted that under the approach to phases in Bošković (2015), D-incorporation voids the phasehood of the DP from which it takes place, so that the issue of DP-phase spell-out would not even arise in this case.

head is *-marked due to extraction out of a conjunct, conjuncts being islands. The *-marked head is then deleted in (30) (due to copy deletion) and (27) (due to the avoid overt conjunction principle, which works on a par with the avoid pronoun principle). On the other hand, the *-marked head is not deleted in examples like (29) and (40). Notice that under this analysis, *-marking on elements which are not realized (i.e. pronounced) in PF would not actually be ignored.[22]

At any rate, I leave teasing apart the analyses of (40) suggested above for future research and continue to assume below that a * induces a violation in PF only if it is present on a PF realized element.[23]

5 On extraction of conjuncts/adjuncts

As noted at the outset, the discussion in this paper is limited to islandhood of conjuncts and adjuncts, i.e. extraction out of conjuncts/adjuncts; it does not deal with extraction of conjuncts/adjuncts. As discussed in §1, while the CSC was traditionally assumed to hold both for extraction out of conjuncts and for extraction of conjuncts, this view is quite clearly wrong, since there are languages that productively allow extraction of conjuncts but still disallow extraction out of conjuncts. This is the reason why I have put the discussion of extraction of conjuncts, i.e. (2), aside above. In this section, I will, however, make some brief remarks on extraction of conjuncts, i.e. the status of (2), the reason being that the rescue-by-PF deletion mechanism, which I have appealed to above, turns out to be relevant to (2), as was in fact explicitly argued in Stjepanović (2014) and Oda (2017).

Notice first that the CSC is not completely divorced from the AC even when it comes to (2), i.e. extraction of the conjunct/adjunct. Both are in principle possible, but there is a productivity difference here in that extraction of adjuncts is more readily available crosslinguistically than extraction of conjuncts. In this respect, we have the following situation: there are languages like Japanese and SC that in principle allow both extraction of conjuncts and extraction of adjuncts; there are languages like English that allow extraction of adjuncts but not extraction of conjuncts. I am, however, not aware of any languages that would allow extraction of conjuncts but not extraction of adjuncts. In other words, we have a small implicational hierarchy here, where the possibility of extraction of adjuncts entails the possibility of extraction of conjuncts. It turns out that there is a way of

[22]For an argument that it should not be, see Bošković (2011).

[23]The discussion below can be easily adjusted to the last account of (40) suggested above, if it turns out to be the most appropriate one.

making sense of this state of affairs under the rescue-by-PF deletion approach discussed above.

Recall that Oda (2017) argues that both individual conjuncts and ConjP are islands. When it comes to extraction of conjuncts themselves, i.e. (2), what is relevant is the islandhood of ConjP: the island that is crossed when a conjunct is extracted is ConjP. This means that what is *-marked when a conjunct is extracted is the head of ConjP (given that what is *-marked is the head of an island).

Importantly, in languages where extraction of a conjunct is allowed, it has been shown that the ConjP head is a clitic that undergoes movement. In other words, the head of ConjP is a trace. This immediately makes (28) relevant here: the cliticization voids the islandhood of ConjP, making extraction of a conjunct possible. In fact, Oda (2017) and Stjepanović (2014) argue for exactly this account of the exceptional possibility of extraction of conjuncts in Japanese and SC. In both languages the conjunction head is a clitic, which Oda and Stjepanović argue undergoes movement. In Japanese, the conjunction is an enclitic and in SC it is a proclitic. In Japanese (41), the conjunction cliticizes to the first conjunct and is in fact carried along under the movement of the first conjunct, which quite conclusively shows that the conjunction head does not remain in its in situ position.

(41) Japanese (Oda 2017)

a. ? Kyoodai$_i$-to kanojo-wa [t_i Toodai]-ni
Kyoto.University-and she-TOP Tokyo.University-DAT
akogareteiru.
admire
'She admires Kyoto University and Tokyo University.'

b. (?) Nani$_i$-to Taro-ga [t_i mizu]-o katta no?
what-and Taro-NOM water-ACC bought Q?
literally 'What did Taro buy and water?'

In fact, as discussed in Oda (2017), in all languages where extraction of a conjunct is possible the conjunction head is a clitic that undergoes movement.[24]

[24] As discussed in Stjepanović (2014), in SC the conjunction procliticizes to the second conjunct, which makes movement of the first conjunct, as in (i-a), possible. (See Stjepanović 2014 for details of the derivation, which also involves ConjP-internal movement of the second conjunct prior to the procliticization of the conjunction to it. Stjepanović shows that the process in question quite generally applies to SC proclitics; thus, she shows, following Bošković 2013b and Talić 2014, that the proclitic preposition in (i-b) procliticizes to the AP (and is carried along under further movement of the AP, as in (i-c)), with Talić's (2014) prosodic arguments for procliticization in terms of syntactic movement of the preposition in (i-b) extending to the conjunction in (i-a).)

The possibility of conjunct extraction can then be rather straightforwardly accounted for under (28), i.e. in terms of a rescue-by-PF deletion analysis (see Oda 2017; Stjepanović 2014).

As discussed above, with extraction of conjuncts, ConjP functions as an island. This means that what is *-marked when such extraction takes place is the head of ConjP. In Japanese, where the conjunction head undergoes movement, the islandhood effect is voided since the *-marked element is deleted in PF (under copy deletion). The analysis thus unifies acceptable CSC-2 violations like (41) with other acceptable island violations in (30) and (32), all of which are instances of the generalization in (28), which is, as discussed above, unified with the rescuing effect of ellipsis on locality violations, i.e. cases like (33), in terms of the rescue-by-PF deletion mechanism.

Recall now the observation made above regarding the availability of extraction of traditional conjuncts and traditional adjuncts, both of which involve extraction of conjuncts under the current analysis: extraction of traditional adjuncts is much more generally available than extraction of traditional conjuncts. The mechanism of rescue-by-PF deletion provides a straightforward account of why this is the case. The above discussion has indicated that extraction of a traditional conjunct is possible only if the head of ConjP is phonologically null, which we have seen can be captured by the mechanism of rescue-by-PF deletion. Turning to adjunct extraction, under the current analysis adjuncts are conjuncts, with ConjP headed by a null head present in the structure. But this is exactly when extraction of a conjunct is possible even with traditional coordination: when the head of ConjP is phonologically null. True, the reason for this is different (in one case the head is phonologically null as a result of PF copy deletion and in the other case it is null to start with), but that does not matter under the approach

(i) Serbo-Croatian

a. ?Knjige$_i$ je Marko [t$_i$ i filmove] kupio.
books is Marko and movies bought
'Marko bought books and movies.'

b. On je ušao u veliku sobu.
he is entered in big room
'He entered a big room.'

c. U veliku je ušao sobu.

It may also be worth noting here that the clitichood of the conjunction may not be the only requirement for the possibility of a CSC-2 violation. Oda notes that all the languages that he observes can violate CSC-2 lack articles, which may suggest that such violations may be possible only in NP languages under Bošković's (2008; 2012) analysis, where languages without articles lack DP (for an account along these lines, see Bošković 2017).

to rescue by PF deletion discussed above. The reason why the conjunct (a traditional adjunct) in (42) is then able to undergo movement is the same as the reason why the conjunct in (41) (a traditional conjunct) is able to undergo movement.[25] What we see here is that a ConjP that is headed by a trace behaves like traditional adjunction modification, which under the current analysis involves a ConjP with a null head, in that both cases void islandhood, a state of affairs that can be captured by the rescue-by-PF-deletion mechanism.

(42) How did John walk?

The analysis thus unifies the possibility of extraction out of the VP conjunct in (27) and the improvement with extraction out of a traditional conjunct in (34–35) with the possibility of extraction of a traditional conjunct in (41) and the traditional adjunct in (42); what matters in all these cases is that the head of the island, the conjunct and ConjP in the former case and ConjP in the latter case, is phonologically null, which is captured under the rescue-by-PF deletion analysis.

There is an interesting prediction made by the current analysis that is worth noting at this point. Recall that, as argued in Oda (2017), both conjuncts and ConjP are islands. In cases like Galician (34), both of these islands are "violated". In (35), on other hand, the islandhood of the conjunct island is voided since the head of the conjunct is phonologically null as a result of article incorporation. Recall now that in languages like Japanese and SC, the head of ConjP (in traditional coordinations) is actually phonologically null (due to conjunction incorporation). This means that extraction out of a conjunct in Japanese and SC involves extraction out of only one island, the conjunct. As a result, we would expect it to be better than extraction out of a conjunct in English and Galician (34) – it should be more on a par with Galician (35) than Galician (34). The prediction is in fact more general, it holds for all languages where extraction of a conjunct is possible; more precisely, in languages where CSC-2 can be voided by incorporating the conjunction head CSC-1 violations should be somewhat weaker than in languages where this is not the case (unless such languages have a way of incorporating the conjunct head, like Galician). It is obviously difficult to compare the strength of island violations across different languages, but impressionistically, CSC-1 violations do seem to be slightly weaker in Japanese and SC than in English (one bilingual Japanese/English speaker consulted did find that CSC-1 violations with Japanese scrambling are weaker than CSC-1 violations with English

[25] As discussed in Oda (2017), extraction of the second conjunct in traditional coordinations is not possible in Japanese for an independent PF reason that does not arise in (42) (the reason also does not arise with *wh*-in-situ in Japanese, which Oda notes is possible as both the first and the second conjunct).

topicalization). Obviously, a more careful investigation is needed here, which I leave for future research.[26]

The proposed analysis makes a similar prediction regarding the strength of CSC-1 violations and the adjunct condition violation. Consider cases where no islandhood is voided through movement of island heads (cf. 28). As discussed above, both conjuncts and ConjP are islands. Extraction out of a conjunct then involves two island violations. Since adjuncts are treated as conjuncts, extraction out of an adjunct also involves extraction out of a conjunct island and a ConjP island. However, since with adjuncts the head of ConjP is phonologically null, the islandhood effect of ConjP is voided, as discussed above. Extraction out of an adjunct then involves one island violation. We may then expect that CSC-1 violations should be stronger than adjunct condition violations in a language like English. That indeed seems to be the case: CSC-1 violations like (4) seem to be worse than adjunct condition violations like (5) (as noted above, the prediction is also borne out with Galician (31) and (34), (34) being worse than (31)). On the other hand, in a language like SC where the head of ConjP is also phonologically null due to the cliticization of the conjunction, extraction out of both conjuncts and adjuncts involves extraction out of a single island. CSC-1 violations and the adjunct condition violations indeed seem to have more or less the same status in SC. Of course, all the predictions noted in this passage still need to be confirmed with more careful data elicitation.

6 Conclusion

This paper has argued for a unified approach to the islandhood of conjuncts and adjuncts, both of which disallow extraction out of them. The unification was made possible by adopting Higginbotham's semantics of traditional adjunction, on which traditional adjunction actually involves coordination. This paper took

[26] It is worth noting here that Oda (2017) observes a construction in SC where both the conjunct and ConjP are headed by a trace, namely (i).

(i) (?)[U veliku]$_i$ je Ivan ušao [[t$_i$ sobu] i u malu kuhinju].
in big is Ivan entered room and in small kitchen

As noted in footnote 24, the conjunction undergoes procliticization in SC, which means ConjP is headed by a trace in (i). Moreover, as also discussed in footnote 24, the head of the first conjunct, which is a PP, undergoes procliticization to the AP, and is carried along under movement of the AP. As a result of P-procliticization, the conjunct from which the AP is extracted is also headed by a trace. Both the islandhood of ConjP and the first conjunct are then voided in (i) through the rescue-by-PF deletion mechanism, hence the acceptability of (i).

this to be reflected in the syntax, with ConjP present in the syntax of traditional adjunction (see also Progovac 1998; 1999). Not only did this position achieve straightforward syntax-semantics mapping in the case at hand, but it also made possible a unification of the islandhood of conjuncts and traditional adjuncts since the two then involve the same syntactic configuration.

I have shown that there are a number of similarities in the islandhood of conjuncts and adjuncts, including the general resistance of their islandhood to crosslinguistic variation (in contrast to other traditional islands, which are subject to crosslinguistic variation). We have also seen that in a number of environments extraction is exceptionally possible out of conjuncts and adjuncts. Significantly, the environments where extraction is exceptionally possible are the same for conjuncts and adjuncts, which can be captured if the two involve the same syntactic configuration. A number of important issues, however, still remain to be addressed in future research, including the question why conjunctions are typically null with traditional adjuncts and overt with traditional coordination, as well as providing an actual account of the islandhood of conjuncts/adjuncts.

The intuition regarding the former issue seems clear: there are choices when it comes to what heads ConjP in traditional coordinations. Even if we put aside the obvious major distinction here, conjunction vs disjunction, languages often have more than one coordinator, which come with different flavors syntactically and/or semantically (note e.g. that the coordinator that hosts article incorporation in Galician is not simple *e* 'and' but *e mais*); in other words, phonological realization of conjunction is a way of making a choice of which coordinator to use. Traditional adjunction, on the other hand, involves the most neutral, straight coordination which does not add anything else – this is the null $Conj^0$.[27]

Some preliminary remarks were also made regarding the islandhood of conjuncts/adjuncts (an issue that is discussed in more detail from the perspective taken in this paper in Oda 2017 and Bošković 2017; see also Bošković 2020). Importantly, it was shown that in several cases where the islandhood of traditional conjunction configurations is voided (for both individual conjuncts and the conjunction phrase itself), where traditional adjunction configurations also do not

[27]This does not mean that null $Conj^0$ can never be used with traditional coordination (see Progovac 1999 for some such cases) or that an overt $Conj^0$ cannot be used in traditional adjunct modification. Regarding the latter, as noted in §4, Progovac (1998; 1999) discusses examples like *I read his paper, and quickly* and *John read the book and avidly*. Also relevant in the context of the current discussion is Progovac's (1999) economy of pronunciation which works in a similar way as Chomsky's (1981) *avoid pronoun principle*, choosing the null conjunction head when possible (Progovac 1998 in fact adopts *avoid overt conjunction*).

show islandhood (in both respects), the head of the conjunction (and individual conjuncts) is phonologically null, with the parallel situation holding for the traditional adjunction configuration, a state of affairs which was captured by appealing to the rescue-by-PF deletion mechanism. We have also seen that the rescue-by-PF deletion analysis can account in a principled way for a number of differences in the strength of the violation with extraction out of conjuncts and adjuncts in various languages/contexts.

Abbreviations

AC	adjunct condition	INS	instrumental
ACC	accusative	NOM	nominative
ATB	across-the-board	PF	phonetic form
CSC	coordinate structure constraint	PIC	phase impenetrability condition
DAT	dative	Q	question particle
DOM	differential object marking	REFL	reflexive
ECP	empty category principle	SC	Serbo-Croatian
EPP	extended projection principle	TNP	traditional NP
GEN	genitive	TOP	topic

Acknowledgements

It is a pleasure and privilege to be able to dedicate this paper to Ian Roberts, for his invaluable and lasting contributions to the field of linguistics.

For helpful comments on this work I thank two anonymous reviewers and the participants of my 2016 seminar at the University of Connecticut.

References

Abels, Klaus. 2011. Don't repair that island! It ain't broke. Paper presented at Islands in the Contemporary Theory, University of the Basque Country, Victoria-Gasteiz.

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.

Barros, Matthew, Patrick D. Elliott & Gary Thoms. 2014. There is no island repair. Ms., Rutgers University, University College London, and University of Edinburgh.

Boeckx, Cedric & Howard Lasnik. 2006. *Intervention and repair*. DOI: 10.1162/ling.2006.37.1.150.

Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59(1). 1–45. DOI: 10.1111/j.1467-9582.2005.00118.x.

Bošković, Željko. 2008. What will you have, DP or NP? In Emily Elfner & Martin Walkow (eds.), *NELS 37: Proceedings of the 37th annual meeting of the North East Linguistic Society*, 101–114. Amherst, MA: GLSA Publications.

Bošković, Željko. 2011. *Rescue by PF deletion, traces as (non)interveners, and the* that*-trace effect*. DOI: 10.1162/LING_a_00027.

Bošković, Željko. 2012. On NPs and clauses. In Günther Grewendorf & Thomas Ede Zimmermann (eds.), *Discourse and grammar: From sentence types to lexical categories*, 179–246. Berlin: De Gruyter.

Bošković, Željko. 2013a. Principles and parameters theory and minimalism. In Marcel den Dikken (ed.), *The Cambridge handbook of generative syntax*, 95–121. Cambridge: Cambridge University Press.

Bošković, Željko. 2013b. Traces do not head islands: Essays in honor of Mamoru Saito. In Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki & Asako Uchibori (eds.), *Deep insights, broad perspectives*, 56–93. Tokyo: Kaitakusha.

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1). 27–89. DOI: 10.1162/LING_a_00148.

Bošković, Željko. 2015. From the Complex NP Constraint to everything: On deep extractions across categories. *The Linguistic Review* 32(4). 603–669. DOI: 10.1515/tlr-2015-0006.

Bošković, Željko. 2016. What is sent to spell-out is phases, not phasal complements. *Linguistica* 56(1). 25–66. DOI: 10.4312/linguistica.56.1.25-66.

Bošković, Željko. 2017. On the coordinate structure constraint, islandhood, phases, and rescue by PF deletion. Ms., University of Connecticut. https://boskovic.linguistics.uconn.edu/wp-content/uploads/sites/2801/2019/05/CSCFranksShortNoAcknow.pdf.

Bošković, Željko. 2018. On movement out of moved elements, labels, and phases. *Linguistic Inquiry* 49(2). 247–282. DOI: 10.1162/LING_a_00273.

Bošković, Željko. 2020. On the coordinate structure constraint, across-the-board movement, and labeling. In Jeroen van Craenenbroeck, Cora Pots & Tanja Temmerman (eds.), *Recent developments in phase theory*. Berlin: De Gruyter.

Chomsky, Noam. 1972. *Spheres of influence in the age of imperialism: Papers submitted to the Bertrand Russell Centenary Symposium, Linz, Austria, 11–*

15 September 1972. Nottingham: Bertrand Russell Peace Foundation for The Spokesman.
Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8(2). 81–120. DOI: 10.1111/j.1467-9612.2005.00076.x.
Corver, Norbert. 1992. Left branch extraction. In Kimberly Broderick (ed.), *NELS 22: Proceedings of the twenty-second annual meeting of the North East Linguistic Society*, 67–84. Amherst, MA: GLSA Publications.
Davidson, Donald D. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*, 81–95. Pittsburgh: Pittsburgh University Press.
Despić, Miloje. 2011. *Syntax in the absence of determiner phrase*. University of Connecticut. (Doctoral dissertation).
Despić, Miloje. 2013. Binding and the structure of NP in Serbo-Croatian. *Linguistic Inquiry* 44(2). 239–270. DOI: 10.1162/ling_a_00126.
Dowty, David. 1989. On the semantic content of the notion of 'thematic role'. In Gennaro Chierchia, Barbara H. Partee & Raymond Turner (eds.), *Properties, types and meaning*, vol. II: Semantic issues. Dordrecht: Kluwer.
Engdahl, Elisabet. 1986. *Constituent questions*. Dordrecht: Reidel.
Epstein, Samuel David. 1999. Un-principled syntax: The derivation of syntactic relations. In Samuel David Epstein & Norbert Hornstein (eds.), *Working minimalism*, 317–345. Cambridge, MA: MIT Press.
Fox, Danny & Howard Lasnik. 2003. Successive-cyclic movement and island repair: The difference between sluicing and VP-ellipsis. *Linguistic Inquiry* 34(1). 143–154. DOI: 10.1162/00243903763255959.
Franks, Steven. 1993. On parallelism in across-the-board dependencies. *Linguistic Inquiry* 24(3). 509–528.
Grosu, Alexander. 1973. On the nonunitary nature of the Coordinate Structure Constraint. *Linguistic Inquiry* 4(1). 88–92.
Haïk, Isabelle. 1985. *The syntax of operators*. MIT. (Doctoral dissertation).
Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16(4). 547–593.
Holmberg, Anders. 2005. Is there a little pro? Evidence from Finnish. *Linguistic Inquiry* 36(4). 533–564. DOI: 10.1162/002438905774464322.

Hornstein, Norbert, Howard Lasnik & Juan Uriagereka. 2003. The dynamics of islands: Speculations on the locality of movement. *Linguistic Analysis* 33(1–2). 149–175.

Hunter, Tim. 2011. *Syntactic effects of conjunctivist semantics: Unifying movement and adjunction.* Amsterdam: John Benjamins. DOI: 10.1075/la.170.

Huybregts, Riny & Henk C. van Riemsdijk. 1985. Parasitic gaps and ATB. In Stephen Berman, Jae-Woong Choe & Joyce McDonough (eds.), *NELS 15: Proceedings of the fifteenth annual meeting of the North East Linguistic Society*, 168–187. Amherst, MA: GLSA Publications.

Johnson, Kyle. 2009. Gapping is not (VP-) ellipsis. *Linguistic Inquiry* 40(2). 289–328. DOI: 10.1162/ling.2009.40.2.289.

Kayne, Richard S. 1994. *The antisymmetry of syntax.* Cambridge, MA: MIT Press.

Lasnik, Howard. 2001. When can you save a structure by destroying it? In Min-Joo Kim & Uri Strauss (eds.), *NELS 31: Proceedings of the thirty-first annual meeting of the North East Linguistic Society 31*, 301–320. Amherst, MA: GLSA Publications.

Law, Paul. 1993. On the base position of *wh*-adjuncts and extraction. Paper presented at the 67th annual meeting of the Linguistic Society of America, Los Angeles, 7–10 January.

Marelj, Marijana. 2011. Bound-variable anaphora and left branch condition. *Syntax* 14(4). 205–229. DOI: 10.1111/j.1467-9612.2011.00156.x.

Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis.* Oxford: Oxford University Press.

Munn, Alan Boag. 1993. *Topics in the syntax and semantics of coordinate structures.* The University of Maryland. (Doctoral dissertation).

Nunes, Jairo. 2004. *Linearization of chains and sideward movement.* Cambridge, MA: MIT Press.

Oda, Hiromune. 2017. Two types of the coordinate structure constraint and rescue by PF deletion. In Andrew Lamont & Katerina Tetzloff (eds.), *NELS 47: Proceedings of the 47th annual meeting of the North East Linguistic Society*, 343–356. Amherst, MA: GLSA Publications.

Parsons, Terence. 1980. Modifiers and quantifiers in natural language. *Canadian Journal of Philosophy* Supplementary volume 6. 29–60. DOI: 10.1080/00455091.1980.10715756.

Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics.* Cambridge, MA: MIT Press.

Postal, Paul M. 1998. *Three investigations of extraction.* Cambridge, MA: MIT Press. 215.

Progovac, Liljana. 1998. Avoid conjunction, adjunction, and the coordination of likes constraint. In Željko Bošković, Steven Franks & William Snyder (eds.), *Formal approaches to Slavic linguistics: The Connecticut meeting 1997*, 252–266. Ann Arbor, MI: Michigan Slavic Publications.

Progovac, Liljana. 1999. Events and economy of coordination. *Syntax* 2(2). 141–159. DOI: 10.1111/1467-9612.00018.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT. (Doctoral dissertation).

Ross, John Robert. 1969. Guess who? In Robert Binnick, Alice Davison, Georgia M. Green & Jerry L. Morgan (eds.), *Papers from the fifth regional meeting of the Chicago Linguistic Society*, 252–286. Chicago: University of Chicago.

Ross, John Robert. 1974. Three batons for cognitive psychology. In Walter B. Weimer & David S. Palermo (eds.), *Cognition and the symbolic processes*, 63–124. Hillsdale, NJ: Lawrence Erlbaum.

Runić, Jelena. 2014a. A new look at argument ellipsis: Evidence from Slavic. In Hsin-Lun Huang, Ethan Poole & Amanda Rysling (eds.), *NELS 43: Proceedings of the forty-third annual meeting of the North East Linguistic Society*, 91–102. Amherst, MA: GLSA Publications.

Runić, Jelena. 2014b. *A new look at clitics, clitic doubling, and argument ellipsis: Evidence from Slavic*. University of Connecticut, Storrs. (Doctoral dissertation).

Stepanov, Arthur. 2001a. *Cyclic domains in syntactic theory*. University of Connecticut, Storrs. (Doctoral dissertation).

Stepanov, Arthur. 2001b. Late adjunction and minimalist phrase structure. *Syntax* 4(2). 94–125. DOI: 10.1111/1467-9612.00038.

Stjepanović, Sandra. 2014. Left branch extraction and the Coordinate Structure Constraint. In Jyoti Iyer & Leland Kusmer (eds.), *NELS 44: Proceedings of the forty-fourth annual meeting of the North East Linguistic Society*, 157–170. Amherst, MA: GLSA Publications.

Stowell, Timothy. 1981. *Origins of phrase structure*. MIT. (Doctoral dissertation).

Takahashi, Daiko. 1994. *Minimality of movement*. University of Connecticut, Storrs. (Doctoral dissertation).

Takahashi, Masahiko. 2012. Phases and the structure of NP: A comparative study of Japanese and Serbo-Croatian. In Steven Franks, Markus Dickinson, George Fowler, Melissa Whitcombe & Ksenia Zanon (eds.), *Proceedings of Formal Approaches to Slavic Linguistics 21*, 359–372. Ann Arbor, MI: Michigan Slavic Publications.

Talić, Aida. 2014. Extraordinary complement extraction: PP-complements and inherently case-marked nominal complements. *Studies in Polish Linguistics* 8(3). 127–150. DOI: 10.4467/23005920SPL.13.007.1542.

Talić, Aida. 2015. Adverb extraction, specificity, and structural parallelism. *Canadian Journal of Linguistics* 60(3). 417–454. DOI: 10.1017/S000841310002627X.

Trenkić, Danijela. 2004. Definiteness in Serbian/Croatian/Bosnian and some implications for the general structure of the nominal phrase. *Lingua* 114(11). 1401–1427. DOI: 10.1016/j.lingua.2003.09.005.

Truswell, Robert. 2011. *Events, phrases and questions*. Oxford: Oxford University Press.

Uriagereka, Juan. 1988. *On government*. University of Connecticut, Storrs. (Doctoral dissertation).

Uriagereka, Juan. 1996. Determiner clitic placement. In Robert Freiden (ed.), *Current issues in comparative grammar*, 257–295. Dordrecht: Kluwer.

Uriagereka, Juan. 1999. Multiple spell-out. In Norbert Hornstein & Samuel David Epstein (eds.), *Working minimalism*, 251–282. Cambridge, MA: MIT Press.

Williams, Edwin. 1990. The ATB theory of parasitic gaps. *The Linguistic Review* 6(3). 265–279. DOI: 10.1515/tlir.1987.6.3.265.

Zlatić, Larisa. 1997. *The structure of the Serbian noun phrase*. University of Texas, Austin. (Doctoral dissertation).

Chapter 13

Re-thinking re-categorization: Is *that* really a complementizer?

Ellen Brandner
Universität Stuttgart

Following Kayne's (2014) argumentation that the complementizer *that* is indeed a relative pronoun and with it the complement clause a special type of relative clause (explicative, i.e. without a gap), the paper contributes to the discussion whether *that*-complement clauses are also structurally relative clauses. One consequence of this would be that *that*-clauses should not allow long wh-extraction, contrary to what is observed in languages like English at first sight. However, the distribution of resumptive pronouns in Alemannic, a Southern German dialect, indeed points into that direction. Like the Celtic languages, Alemannic has a special particle for relative clauses but can use the d-pronoun strategy as well. Both strategies can be used to build long distance dependencies alike. But resumptive pronouns are nearly obligatory with *that*-clauses in sharp contrast to those involving relative clauses. This difference can find an explanation, if the particle-strategy creates a genuine gap in the embedded clause whereas a *that*-complement clause is always a full-fledged clause and the gap in it is only apparent, its appearance regulated by outer-syntactic criteria.

1 Introduction

The more or less established analysis of complementizers of the English *that*-type is that they evolved out of pronominal elements, most commonly the (distal) demonstrative pronoun:

(1) *That* guy over there gives me a headache (demonstrative)

(2) Do you believe *that*? (anaphoric)

Ellen Brandner. 2020. Re-thinking re-categorization: Is *that* really a complementizer? In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 259–274. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280651

(3) I believe [*that*…] (complementizer)

The diachronic scenario, already proposed in very early[1] work, assumes that *that* (and its equivalents in the other Germanic languages) originated as a (cataphoric) pronoun to the following (independent) clause. A re-bracketing of the clausal boundaries posited the pronoun then to the left edge of the embedded clause, see e.g. Roberts & Roussou (2003) for an explicit proposal:

(4) I say *that*: [main clause] → I say [*that* embedded clause]

This process involves in addition to the re-bracketing a re-categorization of *that* such that the previously pronoun enters into the class of C-elements and thus belongs now to the "word class" of complementizers. As such it occupies the C^0-position, i.e. it has not only changed its word class but also its phrase structural status in that it is re-analyzed as a head. Van Gelderen (2004) takes especially this type of reanalysis (Spec-to-head) as a hallmark of the grammaticalization process. Evidence for the head-status of complementizer-*that* is seen in the fact that *that*-clauses allow already in the early stages (e.g. on Old High German) for long wh-extraction – a process which must rely on an empty specifier in the CP as an available intermediate landing site, see Axel (2009; 2017) for this line of reasoning. This scenario is assumed to not only be true of German; the same process has taken place in English and the other Germanic languages.

Now various authors have cast doubt on the assumption that there is indeed such a re-analysis process and ask whether speaking of a category C (in the sense of a word class) is at best misleading – in the worst case it is blurring the actual problem to be solved, e.g. Kayne (2014); Manzini & Savoia (2003; 2011). These authors suggest that we should follow the "WYSWYG-principle" and under this premise *that* (and its cognates in other languages) is indeed never something else than a pronoun. While Manzini & Savoia remain a bit vague about its actual status – besides the claim that Romance *che* ('what') is a quantificational element whose restrictor can also be a proposition (= acting then as a complementizer), Kayne states plainly that *that* is always a relative pronoun and accordingly complement clauses are always relative clauses, construed with a (possibly empty) correlate pronoun in the matrix clause.

This is essentially the analysis proposed in Axel (2009; 2017). She rejects the re-bracketing analysis, based on data in OHG.[2] Like Kayne (2014), she proposes

[1] For example Müller & Frings (1959), but the idea can already be found in very early work from the 19th century, see Axel (2009; 2017) for a survey and further references.

[2] Recall that in OHG, there is a clear distinction between root and embedded clauses due to the position of the finite verb (V2 order vs. verb final in embedded clauses).

that *that* is a relative pronoun, belonging thus to the embedded clause from the beginning on, and assuming that there is a (possibly silent) head noun in the matrix clause. This is in spirit very close to Kayne (2014).[3]

The scenario in (4) would then look like the one in (4′).

(4′) I say (*that/it*) [*that* …embedded clause (= relative clause)]

By showing that long wh-extractions already exist at this stage of the language, a crucial component for her analysis is the Spec-to-head reanalysis – as only in this configuration, long wh-extraction is possible, due to the now empty specifier.

On the other hand, if one follows the Kayne-analysis according to which the "complementizer" is indeed a relative pronoun, one would expect that long wh-extraction out of a *that*-clause cannot exist at all – given that relative clauses are for sure one of the strongest islands for extraction.

In this paper, I will show that there are good reasons to think that Kayne's position is actually correct: there is evidence from the Alemannic dialect, spoken in Southern Germany and Switzerland, that there is no long (cyclic) wh-movement out of *that*-type complement clauses and what looks like extractions – leaving behind a gap – consists of a base-generated wh-phrase in the matrix clause and an actually full-fledged complement clause with a pronoun filling the "extraction-site". This pronoun can be PF-deleted under a rather weak principle like e.g. the avoid pronoun principle (Chomsky 1981), giving thus merely the impression of actual movement.

However, the grammar has a strategy to build long wh-dependencies (LWDs) with real gaps – but this is only possible if the gap in the embedded clause is a genuine gap, coming into existence via a special type of complementizer, used normally in the formation of relative clauses, turning the embedded clause into a predicate. The situation I am referring to is described and analysed in Adger & Ramchand's (2005) work on LWDs in Gaelic (Celtic). I will present evidence here that the very same strategy is used in some variants of Germanic as well. But in contrast to Adger & Ramchand (2005) who suggest that there is a parametric difference between Celtic and Germanic (English in this case) which allows the derivation of genuine long wh-extractions in the latter, I will show that this is not

[3]The difference to a "usual" relative clause is that there is no overtly detectable gap in it. This has to do with the type of the head noun that is modified by the relative clause: it is clearly a kind of a direct object (realizable as a correlate pronoun). The semantic content of this pronoun is actually a proposition – and the relative clause is delivering the content of this proposition. This might be formally analysed in terms of an *aboutness relative*, i.e. a gap-less one, see van Riemsdijk (2003), Cheng & Sybesma (2005), as suggested in Brandner & Bucheli (2018), also Axel (2009; 2017).

true for at least Alemannic. Further and more detailed research – along the lines that will be presented here – will be necessary to make the point valid also for English and other Germanic languages – actually for all languages that have to be claimed to exhibit long wh-extractions. I am aware that this is a far reaching claim – still the data presented should be taken to be an invitation to re-think in general the issue of long wh-extractions.

The data that support this suggestion come from the Southern German dialect Alemannic (ALM). A large scale study about LWDs in the whole Alemannic speaking area revealed that this language uses the same strategy to build LWDs as the Celtic languages. In addition, however – and in contrast to the Celtic languages – Alemannic shows LWDs with *that*-clauses, indicating that a parametric solution as proposed in Adger & Ramchand (2005) is probably not the right way to look at it. Secondly, it will be shown below that these seemingly extractions are in reality no extractions at all. The main evidence comes from the distribution of resumptive pronouns that occur in these "extractions". They occur to such a high percentage that it leaves no room for an actual extraction analysis. Especially, if one assumes that resumptives are inserted to "rescue" an otherwise impossible structure (island violations) or reduce parsing complexity, see Chao & Sells (1983), it would remain a complete mystery why the very same complexity allows or even requires a gap when the LWD is built via relative clause formation.

2 The two strategies

LWDs in Alemannic show up in several versions. Besides the familiar strategies that are also found in Standard German (or at least the spoken variants of it), see the examples in (5a–c), there is a possibility that has to my knowledge not been noted until now, see for a first description Brandner & Bucheli (2018), illustrated with Standard German wording in (5d):

(5) German
Wen hast du gesagt…

a.	[dass Maria ____ gesehen hat]	*dass*-LWD
b.	[wen Maria____ gesehen hat]	copy const.
c.	(was) [wen Maria ____ gesehen hat]	*was*-*w*-constr.
d.	[wo Maria ____ gesehen hat]	*wo*-LWD

'Who did you say that Mary saw?'

The interesting thing about the strategy in (5d) is that the complementizer in the embedded clause corresponds to the one used regularly in relative clauses in this variety, cf. (6), glossed as RCI (relative clause introducer); note that the declarative complementizer in ALM is *dass*, glossed as CCI (complement clause introducer), like in Standard German:

(6) Alemannic
d'frau [*wo*-n-i geschtert ___ troffe ha]
the woman RCI I yesterday met have

(7) Alemannic
mir het er gseet [*dass* er erscht schpöter kunnt]
me has he told CCI he only later comes

Examples like (5d) showed up first during the survey period of SADS[4] where informants offered it as one possible version to express a LWD of the type given in (5a). In the project SynAlm,[5] these were then examined in more detail and contrasted with the "usual" strategy, i.e. *dass*-LWDs. It turned out that both strategies are possible in Alemannic and are in more or less free variation. The large scale investigation (about 580 speakers) in the whole Alemannic speaking area (Switzerland, Southwest Germany, Alsatian and Austria) conducted by SynAlm concerning *wo*-LWDs revealed the following main results:

- *wo*-LWDs were throughout accepted by more than 50% of the speakers, notably the acceptance/rejection is essentially the same as with *dass*-LWDs[6]
- no clear areal patterns could be detected, i.e. it is not the case that there are certain (areally definable) sub-dialects of Alemannic that allow for *wo*-LWDs whereas others do not. Instead, it seems that Alemannic speakers have simply both possibilities at their disposal.

[4] *Syntaktischer Atlas der deutschen Schweiz*, (http://www.dialektsyntax.uzh.ch/de.html).

[5] The study was conducted within the DFG-supported project SynAlm (https://ilg-server.ling.uni-stuttgart.de/synalm/html/). Its funding time was from 2011–2015. SynAlm gathered its data via written questionnaires, mostly using judgments (5-point scale) for examples constructed as minimal pairs. Seven questionnaires were sent out. The number of informants range from 580 to 1000. No informant was excluded but data concerning age, social status, and origin (also of the parents) were collected.

[6] LWDs are generally accepted only by a certain amount of speakers. This holds for Standard German as well as for the dialects. It should also be kept in mind that there are various strategies at the disposal (copy-construction, scope marking etc.). The informants had always the possibility to give an own version of the sentence asked for. In many cases, the informants judged the presented example as bad and chose a parenthetical construction as an alternative, i.e. where there is no extraction at all.

- there was no effect with respect to age: younger speakers accepted the construction to the same percentage as older speakers.

Now Alemannic is not the only language that has a special complementizer in relative clauses (RCs). The Celtic languages are well-known for using a similar strategy like Alemannic by employing a specialized particle in RCs, see e.g. (McCloskey 2001; 2002 and following work) for Irish. The "typical" complementizer for complement clauses is illustrated in (8a). (8b) illustrates an RC, compare these with the ALM clauses in (6) and (7):

(8) Irish

a. Deir siad [*gur* ghoid na síogaí í].
say they go-PST stole the fairies her
'They say that the fairies stole her away.'

b. an ghirseach [*a* ghoid na síogaí ____]
the girl RCI stole the fairies
'the girl that the fairies stole away'

The LWDs in (9) and (10) show that it is the RCI that occurs in LWDs, whereas LWDs out of a *go* (= *gun*)-clause are impossible:

(9) Irish

a. Cé a mheas tú *a* chonaic tú?
who aL thought you aL saw you
'Who did you think that you saw?'

b. Cén t-úrscéal a mheas mé *a* dúirt sé *a* thuig sé.
which novel aL thought I aL said he aL understood he
'Which novel did I think he said he understood?'

(10) Irish

*Dè a thuirt sibh *gun* sgrìobh i?
what C-REL said you that wrote she
'What did you say that she wrote?'

Welsh shows a comparable pattern – although the fact that the LWD is built on a relative clause can be seen here only indirectly since the relative particle does not show up overtly: however, the embedded verb in LWDs is in the so-called "relative form", the morpho-syntactic reflex of having a gap in the clause. Welsh examples taken from Willis (2000: 555).

(11) Welsh
Beth ych chi 'n gredu *sy* 'n wir bwysig miwn
what are you PROG believe-VN IS-REL PRED truly important in
cymdeithas?
society
'What do you think is truly important in society.'

Even other Germanic languages are reported to allow for structures similar to the one in (5d). The following pattern is from Norwegian (Westergaard et al. 2012):

(12) Norwegian
a. Hvem tror du [*som* ___ har gjort det]?
who think you RCI has done it
'Who do you think has done it?'
b. Hvem tror du [*at* ___ har gjort det]?
who think you *that* has done it
'Who do you think has done it?'

In sum, LWDs based on an RC-structure are quite common – also in the Germanic languages – and they occur as an alternative to the (until now) more widely attested *dass*-LWDs, together with the scope-marking and copying constructions – and of course with parenthetical constructions – which seem to be always a possibility.

In SynALm, the acceptance/rejection of resumptive pronouns was systematically tested against these various types of LWDs and it is this last set of data that gave the crucial clue for the claim from above, namely that in *dass*-clauses, there is merely an apparent "gap" and it is only in *wo*-LWDs where genuine gaps show up.

3 Distribution of resumptive pronouns

Until now, we have only seen that Alemannic is similar to the Celtic languages in that it allows LWDs based on RCs. However, the important difference is that Alemannic (together with Norwegian) allows LWDs based on *dass*-clauses as well – in sharp contrast to Celtic. Given the considerations from above, namely that *dass* is a real relative pronoun, it is the Celtic languages that behave as expected. The possibility of LWDs in the Germanic languages (including of course English) is then the fact to be explained.

In the following, I will use the distribution of resumptive pronouns in the various types of LWDs to show that "extraction" out of *dass*-clauses is indeed an illusion: all the extracted arguments can be realized as pronouns and whether they are spelled-out overtly or not is a matter of phonetic form (PF) – where (non-syntactic) factors like distance etc. play a role.

3.1 Resumptive pronouns in Alemannic relative clauses

Before going into the details of the distribution of resumptives in LWDs, a brief illustration of the occurrence of resumptive pronouns in simple RCs in Alemannic is necessary: it has often been claimed in the literature on Alemannic RCs (in this case specifically on Zürich German), that in case of datives and the oblique positions further down in the Keenan/Comrie hierarchy, resumptives occur obligatorily, see van Riemsdijk (2003), Salzmann (2006) among others. Thus, whereas with subjects and objects, resumptives never show up, they occur from the dative-position on, illustrated here only with a dative-argument and a subject-relativization:

(13) Zürich German

a. der Bue [wo ma *em* s'Velo verschprooche het]
the boy RCI one him the=bike promised has
'the boy, who was promised to get a bike'

b. der Bue [wo-n[7]-(**er*) zschpot kummen-isch]
the boy RCI he too late come is
'The boy who arrived too late'

In SynAlm, it could be shown, that this claim is empirically not tenable. Although it is true that there never occur resumptives with subjects and (direct) objects, one can hardly speak of "obligatoriness of dative-resumptives" in light of an acceptance rate ranging between 9–15%.[8] With the oblique-positions further down in the Keenan/Comrie hierarchy, the acceptance/requirement of a resumptive increases accordingly. So we can safely conclude that the occurrence of resumptives in simple RCs follows the expected distribution – whatever the ultimate (syntactic) reason behind the pattern described in the Keenan/Comrie hierarchy – may be.[9]

[7] *-n-* is an epenthetic consonant and is of no relevance here.

[8] Many more sentences with dative-resumptives were tested and the result was basically the same with some minor variation – having probably more to do with the general naturalness of the example and other linguistically insignificant factors.

[9] I will not take a stand here whether this has to do with the necessity to realize oblique/

3.2 Resumptive pronouns in simple LWDs

Equipped with this background let us now turn to the distribution of resumptives in LWDs, both based on *wo*-RCs and *dass*-clauses. The expectation for the *wo*-LWDs is that they show a comparable distribution of resumptives as in simple RCs – given that they have both the same underlying syntax.[10] In *dass*-clauses on the other hand, the assumption of an extraction strategy would one lead to expect that gaps are predominant. However, it turns out that the results are essentially the opposite: resumptives are accepted to a much higher degree in *dass*-LWDs. The results concerning the acceptance of resumptives are given in Table 13.1.

Table 13.1: Acceptance of resumptive pronouns in different types of LWDs and RCs ($n = 580$).

Type of "extracted" phrase	*dass*-LWD	*wo*-LWD	*wo*-RC
subject	70%	9%	–
direct object	30%	5%	–
dative object	43%	12%	15%
adjunct	60%	62%	51%

Although there occur resumptives also with *wo*-LWDs with subjects[11] and (direct) objects to a certain extent – whereas they are categorically excluded in genuine relative clauses – the important difference is the acceptance rate of resumptives in *dass*-LWDs. For subjects, it is evident. The lower acceptance of resumptives (or rather the possibility to have a gap) in direct object position may have to with the fact that many simple transitive verbs have a grammatical output when used as a mere activity verb (*I read a book* vs. *I read*). But this has to be investigated in more detail in future research.

On the other hand, resumptives for datives and obliques in *wo*-LWDs show a rather even distribution with their occurrence in *wo*-RCs. In *dass*-LWDs again,

morphological case – as suggested in Salzmann (2006) or whether different factors are at stake, see for some speculations Brandner & Bucheli (2018). It should be noted that informants who did neither accept a gap nor a resumptive in the relativization of oblique positions adhered simply to a bi-clausal structure, i.e. the formation of an RC was avoided.

[10] Recall that I assume with Adger & Ramchand (2005) that the wh-phrase in the matrix is base-generated there and the gap in the embedded clause is licensed by a local configuration with the respective complementizer whose internal lexical specification allows/requires a gap in its complement (the so-called lambda-feature). I refer to their work for the technical details.

[11] The high acceptance of a resumptive in subject-LWDs does not really come as a surprise – since – as is well known since the work by Engdahl (1985), resumptives in subject positions may occur to avoid an ECP-violation (*that*-trace-effect). In light of the discussion, this fact should be reconsidered again.

datives have a considerably higher acceptance whereas the adjunct behaves similar under all conditions. I will not go into a thorough discussion of these results – since I will take them here merely as a first hint that the resumptives in *dass*-LWDs are maybe not really "resumptives" – but that the embedded clause in a *dass*-LWD is full-fledged in the sense that there are no syntactic gaps – but that all positions are syntactically occupied by a co-referent pronoun – and its PF-realization is subject to non-syntactic conditions. The next set of data shows this difference very clearly.

3.3 Resumptive pronouns in LWDs across two clause boundaries

The acceptance of resumptives was also tested across two clause boundaries, i.e. a situation where the occurrence/acceptance of resumptives can more easily attributed to outer-syntactic (i.e. parsing) properties. The test sentence is given in English wording in (14):

(14) Who did you say [dass / wo Mary heard [dass/wo ___ had an accident]]

We varied the complementizers and resumptives as shown in Table 13.2.

Table 13.2: Acceptance of gap/resumptive in subject position in LWDs crossing two clause boundaries (n = 580).

Variation of comps	Acceptance (1–2)	Complete rejection (5)
dass…dass…gap	30%	22%
dass…dass…resumptive	70%	5%
wo…wo…gap	31%	23%
wo…wo…resumptive	8%	45%

The results show clearly that the acceptance of resumptives is directly connected to the type of the complementizer. Again: *dass*-LWDs nearly obligatorily require an overt pronoun on the "extraction-site" (70% with a rejection rate of 5%) whereas this is nearly impossible with subjects in *wo*-LWDs. The results of this test sentence reproduces nicely a similar result, asked in an earlier questionnaire. There, we didn't head for LWDs but rather what is called *long relativization*; the sentence is again given in English wording:

(15) This is the man [dass/wo I know [dass/wo (he) lives in D.]]

Table 13.3: Acceptance of gap/resumptive in long relativization (two clause boundaries)

Variation of comps	Acceptance (1–2)	Complete rejection (5)
wo…dass…gap	12%	50%
wo…dass…resumptive	87%	3%
wo…wo…gap	44%	19%
wo…wo…resumptive	5%	61%

The results are presented in Table 13.3.

The same template was used for long relativization of a dative argument and here, the acceptance of the resumptive in the *wo…wo*-configuration showed essentially the same result as with simple relativization, namely about 18% – whereas the *dass*-complement clause yielded a result of 83% acceptance for the dative resumptive.

These results are more interesting than the ones from the simple LWDs – since they show that the acceptance of a resumptive is not dependent on distance but rather on the choice of the complementizer. Note that in Table 13.2, all variants with a gap reach a result of only 30%. However, in the case of a *dass*-LWD, the sentence can be saved by inserting the resumptive (by a rejection rate of 5%). This possibility is essentially excluded for *wo*-LWDs.

3.4 Resumptive pronouns in different shapes

A final piece of evidence for the idea that the "extraction out of *dass*-clauses" is maybe an illusion comes from the type of pronoun used as a resumptive. In these test-sentences, we didn't offer the "usual resumptive pronoun", namely the simple personal pronoun as the least marked ones available in Alemannic, see Adger (2011) for discussion, but a pronoun of the *d*-series:

(16) simple pronouns: er – (s)ie – es; *d*-series: d-er – d-ie – d-as

The *d*-series pronouns normally force a disjoint reference interpretation in a binding configuration across a clause-boundary (Wiltschko 1998):

(17) German
Hans$_i$ glaubt, dass er$_{i/j}$ / der$_{*i/j}$ der Beste ist
Hans believes that he *d*-series the best (one) is

Anecdotal observations about a much higher rate of d-pronouns in Alemannic lead us to the idea to test systematically the acceptance of these pronouns as resumptives. And indeed, although the acceptance rate is by far lower than with personal resumptives, it is remarkable that they show up to a much higher degree in *dass*-LWDs, namely 35% acceptance – but only 15% with *wo*-LWDs, This difference in acceptance co-varying with the choice of the complementizer again hints at the conclusion that a *dass*-clause is more encapsulated with respect to its syntactic surrounding as a *wo*-clause, strengthening the idea that it is a full-fledged clause – even if construed with an LWD.[12]

3.5 Resumptives in Celtic

What I left out until now is a discussion of resumptives in the Celtic languages. As discussed in McCloskey's work, Irish exhibits two types of RCI, traditionally named aL and aN. While aL never allows resumptives in RCs, aN requires them. A classical example is given below:

(18) Irish

a. an ghirseach a ghoid na síogaí
the girl aL stole the fairies ___

b. an ghirseach a-r ghoid na síogaí í
the girl aN-PST stole the fairies her

As can be seen, the RCI requiring the resumptive has the tense morpheme attached to it, indicating that it occupies a different, probably lower position in the functional extension of the clause, i.e. closer to Tense, see also Roberts (2005) for such an assumption. Without committing myself to a detailed account in terms of a split C-projection in a Rizzi (1997)-style, it is of course striking that aN shows the same behavior as the complementizer *go* – which also combines with the tense morpheme, yielding these different forms shown above (*gu-r*, *gu-n*, etc. depending on the variant). Clearly, these pattern with the *dass*-LWDs in Alemannic whereas *wo* in Alemannic is the direct parallel to aL.

This would mean then that Alemannic *wo* and Irish *aL* are genuine complementizers – whereas *dass/that* are indeed relative pronouns with the head consisting of a possibly silent correlate pronoun, cf. the structure given in (4′). This then implies that a complement clause introduced by *dass/that/go* is always an island and that the seemingly extraction is not extraction at all. The data discussed here favor such an analysis.

[12]Clearly, the impossibility of binding of the d-pronoun in (17) must then find a different interpretation, see van Kampen (2012) for further observations with respect to these pronouns – where they can even act in some cases as bound variables.

The reason that there is no way in Celtic to build a LWD with a *go*-clause – in contrast to an Alemannic *dass*-LWD – has probably to do with the fact that *go* is originally a preposition (see Braesicke 2019; Elliot Lash, p.c.). As such, its "clausal complement" has probably still a nominal core in it and is thus an island for independent reasons. Furthermore, Celtic has to my knowledge never shown an RC-formation strategy using pronouns. In contrast, in Germanic (and also Alemannic) RCs can be built with pronouns – and indeed – if not used as an aboutness relative and thus a complement clause, as I suggested above, cf. footnote 3, it can occur with a clause-internal gap. Thus, this is a pattern which is encountered in Germanic – but not in Celtic:

(19) German
das Buch, das du ___ gelesen hast, …
the book that you read have
'The book that you've read, …'

The exact details have to be worked out in future work – but the difference in building clausal complements and relative clauses in Germanic in Celtic must be the clue to understand the different behavior when it comes to LWDs. Alemannic is interesting as it has both strategies at its disposal for building RCs and LWDs and the difference in behavior concerning resumptives shows that there are deep syntactic differences between these structures.

4 Conclusion and outlook

I started with taking seriously the doubts on *dass* as having been re-categorized to the word-class of complementizer (and with it its head-status, resp. belonging to the extended projection of the verb). I asked which kind of evidence could be relevant to show whether *dass* is still what it looks like, namely a *d*-series pronoun, resp. a relative pronoun, implying that the complement clause is essentially a relative clause, as assumed in Kayne (2014). The consequence of this view is that complement clauses introduced by *dass* should be opaque to extraction. And indeed, I showed that the unexpected high acceptance rates of resumptive pronouns hint to the conclusion that all arguments in these embedded clauses are syntactically present as pronouns in LWDS. However, they may be subject to a rather "weak" principle like the avoid pronoun principle in being merely not pronounced if too close to the antecedent. This was contrasted with constructions containing a genuine gap, coming into existence via a relative clause formation strategy involving a specialized particle, requiring a gap in its clausal

complement and thus resumptives are essentially not possible – besides in those cases where they appear also in relative clauses – for reasons that I did not discuss here. If this is on the right track, it may have far reaching consequences for a whole bunch of assumptions about the cyclic nature of movement (re-merge). What it essentially means is that there is no cross-clausal movement at all. In light of the idea that re-merge should obey the extension condition in a strict way, this is a welcome result – since long cyclic-successive movement is until now the problematic exception to this condition.

The task for the future will then be to find more languages of the Alemannic type to see whether the correlations outlined in §3.4 hold as well. The Scandinavian languages that allow LWDs with *som* immediately come to mind. Another area of investigation would be the wh-in-situ languages which have LWDs but arguably no clause-internal wh-movement. A base generation approach together with maybe different licensing conditions for gaps/resumptives could shed new light on these long standing issues in generative syntax.

Abbreviations

CCI	complement clause introducer
ECP	empty category principle
LWD	long wh-dependency
OHG	Old High German
PF	phonetic form
PRED	predicative
PROG	progressive
PST	past
RC	relative clause
RCI	relative clause introducer
REL	relative
VN	verbal noun

References

Adger, David. 2011. Bare resumptives. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces*, 343–366. Amsterdam: John Benjamins. DOI: 10.1075/lfab.5.09adg.

Adger, David & Gillian Ramchand. 2005. Merge and move: Wh-dependencies revisited. *Linguistic Inquiry* 36(2). 161–193. DOI: 10.1162/0024389053710729.

Axel, Katrin. 2009. Die Entstehung des *dass*-Satzes: Ein neues Szenario. In Veronika Ehrich, Christian Fortmann, Ingo Reich & Marga Reis (eds.), *Koordination und Subordination im Deutschen*. Linguistische Berichte Sonderheft 16, 21–42.

Axel, Katrin. 2017. The development oft he declarative complementizer in German. *Language* 93(2). 29–65.

Braesicke, Lars. 2019. Gramadach na Gaeilge [Irish grammar]. http://nualeargais.ie/gnag/gram.htm (30 January, 2019).

Brandner, Ellen & Claudia Bucheli. 2018. Über lange w-Extraktion im Alemannischen. In Augustin Speyer & Philipp Rauth (eds.), *Syntax aus Saarbrücker Sicht 2*. Sonderheft der Zeitschrift für Dialektologie und Linguistik, 31–69.

Chao, Wynn & Peter Sells. 1983. On the interpretation of resumptive pronouns. In *NELS 13: Proceedings of the Thirteenth Annual Meeting of the North East Linguistic Society*, 47–61. Amherst, MA: GLSA Publications.

Cheng, Lisa Lai-shen & Rint Sybesma. 2005. A Chinese relative. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinheinz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, 69–76. Berlin: De Gruyter.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Engdahl, Elisabet. 1985. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23(1). 3–44. DOI: 10.1515/ling.1985.23.1.3.

Kayne, Richard S. 2014. Why isn't *this* a complementizer? In Peter Svenonius (ed.), *Functional structure from top to toe* (The Cartography of Syntactic Structures 9), 188–231. Oxford: Oxford University Press.

Manzini, M. Rita & Leonardo M. Savoia. 2003. The nature of complementizers. *Rivista di Grammatica Generativa* 28. 87-110.

Manzini, M. Rita & Leonardo M. Savoia. 2011. *Grammatical categories: Variation in Romance languages*. Cambridge: Cambridge University Press.

McCloskey, James. 2001. The morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37. 67–100. DOI: 10.1017/S0022226701008775.

McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and explanation in the minimalist program*, 184–226. Malden, MA: Blackwell.

Müller, Gertraud & Theodor Frings. 1959. *Die Entstehung der deutschen daß-Sätze*. Akademie-Verlag.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Springer.

Roberts, Ian. 2005. *Principles and parameters in a VSO language: A case study in Welsh*. Oxford: Oxford University Press.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.

Salzmann, Martin. 2006. *A study in indirect A′-dependencies*. Utrecht: LOT.

van Gelderen, Elly. 2004. *Grammaticalization as economy*. Amsterdam: John Benjamins. DOI: 10.1075/la.71.

van Kampen, Jacqueline. 2012. Obviation properties of the Dutch *d*-pronoun. *Linguistics in the Netherlands* 29. 68–82.

van Riemsdijk, Henk C. 2003. East meets West. In Jan Koster & Henk C. van Riemsdijk (eds.), *Germania et alia: A linguistic webschrift for Hans den Besten*. http://odur.let.rug.nl/%5Ctextasciitilde%7B%7Dkoster/DenBesten/contents.htm.

Westergaard, Marit, Øystein A. Vangsnes & Terje Lohndal. 2012. Norwegian *som*: The complementizer that climbed to the matrix left periphery and caused Verb Second violations. In Valentina Bianchi & Cristiano Chesi (eds.), *Enjoy linguistics!: Papers offered to Luigi Rizzi on the occasion of his 60th birthday*, 329–343. Siena: CISCL Press. http://www.ciscl.unisi.it/gg60/papers/westergaard_et_al.pdf.

Willis, David. 2000. On the distribution of resumptive pronouns and wh-trace in Welsh. 36(3). Journal of Linguistics, 531–573. DOI: 10.1017/S0022226700008380.

Wiltschko, Martina. 1998. On the syntax and semantics of (relative) pronouns and determiners. *The Journal of Comparative Germanic Linguistics* 2(2). 143–181. DOI: 10.1023/A:1009719229992.

Part II

Structural issues in morphosyntax

Chapter 14

Types of relative pronouns

Evangelia Daskalaki
University of Alberta

In this paper, I explore the possibility that relative pronouns, like personal pronouns, show different degrees of strength/deficiency. I show that, at least in Greek, the restricted relative (RR) pronoun *o opios* is semantically deficient compared to its free relative (FR) counterpart *opjos* in two interrelated respects: (i) it is referentially deficient and (ii) it does not license its own range. After showing that both FR and RR pronouns behave like transitive Ds, I propose that their differences lie in their featural composition, rather than in their structural make-up: FR determiners, unlike RR determiners, are semantically definite.

1 Introduction

That pronouns may show a different cluster of properties – diachronically, synchronically, and cross-linguistically – is a well-established fact in the literature. Existing accounts, focusing primarily on the different classes of personal pronouns, suggest two main lines of approach.[1] The first one attributes the different properties of (personal) pronouns to their external category (Cardinaletti & Starke 1999; Déchaine & Wiltschko 2002). The second type of analyses treats all pronouns as determiners projecting a DP and derives their differences from their internal structure and/or featural composition (Abney 1987; Cardinaletti 1994; Uriagereka 1995; among others).

The aim of this paper is to explore whether similar claims can be made for the class of relative pronouns.[2] I argue that, at least in Greek, RR pronouns can

[1]For a detailed overview and application to personal pronouns in Greek, see Mavrogiorgos (2010).

[2]See also Sportiche (2011) for French restrictive relative pronouns, and Wiltschko (1998) for German restrictive relative pronouns.

Evangelia Daskalaki. 2020. Types of relative pronouns. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 277–296. Berlin: Language Science Press.
DOI: 10.5281/zenodo.4280653

be shown to be semantically deficient compared to FR pronouns in two (interrelated) respects: (i) RR pronouns are not inherently definite/referential, and (ii) RR pronouns do not license their own range. After showing that both FR and RR pronouns behave like transitive Ds, and are therefore categorially equivalent, I propose that their differences derive from their featural composition: FR determiners, unlike RR determiners, are semantically definite/referential. Because they are definite/referential determiners, they need a range that may take the form of a lexical NP complement or of an animacy restrictor.

The paper is structured as follows: §2 provides some background information concerning (Greek) relative clauses and pronouns. §3 establishes at an empirical level the semantic deficiency of RR pronouns and §4 develops an analysis that capitalizes on the featural composition of the FR and RR D head. Finally, §5 concludes the discussion.

2 Background information on relative clauses and pronouns

2.1 (Greek) relative clauses

Restrictive and free relatives are A′ movement dependencies with different functions. Whereas *restrictive relatives* function as modifiers of nominal heads, *free relatives* function as arguments/adjuncts of lexical predicates (Alexiadou et al. 2000; Bianchi 2002; Grosu & Landman 1998). This is illustrated below with Greek:[3]

(1) Greek
ðjaleksa tus maθites$_i$ [tus opius$_i$ protines t$_i$].
chose.1SG the students.M.PL.ACC which.M.PL.ACC recommended.2SG
'I chose the students who you recommended.'

(2) Greek
ðjaleksa [opjus$_i$ protines t$_i$].
chose.1SG who.M.PL.ACC recommended.2SG
'I chose who you recommended.'

In (1), the RR modifies the nominal head *maθites* 'students'. In (2), the FR complements the verbal head *ðjaleksa* 'chose'.

As far as their semantic interpretation is concerned, FRs in DP position are semantically equivalent with strong DPs (Jacobson 1995). For instance, the FR in (2) can be paraphrased with an RR headed by a demonstrative (3):

[3]On Greek RRs see Alexopoulou (2006); on Greek FRs see Alexiadou & Varlokosta (1997).

(3) Greek
ðjaleksa [aftus [tus opius$_i$ protines t_i]].
chose.1SG those.M.PL.ACC which.M.PL.ACC recommended.2SG
'I chose those ones you recommended.'

2.2 (Greek) relative pronouns

With respect to restrictive and free relative pronouns, languages differ as to whether they draw them from the same paradigm. Thus, English draws both RR and FR pronouns from the paradigm of interrogative pronouns. German, on the other hand, uses interrogative pronouns to introduce FRs and morphologically definite determiners to introduce RRs (Wiltschko 1998).

Greek stands somewhere in between: RR and FR pronouns are similar in that they both combine interrogative and definite morphology.[4] However, they are not identical and replacing one with the other leads to strong ungrammaticality:

(4) Greek
* ðjaleksa tus maθites$_i$ [opjus$_i$ protines t_i].
chose.1SG the students.ACC who.ACC recommended.2SG
* 'I chose the students whoever you recommended.'

(5) Greek
* ðjaleksa [tus opius$_i$ protines t_i].
chose.1SG which.ACC recommended.2SG
* 'I chose which you recommended.'

Furthermore, both types of pronouns are inflected for the same range of categories. Thus, they inflect for number (singular, plural), gender (masculine, feminine, neuter), and case (nominative, accusative, genitive), displaying in this respect the main features characterizing Greek nominal inflection. The complete morphological paradigm of *opjos* and *o opios* is provided in Tables 14.1 and 14.2, respectively (Holton et al. 2004: 100).

[4]Thus, the RR pronoun *o opios* consists of the morphologically definite determiner *o* and the word *opios*. The latter, being itself complex, can be decomposed into the determiner-like prefix *o-* and the interrogative *pios* 'who' (on the morphological decomposition of the RR *o opios*, see Alexiadou 1998). A similar pattern is shown by the FR pronoun *opjos*. Like its RR counterpart, it is a complex word, consisting of the determiner-like prefix *o-* and the interrogative *pjos* 'who'. Unlike its RR counterpart though, it is not introduced by a free determiner (on the etymological decomposition of the FR *opjos*, see Chila-Markopoulou 1994).

Table 14.1: The morphological paradigm of the FR pronoun *opjos-a-o*

	Singular			Plural		
	Masc	Fem	Neut	Masc	Fem	Neut
NOM	*opjos*	*opja*	*opjo*	*opji*	*opjes*	*opja*
ACC	*opjon*	*opja(n)*	*opju*	*opjus*	*opjes*	*opja*
GEN	*opju*	*opjas*	*opjo*	*opjon*	*opjon*	*opjon*

Table 14.2: The morphological paradigm of the RR pronoun *o opios-i opia-to opio*

	Singular			Plural		
	Masc	Fem	Neut	Masc	Fem	Neut
NOM	*o opios*	*i opia*	*to opio*	*i opii*	*i opies*	*ta opia*
ACC	*ton opio*	*tin opia*	*tu opiu*	*tus opius*	*tis opies*	*ta opia*
GEN	*tu opju*	*tis opias*	*to opio*	*ton opion*	*ton opion*	*ton opion*

3 On the deficiency of RR pronouns

Despite being amenable to a similar etymological decomposition and despite being marked for the same range of morphological features, RR pronouns can be shown to be deficient compared to their FR counterparts in a number of ways that recall the differences identified between strong and weak personal pronouns. Let us consider them in turn.

3.1 Contrastive focus

To begin with, only FR pronouns may bear contrastive focus. This is shown by the contrast in grammaticality between (6) and (7).

(6) Greek
kalese mono *opjus* tu protines oxi opjes
invited.3SG only who.M.PL CL.3SG.M recommended.2SG not who.F.PL
tu protines
CL.3SG.M recommended.2SG
'He only invited whichever *men* you recommended to him, not whichever women you recommended to him.'

(7) Greek
*kalese mono aftus *tus opius* tu protines
invited.3SG only those.M.PL.ACC which.M.PL CL.3SG recommended.2SG
oxi aftes tis opies tu protines.
not those.F.PL.ACC which.F.PL CL.3SG recommended.2SG
*'He only invited those men *who* you recommended, not those women who you recommended.'

Thus, in (6), the FR pronoun *opjus*, encoding masculine gender, can be contrastively focused with the FR pronoun *opjes*, encoding feminine gender. Crucially, in the same contrastive configuration, the RR pronoun *tus opius* is not permissible with contrastive stress (7).[5]

3.2 Null counterparts

Secondly, only FR pronouns are obligatorily realized (Alexiadou et al. 2000: 22). To this end, example (8) shows that replacing a FR pronoun with the uninflected complementizer *pu* 'that' leads to strong ungrammaticality:

(8) Greek
a. ðjaleksa [opjus protines].
chose.1SG who.ACC recommended.2SG
'I chose whoever you recommended.'

[5]The English translation of (6) and (7) in the main text fails to convey the contrast between FR and RR pronouns with respect to focus. This is because English relative pronouns do not encode gender distinctions (that is *who* can be used to refer to both female and male entities). The same effect, though, can be conveyed with the English FR pronouns *who* (a FR pronoun used for animate entities) and *what* (a FR pronoun used for inanimate entities).

(i) Greek
*ðen θeli mono afta ta opia exis ala ke aftus
NEG want.3SG only those.N.PL.ACC which.N.PL have.2SG but and those.M.PL.ACC
tus opius exis.
who.M.PL have.2SG
intended: 'He doesn't only want those (things) which you have, but also those (persons) *who* you have.'

(ii) Greek
*ðen θeli mono oti exis ala ke *opjon* exis.
NEG want.3SG only what have.2SG but and who.M.SG have.2SG
'He doesn't only want what you have but also *who* you have.'

b. * ðjaleksa [pu protines].
chose.1SG that recommended.2SG
*'I chose that you recommended.'

By contrast, complementizer RRs (9b) are a very common alternative to pronominal RRs (9a) in Greek and in other languages:

(9) Greek
a. ðjaleksa aftus [tus opius protines].
chose.1SG those.M.PL.ACC which recommended.2SG
'I chose those ones who you recommended.'
b. ðjaleksa aftus [pu protines].
chose.1SG those.M.PL.ACC that recommended.2SG
'I chose those ones that you recommended.'

3.3 Animacy

Furthermore, only FR pronouns appear to license an animacy restriction.

Thus, FR pronouns marked for masculine/feminine gender licence by default a [+animate] interpretation, whereas FR pronouns marked for neuter gender license a [−animate] interpretation. For example, the masculine FR pronoun *opjus* in (10a), under its more natural interpretation, refers to a male animate entity, whereas the neuter FR *opja* in (10b), evokes a [−animate] entity.

(10) Greek
a. ðjaleksa opjus protines.
chose.1SG who.M.PL recommended.2SG
'I chose who you recommended.'
b. ðjaleksa opja protines.
chose.1SG what.N.PL recommended.2SG
'I chose what you recommended'

A similar point is made by the minimal pair in (11): whereas the neuter FR pronoun *opjo* is perfectly grammatical as the subject of verbs that typically take thematic/inanimate subjects (11a), it sounds awkward, when it occupies the subject position of verbs that typically require agentive/animate subjects (11b).

(11) Greek
a. opjo espase
what.N.SG broke.3SG
'What(ever) broke.'

b. ## opjo eɣrapse tin epistoli
what.N.SG wrote.3SG the letter.ACC
##'What(ever) wrote the letter.'

The distribution of RR pronouns, on the other hand, does not appear to be regulated by animacy considerations. To illustrate, RR pronouns are admissible with both animate and inanimate antecedents, independently of whether they are marked for masculine (12) or neuter gender (13).

(12) Greek
a. ðjaleksa tus maθites tus opius protines.
chose.1SG the students.ACC which.M.PL recommended.2SG
'I chose the students who you recommended.'
b. ðjaleksa tus pinakes tus opius protines.
chose.1SG the paintings.ACC which.M.PL recommended.2SG
'I chose the paintings which you recommended.'

(13) Greek
a. ðjaleksa ta peðja ta opia protines.
chose.1SG the kids.ACC which.N.PL recommended.2SG
'I chose the kids who you recommended.'
b. ðjaleksa ta pexniða ta opia protines.
chose.1SG the toys.ACC which.N.PL recommended.2SG
'I chose the toys which you recommended.'

3.4 Referentiality

A further difference between FR and RR pronouns concerns their ability to introduce new referents. Consider in this regard the examples in (14) illustrating coordination of FRs:

(14) Greek
a. kalesa opjon simbaθi i Maria ke opjon adipaθi i Lina.
invited.1SG who.ACC like.3SG the Maria.NOM and who.ACC dislike.3SG the Lina.NOM
'I invited whoever Maria likes and whoever Lina dislikes.'
[✓ Maria likes X & Lina dislikes Y; ✓ Maria likes X & Lina dislikes X]

b. kalesa opjon simbaθi i Maria ke adipaθi i
invited.1SG who.ACC like.3SG the Maria.NOM and dislike.3SG the
Lina.
Lina.NOM
'I invited whoever Maria likes and Lina dislikes.'
[*Maria likes X & Lina dislikes Y; ✓ Maria likes X & Lina dislikes X]

When coordination takes place at the FR pronoun level, the coordinated phrases may either refer to two distinct discourse referents or to a single participant (14a). Of the two possible readings, the first one is the preferred one. However, when coordination takes place below the FR pronoun, the coordinated phrases may only refer to a single participant (14b). In other words, there appears to be a correlation between the number of FR pronouns and the number of referents.[6]

The correlation between number of pronouns and number of referents is not replicated by RRs:

(15) Greek
a. kalesa afton ton opio simbaθi i Maria ke
invited.1SG this.one.ACC which.ACC like.3SG the Maria.NOM and
ton opio adipaθi i Lina.
which.ACC dislike.3SG the Lina.NOM
'I invited this one who Maria likes and who Lina dislikes.'
[*Maria likes X & Lina dislikes Y; ✓ Maria likes X & Lina dislikes X]

[6]In this respect the FR pronoun *opjos* behaves like the definite determiner *o* 'the' in argumental DPs. Alexiadou et al. (2007: 67–68), replicating a point originally made by Longobardi (1994) for Italian, show that there appears to be a correlation between the number of definite determiners in coordinated DPs and the number of referents. Thus, whereas there is only one referent in (i), there are two referents in (ii):

(i) Greek
irθ-e/*-an o antiprosopos tis dikastikis arxis ke proedros tis eforeftikis
came-3SG/PL the delegate of.the court and chair of.the elective
epitropis.
committee
'The representative of the court and chair of the elective committee have arrived.'

(ii) Greek
irθ-an/*-e o antiprosopos tis dikastikis arxis ke o proedros tis eforeftikis
came-3PL/SG the delegate of.the court and the chair of.the elective
epitropis.
committee
'The representative of the court and the chair of the elective committee has arrived.'

b. kalesa afton ton opio simbaθi i Maria ke
invited.1SG this.one.ACC which.ACC like.3SG the Maria.NOM and
ton opio adipaθi i Lina.
which.ACC dislike.3SG the Lina.NOM
'I invited this one who Maria likes and Lina dislikes.'
[*Maria likes X & Lina dislikes Y; ✓ Maria likes X & Lina dislikes X]

What the above examples serve to show is that multiple occurrences of an RR pronoun do not produce a multiple index interpretation.

3.5 Overt NP complement

Finally, only FR pronouns may licence overt NP complements. This is shown by the contrast in grammaticality between (16) and (17):[7]

(16) Greek
ðjaleksa opjus (ipopsifius) protines.
chose.1SG who.ACC candidates recommended.2SG
'I chose whichever candidates you recommended.'

(17) Greek
a. * ðjaleksa tus opius ipopsifius protines.
chose.1SG which.ACC candidates recommended.2SG
*'I chose which candidates you recommended.'
b. * ðjaleksa tus ipopsifius tus opius ipopsifius protines.
chose.1SG the candidates which.ACC candidates recommended.2SG
*'I chose the candidates which candidates you recommended.'

Crucially, FR pronouns with overt NP complements (complex FR pronouns, henceforth) differ from the simple FR pronouns discussed so far, in two respects: First, they cannot bear contrastive stress. In instances of contrastive focus it is their complement that is focused (18):

[7]It is only in appositive relatives that *o opios* may take an overt NP complement:

(i) Greek
to computer, to opio computer epemenes na aɣoraso, ðen ðulevi.
the computer, which computer insisted.2SG SBJV buy.1SG NEG work.3SG
'The computer, which you insisted that I buy, is not working.'

(18) Greek

a. * kalese mono *opjus* maθites tu
invited.3SG only which.M.PL students.M.PL CL.3SG.M
protines oxi opjes maθitries tu
recommended.2SG not which.F.PL students.F.PL CL.3SG.M
protines.
recommended.2SG
intended: 'He only invited *whichever* male students you recommended to him, not whichever female students you recommended to him.'

b. kalese mono opjus *maθites* tu protines
invited.3SG only which students.M.PL CL.3SG.M recommended.2SG
oxi opjes maθitries tu protines.
not which students.F.PL CL.3SG.M recommended.2SG
'He only invited whichever *male* students you recommended to him, not whichever female students you recommended to him.'

Second, they may take both animate and inanimate complements, independently of whether they are marked for masculine/feminine gender, as in (19), or for neuter gender, as in (20):

(19) Greek

a. ðjaleksa opjus maθites protines.
chose.1SG which.M.PL students.ACC recommended.2SG
'I chose whichever students you recommended.'

b. ðjaleksa opjus pinakes protines.
chose.1SG which.M.PL paintings.ACC recommended.2SG
'I chose whichever paintings you recommended.'

(20) Greek

a. ðjaleksa opja peðja protines.
chose.1SG which.N.PL kids.ACC recommended.2SG
'I chose whichever kids you recommended.'

b. ðjaleksa opja pexniða protines.
chose.1SG which.N.PL toys.ACC recommended.2SG
'I chose whichever toys you recommended.'

3.6 Summary

A schematic summary of the differences between restrictive and free relative pronouns (simple and complex) is provided in Table 14.3.

Table 14.3: The properties of RR and FR pronouns.

	Pronouns		
	FR (simple)	FR (complex)	RR
Contrastive focus	Yes	No[a]	No
Null counterparts	No	No	Yes
Animacy	Yes	No	No
Disjoint reference under conjunction	Yes	Yes	No
Overt NP complement	No	Yes	No

[a](only their complement)

The list of differences between free and restrictive relative pronouns can be narrowed down into two main points of divergence:

1. FR pronouns (simple/complex), unlike RR pronouns, are referential. This explains the correlation between the number of FR pronouns and the number of referents (14), a correlation that does not hold in the case of RR pronouns (15).

2. FR pronouns (simple/complex), unlike RR pronouns, may license their own range. The range may take the form of an animacy restriction licensed by the FR pronoun (10) and (11) (in the case of simple FR pronouns), or the form of a lexical NP complementing the FR pronoun (19) and (20) (in the case of complex FR pronouns). This explains why simple FR pronouns can be contrastively focused. Being inherently specified as [+animate] or [−animate], they can bear contrastive focus with respect to animacy (6). RR pronouns, on the other hand, not being specified for animacy cannot bear contrastive focus for a property they lack (7).

Under this view, FR pronouns lack null counterparts because their deletion would result in unrecoverable loss of both referentiality and range (8).

4 Towards an analysis

Having established at an empirical level that RR pronouns are deficient compared to FR pronouns, I will now consider the question of theoretical implementation. After showing that both types of pronouns are transitive determiners (§4.1 and §4.2), I will suggest that their differences lie in their featural composition: whereas both RR and FR determiners are morphologically definite, only the latter ones are semantically definite (§4.3).

4.1 Both free and restrictive relative pronouns are DPs

It is possible that the referential deficiency of RR pronouns is reflective of a kind of structural deficiency. Thus, adopting and adapting Déchaine & Wiltschko's (2002) account of personal pronouns, we could assume that whereas FR pronouns are Ds projecting a DP, RR pronouns are the mere spell out of phi features (phi Ps). Within this approach, RR pronouns fail to refer because they lack an external D layer, which is typically taken to be the locus of definiteness/referentiality.

There are two main issues with this approach. First, as mentioned in the introduction, both free and restrictive relative pronouns incorporate a morphologically definite determiner (*o* 'the'). Thus, morphological considerations suggest that they are both Ds. The second issue is syntactic in nature and concerns their distribution. Even though both pronouns surface in [Spec,CP], they can be theta related to all the major argument positions, including the subject of (in)transitive verbs, the subject of primary and secondary predication, the (in)direct object, and the prepositional object position. The latter is illustrated in (21) and (22) with a FR and RR pronoun, respectively:

(21) Greek
jia opjus (maθites / pinakes) mu milises
about which students paintings CL.1SG.GEN talked.2SG
'About whichever (students/paintings) you talked to me.'

(22) Greek
o maθitis / pinakas jia ton opio mu milises
the student.NOM painting.NOM about who.PL.ACC CL.1SG.GEN talked.2SG
'The student/painting about whom you talked to me.'

On the assumption that argumenthood is a property of DPs (Longobardi 1994), it follows that both *opjos*-phrases and *o opios* phrases are associated with a DP projection.

4.2 Both free and restrictive relative pronouns are transitive Ds

Furthermore, it can be argued that in addition to showing the external distribution of DPs, both types of pronouns show the internal syntax of determiners. Complex FR pronouns clearly behave like transitive determiners, since they allow an NP complement. The latter can be overt, as in (23) repeated from (16) above, or elided under identity with a discourse antecedent, as in (24).[8]

(23) Greek
kalesa [opjus ipopsifius protines].
invited.1SG who.ACC candidates recommended.2SG
'I invited whichever candidates you recommended.'

(24) Greek
a. pjus ipopsifius kaleses?
which candidates invited.2SG
'Which candidates did you invite?'
b. opjus mu protines.
who CL.1SG.GEN recommended.2SG
'Whoever you recommended to me.'

In the absence of a salient discourse antecedent, we saw that FR pronouns (simple FR pronouns in our terms) receive a [±animate] interpretation, depending on their gender specification (10–11). One way to implement this observation is to assume that they bear interpretable phi features that are responsible for licensing a null complement. Thus, an interpretable masculine/feminine gender licenses an empty [+animate] NP complement, whereas an interpretable neuter gender licenses a [−animate] NP complement. Within this account, the difference

[8]Evidence suggesting that the FR pronoun in (24b) is a transitive determiner with a deleted NP restrictor comes from its similarities with other instances of nominal subdeletion attested in Greek, such as the one illustrated in (i):

(i) Greek
a. pja fusta aɣorases?
which.F.SG skirt bought.2SG
'Which skirt did you buy?'
b. tin kokini.
the.F.SG red.F.SG

In this regard, see Daskalaki (2009) who shows how the conditions on nominal subdeletion identified by Giannakidou & Stavrou (1999) can be replicated for FR phrases.

between complex and simple FR pronouns does not lie in their (in)transitivity. Rather it depends on whether the FR determiner has entered the derivation with an uninterpretable set of phi features (that will be valued by an overt lexical NP) or with an interpretable set of phi features that is responsible for licensing a null, [±animate] NP complement.[9]

Let us, finally, consider the RR pronoun *o opios*. At a first approximation its treatment as a transitive determiner seems implausible, given that, at least in its restrictive use, it never surfaces with an overt NP complement (17). However, this would be incompatible with both the *raising analysis* (Kayne 1994; for Greek RRs, see Alexiadou & Anagnostopoulou 2000, among others) and the *matching analysis* (Sauerland 1998; for Greek RRs, see Kotzoglou & Varlokosta 2005 of relative clauses). Motivated by independent considerations, such as reconstruction effects, both analyses maintain the claim that the RR pronoun is a determiner taking an NP complement. In the case of the raising analysis, the NP complement is raised to the antecedent position, whereas in the case of the matching analysis it is deleted under identity with an externally Merged antecedent.[10] In view of these independent considerations, I will be assuming that RR pronouns, like FR pronouns, are transitive determiners.[11]

4.3 RR pronouns, unlike FR pronouns, have an expletive D head

If both FR and RR pronouns are transitive Ds, then the referential deficiency of RR pronouns cannot be treated as an instance of structural deficiency. A conceivable alternative would be to treat it as an instance of featural deficiency. Under this view, the difference between FR and RR pronouns depends on whether their D head is semantically definite/referential, as in the case of FR pronouns, or semantically inert, as in the case of RR pronouns.

[9] Alternatively, it could be the case that the phi features of the FR determiner are always uninterpretable. In the case of complex FR pronouns they get valued through agreement with an overt lexical NP, whereas in the case of simple FR pronouns they get valued through agreement with the gender specification of a null NP meaning 'man', 'woman', or 'thing'. An analysis along these lines would be compatible with Panagiotidis (2003) and would allow us to treat homogeneously complex and (apparently) simple FR pronouns. However, it is not clear how it would derive the contrast between the two types of FR pronouns with respect to contrastive focus. In other words, if both simple and complex FR pronouns bear uninterpretable gender it is not clear why only the former ones can bear contrastive focus (compare (6) with (18)).

[10] Thanks to an anonymous reviewer for pointing this out to me.

[11] Within this analysis, (17a) is ungrammatical not because there is no NP position projected in syntax, but because the RR determiner, being expletive (see §4.3) cannot introduce a clause that functions as an argument. Accordingly, (17b) is ungrammatical because due to some economy consideration the complement of the RR determiner needs to be deleted under identity with a c-commanding antecedent.

That the definite morphology of RR pronouns is void of any semantic contribution is not a novel claim (see, among others, Bianchi 1999: 80; for Greek, see Alexiadou 1998). Independent evidence in support of this analysis comes from the expletive uses of the Greek definite determiner in contexts other than RRs. Consider, for example, the phenomenon of polydefiniteness, illustrated in (25):

(25) Greek
to spiti to megalo
the house the big
'the big house'

In (25), a noun (*spiti* 'house') is modified by an adjective (*megalo* 'big'), and noun and adjective are each introduced by a morphologically definite determiner (*to* 'the'). Despite the multiple occurrences of the definite article, the construction does not receive a multiple reference interpretation. Thus, (25) refers to a single entity at the intersection of the set of houses and the set of big entities (Lekakou & Szendrői 2012). This fact has been taken to show that the definite determiner in Greek, at least in some contexts, can be used as an expletive (for an overview of the proposed analyses, see Alexiadou 2014). It is this claim that we reiterate here for the RR determiner.

Our second claim, that the FR pronoun encodes definiteness/referentiality, has been more controversial in the literature. Recall from §2.1 that FRs can be paraphrased with definite DPs. One group of analyses derives the referentiality/definiteness of FRs from the referentiality/definiteness of FR pronouns (see, for instance, Jacobson 1995 and Pancheva 2000, among others). A different group of analyses suggests that the reason why FRs are interpreted like definite DPs is because of a null c-commanding determiner/element that turns them into referential expressions (Groos & van Riemsdijk 1981; Caponigro 2003; Grosu & Landman 1998, among others).

One of the main semantic arguments in favor of the null D analysis is that many languages use the same range of relative pronouns both in definite FRs and in irrealis FRs (Caponigro 2003). Irrealis FRs differ from definite FRs in a number of ways (Caponigro 2003; Pancheva 2000; Grosu & Landman 1998): Irrealis FRs always complement existential predicates (mainly the existential *have* or *be*), they include irrealis verbal morphology, and, crucially, they cannot be paraphrased by definite DPs. Rather they appear to be semantically equivalent with weak NPs. As an illustrative example, we may consider the Polish examples below, illustrating a standard and an irrealis FR, respectively:

(26) Polish (Caponigro 2003: 27)
Posmakowalam [co ugotowałes].
tasted.1SG what cooked.2SG
'I tasted what you cooked.'

(27) Polish (Caponigro 2003: 88)
(Nie) mam [co robić].
not have.1SG what do.INF
'There {is something, isn't anything} I can do.'

As pointed out by Caponigro (2003), the fact that the same range of pronouns is used both in standard/definite (26) and in irrealis FRs (27) is problematic for the claim that these pronouns are inherently definite. Significantly, though, this counterargument does not apply to the Greek data. As illustrated below, FR pronouns fail to introduce irrealis FRs (28a). Rather an interrogative pronoun is used for this purpose (28b):

(28) Greek
a. * ðen exo se opjon na miliso.
NEG have.1SG to who SBJV talk.1SG
b. ðen exo se pjon na miliso.
NEG have.1SG to who SBJV talk.1SG
intended: 'I don't have anyone to talk to.'

If *opjos* is not semantically definite, it is not clear what rules out its use in (28a).

An additional challenge for the extension of the null D analysis to Greek is posed by the fact that the presumed null definite D fails to be replaced by the overt definite determiner *o* 'the' that independently exists in the language (29):

(29) Greek
*Kalese ton opjon θes.
invite.2SG the who want.2SG
*'Invite the whoever you want.'

Of course, it could be the case that the morphologically definite determiner is always expletive and that definiteness is always provided by a null c-commanding functional head.[12] Even in this case though, one would expect that *o opios* would be able to introduce a FR (when embedded under the null definite D) and that *opjos* would be able to introduce an RR (when not embedded) under the null D). As shown below, neither of the two predictions is borne out:

[12]This has actually been proposed by Lekakou & Szendrői (2012) on the basis of polydefinites.

(30) Greek
*Kalese [∅ [ton opio maθiti θes]].
invite.2SG which student want.2SG
intended: 'Invite which student you want.'

(31) Greek
*Kalese afton / ton maθiti [opjon θes].
invite.2SG this.one the student who want.2SG
*'Invite him/the student whoever you want.'

In view of the above facts, I conclude that, at least in Greek, the FR determiner, unlike the RR determiner, is semantically definite/referential.[13] Thus, whereas the RR determiner *o opios* is [−def, +rel], the FR determiner *opjos* is [+def, +rel]. Because it is semantically definite, it needs a range that is provided by its NP complement. The latter can be an overt NP, a deleted NP, or an empty NP that receives a [±animacy] interpretation.

5 Conclusions

In this paper, I explored the possibility that relative pronouns, like personal pronouns, show different degrees of strength/deficiency. I showed that, at least in Greek, the RR pronoun *o opios* is semantically deficient compared to its FR counterpart *opjos* in two interrelated respects: (i) it is referentially deficient and (ii) it does not license its own range. After showing that both FR and RR pronouns behave like transitive Ds, I proposed that their differences lie in their featural composition: FR determiners, unlike RR determiners, are semantically definite. This analysis suggests that, at least in some cases, referential deficiency can be indicative of featural rather than structural deficiency (cf. Cardinaletti & Starke 1999; Déchaine & Wiltschko 2002). Furthermore, it opens up the possibility of attributing the distribution of free and restrictive relative clauses to the properties of their introductory determiners. FR determiners, being [+def], turn a clause into a referential DP. RR determiners, on the other hand, being expletive, turn a clause into a predicate that can function as a nominal modifier. The implications of these conclusions for existing analyses of free and restrictive relatives can be the topic of future research.

[13] If this conclusion is on the right track, then it seems that the semantic import of FR pronouns could be subject to cross-linguistic variation. On the one hand, there are FR pronouns like the Greek *opjos* that may take an NP complement and encode definiteness. On the other hand, there are FR pronouns like the Polish *co* or the English *who* that may not take an NP complement, and, according to Caponigro's convincing analysis (2003), encode animacy (they are mere set restrictors).

Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
CL	clitic
F	feminine
FR	free relative
GEN	genitive
INF	infinitive
M	masculine
N	neuter
NEG	negation
NOM	nominative
PL	plural
RR	restricted relative
SBJV	subjunctive
SG	singular

Acknowledgements

In this paper, I revisit a puzzle that I briefly discussed in my PhD dissertation (University of Cambridge, 2009). I would like to thank Prof. Ian Roberts, my PhD supervisor, for all his help and support during those years. Moreover, I would like to thank two anonymous reviewers for helpful comments and suggestions. All remaining errors are, of course, my own.

References

Abney, Steven. 1987. *The English noun phrase in its sentential aspect*. MIT. (Doctoral dissertation).

Alexiadou, Artemis. 1998. On the structure of Greek relative clauses. *Studies in Greek Linguistics* 18. 15–29.

Alexiadou, Artemis. 2014. *Multiple determiners and the structure of DPs*. Amsterdam: John Benjamins. DOI: 10.1075/la.211.

Alexiadou, Artemis & Elena Anagnostopoulou. 2000. Asymmetries in the distribution of clitics: The case of Greek restrictive relatives. In Frits Beukema & Marcel den Dikken (eds.), *Clitic phenomena in European languages*, 47–70. Amsterdam: John Benjamins. DOI: 10.1075/la.30.04ale.

Alexiadou, Artemis, Liliane Haegeman & Melita Stavrou. 2007. *Noun phrase in the generative perspective*. Berlin: De Gruyter.

Alexiadou, Artemis, Paul Law, André Meinunger & Chris Wilder. 2000. Introduction. In *The syntax of relative clauses*, 1–51. Amsterdam: John Benjamins. DOI: 10.1075/la.32.01ale.

Alexiadou, Artemis & Spyridoula Varlokosta. 1997. Free relatives in Modern Greek. *Greek Linguistics* 2. Reprinted in Artemis Alexiadou (ed.), *The morpho-syntax of Greek*, 222–251. Newcastle upon Tyne: Cambridge Scholars Publishing, 405–414.

Alexopoulou, Theodora. 2006. Resumption in relative clauses. *Natural Language & Linguistic Theory* 24(1). 57–111. DOI: 10.1007/s11049-005-0898-2.

Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: De Gruyter. DOI: 10.1515/9783110803372.

Bianchi, Valentina. 2002. Headed relative clauses in generative syntax Part I. *GLOT International* 6(7). 197–204.

Caponigro, Ivano. 2003. *Free not to ask: On the semantics of free relatives and* wh*-words cross-linguistically*. UCLA. (Doctoral dissertation).

Cardinaletti, Anna. 1994. On the internal structure of pronominal DPs. *The Linguistic Review* 11(3–4). 195–219. DOI: 10.1515/tlir.1994.11.3-4.195.

Cardinaletti, Anna & Michal Starke. 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In Henk C. van Riemsdijk (ed.), *Eurotyp: Volume 5: Clitics in the languages of Europe, Part 1*, 145–234. DOI: 10.1515/9783110804010.145.

Chila-Markopoulou, Despina. 1994. Problems of diachronic syntax: Free relatives in Medieval and Modern Greek. In Irene Philippaki-Warburton, Katerina Nicolaidis & Maria Sifianou (eds.), *Themes in Greek linguistics: Papers from the first international conference on Greek linguistics, Reading, September 1993*, 413–420. Amsterdam: John Benjamins. DOI: 10.1075/cilt.117.59chi.

Daskalaki, Evangelia. 2009. *(Mis)matching patterns in Greek free relatives*. University of Cambridge. (Doctoral dissertation).

Déchaine, Rose-Marie & Martina Wiltschko. 2002. Decomposing pronouns. *Linguistic Inquiry* 33(3). 409–442. DOI: 10.1162/002438902760168554.

Giannakidou, Anastasia & Melita Stavrou. 1999. Nominalization and ellipsis. *The Linguistic Review* 16(4). 295–331. DOI: 10.1515/tlir.1999.16.4.295.

Groos, Anneke & Henk C. van Riemsdijk. 1981. Matching effects in free relatives: A parameter of core grammar. In Adriana Belletti, Luciana Brandi & Luigi Rizzi (eds.), *Theory of markedness in generative grammar: Proceedings of the 1979 GLOW conference*, 171–216. Pisa: Scuole Normale Superiore.

Grosu, Alexander & Fred Landman. 1998. Strange relatives of the third kind. *Natural Language Semantics* 6(2). 125–170. DOI: 10.1023/A:1008268401837.

Holton, David, Peter Mackridge & Irene Philippaki-Warburton. 2004. *Greek*. London: Routledge.

Jacobson, Pauline. 1995. On the quantificational force of English free relatives. In Emmon Bach, Eloisie Jelinek, Angelika Kratzer & Barbara H. Partee (eds.), *Quantification in natural languages*, 451–486. Dordrecht: Kluwer.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kotzoglou, George & Spyridoula Varlokosta. 2005. Clitics in Greek restrictive relative clauses: An integrated approach. *Reading Working Papers in Linguistics* 8. 27–49.

Lekakou, Marika & Kriszta Szendrői. 2012. Polydefinites in Greek: Ellipsis, close apposition and expletive determiners. *Journal of Linguistics* 48(1). 107–149. DOI: 10.1017/S0022226711000326.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4). 609–665.

Mavrogiorgos, Marios. 2010. Internal structure of clitics and cliticization. *Journal of Greek Linguistics* 10(1). 3–44. DOI: 10.1163/156658410X499676.

Panagiotidis, Phoevos. 2003. Empty nouns. *Natural Language & Linguistic Theory* 21(2). 381–432. DOI: 10.1023/A:1023384924981.

Pancheva, Roumyana. 2000. *Free relatives and related matters*. University of Pennsylvania. (Doctoral dissertation).

Sauerland, Uli. 1998. *The meaning of chains*. MIT. (Doctoral dissertation).

Sportiche, Dominique. 2011. French relative *qui*. *Linguistic Inquiry* 42(1). 83–124. DOI: 10.1162/LING_a_00029.

Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26(1). 79–123.

Wiltschko, Martina. 1998. On the syntax and semantics of (relative) pronouns and determiners. *The Journal of Comparative Germanic Linguistics* 2(2). 143–181. DOI: 10.1023/A:1009719229992.

Chapter 15

Rethinking relatives

Jamie Douglas
University of Cambridge

This chapter is concerned with the syntactic size of finite and infinitival relative clauses in English. I claim that these fall into three (or even four) distinct structural sizes. Assuming a cartographic descriptive framework, I provide evidence for this claim from novel observations concerning the (un)availability of adverbial and argument fronting in the different types of relative clause (following Haegeman 2012). Specifically, some relative clauses permit both adverbial and argument fronting, some permit adverbial fronting only, whilst others do not permit fronting at all. Additional support for my claim comes from three instances of categorial distinctness effect (in the sense of Richards 2010), which I argue instantiate a distinctness effect between elements in SpecTopP and SpecFocP.

1 Introduction

Relative clauses (RCs) have been a subject of study within generative frameworks for decades. It is probably fair to say that the syntactic literature has been primarily concerned with how the RC head (the noun modified by the RC) is related to the RC-internal gap, with reconstruction effects playing a prominent role in discussions and analyses. However, rather than focussing on the RC head, I will consider the RC itself. More specifically, I will investigate the syntactic structure and the structural size of English RCs.

The literature typically recognises two distinct structural sizes as far as RCs are concerned: clausal RCs, as in (1), and reduced RCs, as in (2).

(1) Clausal RCs
the man [(who(m)/that) I met yesterday]

Jamie Douglas. 2020. Rethinking relatives. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 297–326. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280655

(2) Reduced RCs
the man [(being) arrested by police yesterday]

I will not discuss reduced RCs here (for recent discussion, see Douglas 2016; Harwood 2017) but will focus exclusively on clausal RCs, simply calling them RCs from now on. I argue that RCs are not homogeneous in their structural size, i.e. they vary in terms of how much syntactic structure they contain. The different types of RC that I will investigate are exemplified below:

(3) Finite *wh*-RCs
   a. The man [who saw me] is John.
   b. The house [which I lived in] fell down.
   c. The house [in which I lived] fell down.

(4) (Finite) *that*-RCs
   a. The man [that saw me] is John.
   b. The man [that I saw] is John.
   c. The house [that I lived in] fell down.

(5) Finite ∅-RCs
   a. The man [I saw] is John.
   b. The house [I lived in] fell down.

(6) Infinitival *wh*-RCs
   a. The man [to whom to speak] is John.
   b. The house [in which to live] is that one.
   c. For a beginner, the course will likely provide a good atmosphere [in which for you to fire your first shots].[1]

(7) Infinitival *for*-RCs
   a. The man [for you to see] is John.
   b. The man [for her to speak to] is John.

(8) Infinitival ∅-RCs
   a. The man [to see] is John.
   b. The man [to speak to] is John.

[1]This example is from: http://hunting.about.com/od/hunting-for-beginners/a/Hunting-For-Beginners.htm. Such examples are not acceptable to all speakers (see, e.g., the judgements in Chomsky & Lasnik 1977; Huddleston et al. 2002), though there are speakers for whom they are acceptable.

The names for the different types of RC should be reasonably transparent. I do not refer to *wh*-RCs with and without preposition pied-piping as different types. Furthermore, I classify examples like (6c) as infinitival *wh*-RCs rather than infinitival *for*-RCs since the *wh*-phrase is further to the left. ∅-RCs are those without an overt *wh*-relative pronoun, *that* or *for*.

The idea that RCs might vary in structural size is not new, with a number of authors claiming a size difference between finite RCs introduced by an overt relative pronoun or complementiser and those not (Bošković 1994; 1996; 1997; 2016; Weisler 1980; Doherty 1993; 2000), or between infinitival RCs relativising on subjects and those relativising on non-subjects (Bhatt 1999). However, previous studies tend not to consider finite and infinitival RCs together, nor to consider the issue from a serious cartographic perspective (though see Haegeman 2012 for the application of such an approach to a range of clause types in English).

My more specific aim is thus to determine the structure and size of the left periphery of full clausal RCs. To investigate this question, I test whether full clausal RCs of the various types illustrated above are compatible with adverbial and argument fronting (including negative preposing), as done in Haegeman (2012) for a range of clause types following the cartographic tradition (Rizzi 1997 et seq. among many others). Unlike Haegeman (2012), I focus exclusively on RCs, demonstrating that there is a lot more to say about RCs and fronting possibilities in their left peripheries. This is largely a result of empirical differences. Haegeman writes:

> In the following discussion judgments are based on the literature and on a number of informants, all speakers of British English. There is, however, interspeaker variation, and some speakers are much more liberal when it comes to the distribution of fronted arguments in English. These speakers may well find that their judgments deviate systematically from those discussed here. Given that the divergence is systematic, I tentatively conclude that their grammar must differ from that of the speakers on whom this work is based. (Haegeman 2012: 54)

I, and some that I have informally consulted, seem to belong to the "much more liberal" speakers of British English (others that I have consulted seem to belong to Haegeman's "not-so-liberal" group).[2] The biggest difference between Haegeman's (2012) reported judgements and those to be reported below is that Haegeman essentially rejects argument fronting in all RCs (a long-standing and

[2]Haegeman (2012) notes where some authors seem to be more liberal, e.g. Radford (2009a).

widespread claim in the literature, see Chomsky 1977 and Bak 1984), whilst I (and some of my consultants) accept it in some (but not all) RC-types. Nonetheless, even when it is permitted, argument fronting is constrained. I will argue that argument fronting is subject to what will be called a *categorial distinctness effect* (see Richards 2010), i.e. an argument that is fronted inside an RC must be of a different phrasal category from whatever is relativised. This will become apparent in §3.

The structure of this chapter is as follows. The adverbial fronting data is laid out in §2, whilst the argument fronting data and the aforementioned categorial distinctness effect are presented in §3. My analysis is laid out in §4 and suggests a close formal relation between relativisation and topicalisation (at least in finite RC contexts). §5 concludes.

2 Adverbial fronting

2.1 Finite *wh*-RCs

Adverbial fronting and adverbial negative preposing seem to behave in more or less the same way, except that adverbial negative preposing triggers so-called subject–auxiliary inversion. In this section, I will show that adverbial fronting is permitted in finite *wh*- and *that*-RCs and in infinitival *wh*-RCs, but is not permitted in the other RC-types.

Adverbial fronting is permitted in *wh*-RCs, both in non-subject RCs, as in (9), and in subject RCs, as in (10) (see also Doherty 1993; 2000). The same applies to adverbial negative preposing, as in (11) (non-subject RCs) and (12) (subject RCs).

(9) a. I met a man who *next year* Mary might (actually) date.
b. I bought a dress which *next year* Mary might (actually) wear.

(10) a. I met a man who *next year* might (actually) date Mary.
b. I bought a dress which *next year* might (actually) make Mary popular.

(11) a. I met a man who *under no circumstances* would Mary ever date.
b. I bought a dress which *under no circumstances* would Mary ever wear.

(12) a. I met a man who *under no circumstances* would ever go out with Mary.
b. I bought a dress which *under no circumstances* would ever make Mary popular.

The *wh*-relative pronoun may or may not pied-pipe a preposition. Adverbial fronting is compatible with either option, as in (13). The same applies to adverbial negative preposing, as in (14).

(13) a. I met a man who *next year* Mary might (actually) grant a second date to.
b. I met a man to whom *next year* Mary might (actually) grant a second date.

(14) a. I met a man who *under no circumstances* would Mary ever grant a first date to.
b. I met a man to whom *under no circumstances* would Mary ever grant a first date.

2.2 Finite *that*-RCs

Adverbial fronting is permitted in *that*-RCs, both in non-subject RCs, as in (15), and in subject RCs, as in (16) (see also Doherty 1993; 2000). The same applies to adverbial negative preposing, as in (17) (non-subject RCs) and (18) (subject RCs).

(15) a. I met a man that *next year* Mary might (actually) date.
b. I bought a dress that *next year* Mary might (actually) wear.

(16) a. I met a man that *next year* might (actually) date Mary.
b. I bought a dress that *next year* might (actually) make Mary popular.

(17) a. I met a man that *under no circumstances* would Mary ever date.
b. I bought a dress that *under no circumstances* would Mary ever wear.

(18) a. I met a man that *under no circumstances* would ever go out with Mary.
b. I bought a dress that *under no circumstances* would ever make Mary popular.

that-RCs do not permit pied-piping of prepositions at all so (19b) and (20b) are ungrammatical independently of adverbial fronting and adverbial negative preposing respectively.

(19) a. I met a man that *next year* Mary might (actually) grant a second date to.
b. * I met a man to that (*next year*) Mary might (actually) grant a second date.

(20) a. I met a man that *under no circumstances* would Mary ever grant a first date to.
b. * I met a man to that (*under no circumstances*) would Mary ever grant a first date.

2.3 Finite ∅-RCs

Unlike in finite *wh*-RCs and finite *that*-RCs, adverbial fronting is not permitted in finite ∅-RCs (see also Doherty 1993; 2000). This applies to both non-subject RCs, as in (21), and subject RCs, as in (22). Note, however, that finite subject ∅-RCs are generally impossible in (standard) English.[3] In other words, the examples in (22) are ungrammatical independently of adverbial fronting. Exactly the same holds of adverbial negative preposing, as in (23) (non-subject RCs) and (24) (subject RCs).

(21) a. * I met a man *next year* Mary might (actually) date.
b. * I bought a dress *next year* Mary might (actually) wear.

(22) a. * I met a man (*next year*) might (actually) date Mary.
b. * I bought a dress (*next year*) might (actually) make Mary popular.

(23) a. * I met a man *under no circumstances* would Mary ever date.
b. * I bought a dress *under no circumstances* would Mary ever wear.

(24) a. * I met a man (*under no circumstances*) would (ever) go out with Mary.
b. * I bought a dress (*under no circumstances*) would (ever) make Mary popular.

Ø-RCs do not permit pied-piping of prepositions in general. Hence (25b) and (26b) are ungrammatical independently of adverbial fronting or adverbial negative preposing respectively.

(25) a. * I met a man *next year* Mary might (actually) grant a second date to.
b. * I met a man to (*next year*) Mary might (actually) grant a second date.

[3] There are apparent counterexamples, such as (i):

(i) There's a man sells vegetables at the market.

However, there is good reason to believe that these are not instances of genuine ∅-RCs (see den Dikken 2005; Harris & Vincent 1980; Henry 1995; Lambrecht 1988; McCawley 1998), so I set these aside (pace Doherty 1993; 2000).

(26) a. * I met a man *under no circumstances* would Mary ever grant a first date.
b. * I met a man to (*under no circumstances*) would Mary ever grant a first date.

2.4 Infinitival *wh*-RCs

In English, infinitival *wh*-RCs obligatorily involve a pied-piped preposition. Subject infinitival *wh*-RCs are consequently impossible because subjects do not have any prepositions to pied-pipe. All of the examples therefore involve non-subject relativisation. As can be seen, adverbial fronting and adverbial negative preposing is permitted, as in (27) and (28) respectively.

(27) Mary's the woman to whom *next week* to hand these documents.

(28) Mary's the woman to whom *under no circumstances* to ever hand these documents.

Some speakers allow the complementiser *for* and an overt subject in infinitival *wh*-RCs, though even then it is typically judged as somewhat degraded. Other speakers judge it ungrammatical (see Chomsky & Lasnik 1977; Huddleston et al. 2002: 1067). For those that do accept such structures, adverbial fronting is permitted in such cases. The fronted adverbial obligatorily precedes *for*, as in (29).

(29) a. ?? Mary's the woman to whom *next week* for you to hand these documents.
b. * Mary's the woman to whom for you *next week* to hand these documents.

The same seems to be true for adverbial negative preposing, as in (30).

(30) a. ?? Mary's the woman to whom *under no circumstances* for you to ever hand these documents.
b. * Mary's the woman to whom for you *under no circumstances* to ever hand these documents.

2.5 Infinitival *for*-RCs

Unlike in infinitival *wh*-RCs (with and without *for*), adverbial fronting is not permitted in infinitival *for*-RCs, i.e. infinitival RCs with overt *for* but no *wh*-relative pronoun, as in (31). The same applies to adverbial negative preposing, as in (32).

(31) a. ?* I met a man *next year* for you to bring to the party.
b. * I met a man for you *next year* to bring to the party.

(32) a. * I met a man *under no circumstances* for you to ever bring to the party.
b. * I met a man for you *under no circumstances* to ever bring to the party.

Infinitival *for*-RCs do not permit pied-piping of prepositions in general. Hence (33) and (34) are ungrammatical independently of adverbial fronting and adverbial negative preposing.

(33) * Mary's the woman to (*next week*) for you to hand these documents.

(34) * Mary's the woman to (*under no circumstances*) for you to ever hand these documents.

2.6 Infinitival ∅-RCs

Like in infinitival *for*-RCs, adverbial fronting is not permitted in infinitival ∅-RCs, i.e. infinitival RCs with neither *for* nor a *wh*-relative pronoun, as in (35). The same applies to adverbial negative preposing, as in (36).

(35) * I met a man *next year* to bring to the party.

(36) * I met a man *under no circumstances* to ever bring to the party.

Infinitival ∅-RCs do not permit pied-piping of prepositions in general, hence (37) and (38) are ungrammatical independently of adverbial fronting and adverbial negative preposing.

(37) * Mary's the woman to (*next week*) to hand these documents.

(38) * Mary's the woman to (*under no circumstances*) to ever hand these documents.

2.7 Summary

Adverbial fronting and adverbial negative preposing are permitted in finite *wh*-RCs, finite *that*-RCs, and infinitival *wh*-RCs (with and without *for*). They are not permitted in finite ∅-RCs, infinitival *for*-RCs, and infinitival ∅-RCs. Furthermore, they do not seem to interact with preposition pied-piping in any way.

3 Argument fronting

3.1 Finite *wh*-RCs

I turn now to argument fronting. As I will show, argument fronting is more constrained than adverbial fronting. Indeed, as pointed out in §1, Haegeman's (2012) analysis is based on cases where argument fronting in RCs is generally impossible. This seems to be true for some of the speakers I have consulted as well. However, other speakers are "more liberal". Nevertheless, even for these more liberal speakers it is not the case that fronted arguments are freely permitted in all types of RC. As will be seen, argument fronting exhibits a *categorial distinctness effect*. Anticipating the findings, argument fronting is permitted in finite *wh*- and *that*-RCs but not in the other RC-types.

Let us first consider non-subject RCs. Fronted arguments are acceptable to "more liberal" informants, as in (39).[4] The fronted argument obligatorily follows the relative pronoun, as shown by the ungrammaticality of (40).

(39) a. I met a man to whom, *a second date*, Mary might actually grant.
b. I bought a car in which, *muddy shoes*, I would never allow.

(40) a. * I met a man, *a second date*, to whom Mary might actually grant.
b. * I bought a car, *muddy shoes*, in which I would never allow.

However, argument fronting is restricted. Observe that in (39) the *wh*-relative pronouns have pied-piped a preposition. Interestingly, without such pied-piping, the examples become degraded or unacceptable, as in (41).

(41) a. ?* I met a man who(m), *a second date*, Mary might actually grant to.
b. ?* I bought a car which, *muddy shoes*, I would never allow in.

The same effect can be seen when it is the fronted argument rather than the relative pronoun that has the option of pied-piping a preposition. In (42), the fronted argument has pied-piped a preposition and the result is acceptable, whilst in (43), it has not pied-piped a preposition and the result is unacceptable.

(42) I witnessed the second date which, *to that man*, Mary should never have granted.

[4] Similarly, Radford (2009a: 282) judges the following example as acceptable:

(i) A university is the kind of place in which, that kind of behaviour, we cannot tolerate.

(43) * I witnessed the second date which, *that man*, Mary should never have granted to.

What these data tell us is that the relative pronoun and fronted argument cannot both be nominal phrases (DPs). If one is a DP, the other must pied-pipe a preposition, i.e. be a prepositional phrase (PP). To my knowledge, this is a novel empirical generalisation. Adopting Richards's (2010) terminology, I refer to this as a *categorial distinctness effect*.

This raises the question of what happens when both the relative pronoun and fronted argument pied-pipe a preposition. The result is grammatical (example adapted from Totsuka (2014).

(44) I met a man *with whom*, *about linguistics*, I could talk all day.

However, there is an issue about whether the fronted PP in such examples is actually an argument (see Rizzi 1997: 294, 322–325). I leave such examples aside for now but will return to them in §4.4.

The categorial distinctness effect is particularly important when it comes to argument fronting in subject RCs. It has been claimed that fronted topics, or fronted arguments more generally, are impossible in subject RCs (Haegeman 2012: 58; Rizzi 1997: 307). The following examples, taken from Rizzi (1997: 307), are intended to show that fronted arguments are possible in non-subject RCs, as in (45a) and (46a), but impossible in subject RCs, as in (45b) and (46b) (judgements as in the original).[5]

(45) a. ?? the man to whom, *that book*, I gave
b. * the man who, *that book*, gave to me

(46) a. ? a man to whom, *liberty*, we should never grant[6]
b. * a man who, *liberty*, should never grant to us

[5]Haegeman (2012: Ch. 2, note 6) notes via personal communication with Andrew Radford that he accepts the following:

(i) He's the kind of person who, a noble gesture like that, would simply not appreciate.

I, and others, find this example odd. We feel that it needs a subject resumptive pronoun to be even marginally acceptable, as in (ii). Interestingly, an object resumptive does not seem even marginally possible, as in (iii). See §4.4 for discussion.

(ii) ? He's the kind of person who, a noble gesture like that, he would simply not appreciate.

(iii) * He's the kind of person who, a noble gesture like that, would simply not appreciate it.

However, observe that the non-subject RC examples in (45a) and (46a) satisfy categorial distinctness whilst the subject RC examples in (45b) and (46b) do not. If the categorial distinctness effect is responsible for the ungrammaticality of (45b) and (46b), the prediction is that fronted arguments will be allowed in subject RCs provided that the fronted argument pied-pipes a preposition. This prediction is borne out as the contrast between (47) and (48) shows.

(47) a. * I met a man who, *Mary*, might actually grant a second date to.
b. * I bought a car which, *children*, can give hours of entertainment to.

(48) a. ? I met a man who, *to Mary*, might actually grant a second date.
b. I bought a car which, *to children*, can give hours of entertainment.

These data thus show that argument fronting *is* permitted in subject RCs but that the fronted argument must be a PP in line with the categorial distinctness effect.

The same effect can be seen with argument negative preposing. As the contrasts below show, if the relative pronoun has not pied-piped a preposition, the fronted argument must do so. This applies to both non-subject and subject RCs.

(49) a. I met a man who(m), *to no woman*, would I ever recommend (as a date).
b. I bought a dress which, *to no woman*, would I ever give (as a present).
c. I met a man who, *to no woman*, would ever give roses.
d. I bought a dress which, *to no woman*, would ever be given (as a present).

(50) a. * I met a man who(m), *no woman*, would I ever recommend to (as a date).
b. * I bought a dress which, *no woman*, would I ever give to (as a present).
c. * I met a man who, *no woman*, would ever give roses to.
d. * I bought a dress which, *no woman*, would ever be given to (as a present).

The negative preposed argument can only be a DP if the relative pronoun pied-pipes a preposition.

(51) a. I met a man to whom, *no advice* would I ever give.
b. I met a woman to whom, *no roses* would a man ever give.

[6](46a) is adapted from Baltin (1982: 17). Baltin judges it as acceptable, but notes that not all speakers find it totally acceptable.

(52) a. * I met a man who(m), *no advice* would I ever give to.
b. ?? I met a woman who(m), *no roses* would a man ever give to.

To summarise, I have shown that argument fronting is permitted in finite *wh*-RCs but is subject to a categorial distinctness effect. The categorial distinctness effect says that a relative pronoun and fronted argument cannot both be DPs. If one is a DP, the other must be a PP. This is schematised in Table 15.1.

Table 15.1: Categorial distinctness effect

Relative pronoun	Fronted argument	Combination allowed?
DP	DP	*
DP	PP	✓
PP	DP	✓
PP	PP	See §4.4

3.2 Finite *that*-RCs

Argument fronting is permitted in finite *that*-RCs and is subject to the categorial distinctness effect. However, for whatever reason, preposition pied-piping is not possible with *that*, which rules out PP–DP and PP–PP, and I predict from the categorial distinctness effect that option DP–DP is not available either. Consequently, I predict that DP–PP is the only option, i.e. the fronted argument can only be a PP. This prediction is borne out and applies to both non-subject and subject RCs.

(53) a. I bought a dress that, *to Mary*, I might consider giving (as a present).
b. I bought a dress that, *to Mary*, could be given (as a present).
c. I bought a car that, *to children*, would give hours of entertainment.

(54) a. * I bought a dress that, *Mary*, I might consider giving to (as a present).
b. * I bought a dress that, *Mary*, could be given to (as a present).
c. * I bought a car that, *children*, would give hours of entertainment to.

The same applies to argument negative preposing.

(55) a. I bought a dress that, *to no woman*, would I ever give (as a present).
b. I bought a dress that, *to no woman*, would ever be given (as a present).

(56) a. * I bought a dress that, *no woman*, would I ever give to (as a present).
b. * I bought a dress that, *no woman*, would ever be given to (as a present).

If there is no preposition for the fronted argument to pied-pipe in the first place, we predict that argument fronting will simply be unavailable. This prediction is also borne out as the following examples show.

(57) a. * I bought a car that, *muddy shoes*, I would never allow in.
b. * I bought a car that, *hours of entertainment*, would give to children.
c. * I bought a car that, *the children*, can keep entertained.

(58) a. ?? I bought a car that, *not a single muddy shoe* would I ever allow in.
b. * I bought a car that, *not a single hour of entertainment*, would ever give to any child.
c. * I bought a car that, *no child*, can keep entertained.

3.3 Finite ∅-RCs

Unlike finite *wh*-RCs and finite *that*-RCs, argument fronting is not permitted in finite ∅-RCs at all, even if the fronted argument is a PP. Since subject ∅-RCs are generally impossible in English, only non-subject ∅-RCs are illustrated.

(59) a. * I met a man, *Mary*, I might recommend to (as a date).
b. * I bought a dress, *Mary*, I could give to (as a present).

(60) a. * I met a man, *to Mary*, I might recommend (as a date).
b. * I bought a dress, *to Mary*, I could give (as a present).

Pied-piping of prepositions is not permitted with ∅. Therefore, if argument fronting were possible at all, we would expect PP fronted arguments to be possible, as they were with *that*-RCs. Since PP fronted arguments are impossible, I conclude that argument fronting is generally impossible in finite ∅-RCs.

Argument negative preposing behaves in exactly the same way.

(61) a. * I met a man, *no woman* would I ever recommend to (as a date).
b. * I bought a dress, *no woman* would I ever give to (as a present).

(62) a. * I met a man, *to no woman* would I ever recommend (as a date).
b. * I bought a dress, *to no woman* would I ever give (as a present).

3.4 Infinitival *wh*-RCs

Argument fronting is not permitted in infinitival *wh*-RCs (regardless of whether *for* is present or not), even if the fronted argument is a DP. Since infinitival *wh*-RCs obligatorily involve pied-piping of a preposition, if argument fronting were possible at all, we would expect DP fronted arguments to be possible. Since they are not, I conclude that argument fronting is generally impossible in infinitival *wh*-RCs.

(63) a. * I found an ideal venue in which, *Mary*, to propose to.
b. * I found an ideal venue in which, *Mary*, for you to propose to.
c. * I found an ideal venue in which for you, *Mary*, to propose to.

(64) a. * I found an ideal venue in which, *to Mary*, to propose.
b. * I found an ideal venue in which, *to Mary*, for you to propose.
c. * I found an ideal venue in which for you, *to Mary*, to propose.

Similarly, argument negative preposing is not permitted (regardless of whether *for* is present or not, and regardless of whether the fronted argument is a PP or a DP).

(65) a. * This is a place in which, *no man*, to ever give one's real name to.
b. * This is a place in which, *no man*, for you to ever give your real name to.
c. * This is a place in which for you, *no man*, to ever give your real name to.

(66) a. * This is a place in which, *to no man*, to ever give one's real name.
b. * This is a place in which, *to no man*, for you to ever give your real name.
c. * This is a place in which for you, *to no man*, to ever give your real name.

3.5 Infinitival *for*-RCs

As with infinitival *wh*-RCs, argument fronting is not permitted in infinitival *for*-RCs at all, regardless of whether the fronted argument is a DP or a PP.

(67) a. * I found an ideal venue, *Mary*, for you to propose to in.
b. * I found an ideal venue for you, *Mary*, to propose to in.

(68) a. * I found an ideal venue, *to Mary*, for you to propose in.
b. * I found an ideal venue for you, *to Mary*, to propose in.

The same applies to argument negative preposing.

(69) a. * I saw a venue, *no woman*, for one to propose to in.
b. * I saw a venue for one, *no woman*, to propose to in.

(70) a. * I saw a venue, *to no woman*, for one to propose in.
b. * I saw a venue for one, *to no woman*, to propose in.

3.6 Infinitival ∅-RCs

Finally, as with all other infinitival RCs so far, argument fronting is not permitted in infinitival ∅-RCs, regardless of whether the fronted argument is a DP or a PP.

(71) * I found an ideal venue, *Mary*, to propose to in.

(72) * I found an ideal venue, *to Mary*, to propose in.

The same applies to argument negative preposing.

(73) * I saw a venue, *no woman*, for one to propose to in.

(74) * I saw a venue, *to no woman*, for one to propose in.

3.7 Summary

Argument fronting is permitted in finite *wh*-RCs and *that*-RCs, and is prohibited in finite ∅-RCs and all infinitival RCs. Where argument fronting is permitted, it is subject to a categorial distinctness effect. The relative pronoun (or relative operator in the case of *that*-RCs) and fronted argument cannot both be DPs. If one is a DP, the other must be a PP. Exactly the same pattern is found with argument negative preposing.

4 Analysis and discussion

4.1 The distribution of adverbial and argument fronting

Putting the conclusions from §2 and §3 together, we have the empirical situation regarding the distribution of adverbial and argument fronting in English RCs shown in Table 15.2 (note that the terms *adverbial fronting* and *argument fronting* will now be used to cover their negative preposing counterparts as well).

Table 15.2: Distribution of adverbial and argument fronting in full clausal RCs in English. ✓: allowed; (✓): allowed subject to restrictions; *: not allowed.

	Finite			Infinitival		
	wh-RCs	*that*-RCs	∅-RCs	*wh*-RCs	*for*-RCs	∅-RCs
Adverbial fronting	✓	✓	*	✓	*	*
Argument fronting	(✓)	(✓)	*	*	*	*

I propose that this distribution can be captured by positing (at least) three distinct sizes of RC in English, which I will describe in cartographic terms. Rizzi (2004: 242) proposes the following articulation of the C-domain (* here means "iterable"):

(75) Force > Top* > Int > Top* > Focus > Mod* > Top* > Fin > IP

SpecTopP hosts topic phrases, SpecFocusP hosts focus phrases, SpecIntP hosts high *wh*-elements such as Italian *perché* 'why', and SpecModP hosts fronted adverbials in all but "very special discourse contexts" (Rizzi 2004). I will adopt the simplified version in (76).

(76) Force > Top > Foc > Mod* > Fin > IP

The reasons for this simplification are: (i) I am not concerned with Int; (ii) English does not permit multiple topics (see Haegeman 2012 and references therein); and (iii) English topics can never follow foci (see Haegeman 2012 and references therein). Fronted arguments can be topics or foci. Below, I will address the issue of whether the fronted argument in RCs is a topic or a focus.

I am now in a position to account for the distribution of argument fronting and adverbial fronting in RCs. In brief, I propose that finite *wh*-RCs and *that*-RCs are TopPs, infinitival *wh*-RCs are FocPs, and finite ∅-RCs, infinitival *for*-RCs and infinitival ∅-RCs are FinPs (or alternatively, unsplit CPs). This proposal is summarised in Table 15.3.

FinPs are too small to contain TopP, FocP or ModP. Consequently, they permit neither argument nor adverbial fronting. In finite ∅-RCs, Fin is ∅, whilst in infinitival *for*-RCs, Fin is lexicalised as *for*, in line with previous proposals (Haegeman 2012; Radford 2009b; Rizzi 1997). If infinitival ∅-RCs are FinPs, Fin is also ∅ in these cases. FocPs contain ModP, so permit adverbial fronting. Argument fronting is not permitted because FocP is too small to contain TopP and because

Table 15.3: RC structures

	Finite *wh*-RCs	Finite *that*-RCs	Infinitival *wh*-RCs	Finite ∅-RCs	Infinitival *for*-RCs	Infinitival ∅-RCs
Adverbial fronting	✓	✓	✓	*	*	*
Argument fronting	(✓)	(✓)	*	*	*	*
Structural size	TopP		FocP	FinP (or unsplit CP)		

relativisation in infinitival *wh*-RCs targets SpecFocP. Finally, TopPs contain FocP and ModP. Consequently, they permit argument fronting (focus fronting) and adverbial fronting. I assume that Top is lexicalised as *that* in *that*-RCs, but is ∅ in finite *wh*-RCs (where the *wh*-relative pronoun occupies SpecTopP). In the following subsections, I will expand on and discuss various aspects of this proposal.

4.2 FinP RCs

There is potentially a size difference between finite ∅-RCs and infinitival *for*-RCs on the one hand, and infinitival ∅-RCs on the other. The evidence comes from accessibility in the sense of Keenan & Comrie (1977), i.e. the grammatical functions that can be relativised. Finite ∅-RCs and infinitival *for*-RCs can relativise any argument (except the subject), including arguments embedded inside (finite) clauses. Infinitival ∅-RCs can also relativise any argument (including the subject), but cannot relativise out of an embedded finite clause (Longenbaugh 2016), at least for some speakers.[7] This is shown in the following examples (the (e) and (f) examples in (77) to (79) are taken or adapted from Longenbaugh 2016).

(77) Finite ∅-RCs

a. * I found a man can fix the sink.
b. I found a sink you can fix.
c. I found a woman you can give a present to.
d. I found a boy you can force to run faster.
e. I found a play you can prove was written by Shakespeare.
f. I found a play you can prove (that) Shakespeare wrote.

[7] I have found the judgements of (79e,f) to be somewhat variable.

(78) Infinitival *for*-RCs

a. * I found a man for to fix the sink.
b. I found a sink for you to fix.
c. I found a woman for you to give a present to.
d. I found a boy for you to force to run faster.
e. I found a play for you to prove was written by Shakespeare.
f. I found a play for you to prove (that) Shakespeare wrote.

(79) Infinitival ∅-RCs

a. I found a man to fix the sink.
b. I found a sink to fix.
c. I found a woman to give a present to.
d. I found a boy to force to run faster.
e. * I found a play to prove was written by Shakespeare.
f. ?* I found a play to prove (that) Shakespeare wrote.

If this is correct, infinitival ∅-RCs seem to exhibit A′-properties in that arguments can be relativised without higher arguments intervening with such movement, as well as A-properties in that such movement is clause-bound (at least for some speakers), as shown by the ungrammaticality of relativising an element from an embedded finite clause in (79e, f). In contrast, finite ∅-RCs and infinitival *for*-RCs exhibit A′-properties. Longenbaugh (2016) suggests that the hybrid A′/A-properties are the result of a composite probe, i.e. one seeking both A- and A′-related features. One could hypothesise that, if a C-domain is absent, both A- and A′-features are present on T, whilst if a C-domain is present, the A-features are on T and the A′-features in the C-domain. If this is correct, this suggests the following three things. First, finite ∅-RCs, infinitival *for*-RCs and infinitival ∅-RCs all lack the requisite structure to host fronted adverbials and fronted arguments, i.e. their C-domains contain no structure higher than FinP. Second, finite ∅-RCs and infinitival *for*-RCs do have at least some portion of the C-domain. Third, infinitival ∅-RCs may lack a C-domain altogether.

4.3 FocP RCs

According to my proposal, infinitival *wh*-RCs do not permit argument fronting because relativisation and argument fronting would be competing for the same position, namely SpecFocP. However, it has also been claimed in the literature that argument fronting is generally impossible in infinitival clauses (see Bianchi

1999: 206–208). Evidence comes from the impossibility of argument fronting in raising and control infinitivals (Haegeman 2012: 67–68; see also Hooper & Thompson 1973: 484–485).

(80) Hooper & Thompson (1973: 485)
  a. * My friends tend, the more liberal candidates, to support.
  b. * I have decided, your book, to read.

Argument fronting is also prohibited in ECM complements (Haegeman 2012: Ch. 2, note 20).

(81) Culicover & Levine (2001: 297, fn. 14)
  a. * I really want, that solution, Robin to explore thoroughly.
  b. * Police believe, the London area, the suspect to have left.

However, this evidence does not rule out structural size being relevant since these infinitival clauses could themselves be too small to host fronted arguments. Instead, we need to test an infinitival clause that is independently considered to be quite large. If argument fronting is impossible in such cases, this is evidence that argument fronting is simply impossible in infinitival clauses regardless of their size. However, if argument fronting is possible, it suggests that structural size does play a role in the availability of argument fronting. In this respect, consider embedded questions. It is typically said that *wh*-phrases in embedded finite contexts target a higher position in the left periphery (SpecForceP) than in matrix contexts (SpecFocP) (see Haegeman 2012; Pesetsky 1995), thereby capturing the observation that matrix *wh*-phrases follow topics but embedded *wh*-phrases precede them. The high position of *wh*-phrases in embedded clauses is potentially related to clause-typing (Cheng 1991). Now, assuming that *wh*-phrases in embedded infinitival questions also occupy a high left peripheral position for clause-typing, observe that argument fronting seems to be possible. The examples may not be perfect, but they certainly seem better than those in (80) and (81).

(82) a. ? John didn't know what, to Mary especially, to say at a time like that.
  b. ? I asked to whom, this particular form, to give so that it would be processed promptly.

Therefore, it seems that argument fronting is not incompatible with infinitival contexts per se (pace Bianchi 1999), and I thus conclude that infinitival *wh*-RCs do not permit argument fronting because they are structurally too small and not because they are infinitival.

Finally, a potential problem is that infinitival *wh*-RCs do not seem to be necessarily associated with focus interpretations (Luigi Rizzi, p.c.). This may be due to us erroneously associating the lowest position for fronted arguments in the C-domain with SpecFocP. The crucial proposal that I am making is that infinitival *wh*-RCs have only a single position for fronted arguments in their left-periphery. This is targeted by relativisation and hence blocks all other argument fronting. If it turns out that there is a position for fronted arguments below FocP (see Douglas 2016: 83, fn. 15), what I have been calling FocP RCs would actually be slightly smaller than FocP. However, the essence of the present proposal would remain unaffected.

4.4 TopP RCs

I now return to finite *wh*-RCs and *that*-RCs, which I have proposed are TopPs. This proposal makes several (correct) predictions. First, if relativisation targets SpecTopP, we predict that there is only a single position left for argument fronting. Thus, we expect multiple argument fronting to be permitted in non-RC contexts, but only single argument fronting in RC contexts. This prediction is borne out. English permits multiple fronted arguments in non-RC contexts always in the order topic–focus (Culicover 1991; Haegeman 2012).[8]

(83) That book, *to John* Mary gave in 1979.

However, it is extremely difficult if not impossible to have multiple fronted arguments within RCs.

(84) * the year in which, that book, *to John* Mary gave

Alternatively, the difficulty with multiple argument fronting in RCs may be due to the categorial distinctness effect, i.e. it may simply be too difficult to front two arguments and relativise an element whilst simultaneously respecting categorial distinctness. To tease these two options apart, I will consider a second prediction made by the present analysis.

My analysis predicts that fronted arguments in finite *wh*- and *that*-RCs will target SpecFocP, i.e. the fronted argument will be a focus rather than a topic. On a hypothetical alternative analysis, multiple argument fronting is allowed in principle but ruled out by categorial distinctness. This means that a single fronted

[8]The standard claim is that multiple topics are not permitted in English (Haegeman 2012 and references therein), and that multiple foci are not permitted generally (Haegeman 2012; Rizzi 1997).

argument could be either a focus or a topic in principle. To distinguish these two hypotheses, we must thus ask whether the fronted argument behaves like a topic at all. The empirical situation is difficult, but overall the fronted argument in RCs seems to be a focus rather than a topic, as will be shown below, thereby supporting our analysis rather than the hypothetical alternative.

I will apply two of Rizzi's (1997) topic/focus diagnostics. Rizzi shows that foci exhibit weak crossover (WCO) whilst topics do not. As the following data show, the fronted argument always seems to be sensitive to WCO suggesting that it must be a focus and cannot be a topic (the judgements may be quite subtle in some cases).

(85) a. ?* the school to which, John Smith$_i$ his$_i$ mother is planning to send
b. the school to which, John Smith Mary is planning to send

(86) a. ? the person to whom, this book$_i$ its$_i$ author is happy to give for free
b. the person to whom, this book Mary is happy to give for free

As a second diagnostic, Rizzi (1997) notes that topics can be resumed by resumptive pronouns, but foci cannot (at least in Italian). Although English does not typically make use of resumptive pronouns (unless with hanging topics or to repair certain island violations), it seems that the fronted argument is not very readily resumed by a resumptive pronoun. In fact, it seems more acceptable to resume the RC head (or relative pronoun) than the fronted argument (recall footnote 5). This suggests that the fronted argument must be a focus and cannot be a topic. Consider the following contrasts:

(87) a. ?* a man to whom, *unfettered liberty* we would never grant *it*
b. ? a man *to whom*, unfettered liberty we would never grant *to him*

(88) a. ?* a man to whom, *this book* Mary would happily give *it*
b. ? a man *to whom*, this book Mary would happily give *to him*

Although none of these considerations are conclusive in isolation, they nevertheless both seem to converge on the conclusion that argument fronting in English RCs is always focalisation and never topicalisation. This in turn suggests that the ban on multiple argument fronting in RCs in English, as in (84), is due to the idea that SpecTopP is targeted by relativisation and so cannot be targeted by topicalisation as well. This thus suggests that relativisation and topicalisation compete for the same position, i.e. SpecTopP. This formally captures the long-standing intuition that relativisation and topicalisation are intimately related (see Abels

2012; Bianchi 1999; Kuno 1973; 1976; Williams 2011) and could in fact suggest that topicalisation feeds relativisation in English and other languages (see Douglas 2016 for discussion of English and Malagasy in this respect).

The third prediction made by our analysis concerns the categorial distinctness effect. As seen above, this effect holds between the fronted argument and the relative pronoun/operator, i.e. between the constituents in SpecFocP and SpecTopP. If this is correct, we might also expect to find the categorial distinctness effect between foci and topics more generally. This is indeed what we find.

(89) a. This present, *to Mary* I would give.
b. * This present, *Mary* I would give to.

(90) a. To Mary, *this present* I would give.
b. * Mary, *this present* I would give to.

(89) shows that, if the topic phrase is a DP, the focus phrase cannot be a DP, as in (89b), and must be a PP, as in (89a). (90) shows that, if the focus phrase is a DP, the topic phrase cannot be a DP, as in (90b), and must be a PP, as in (90a). As far as I am aware, this is a novel observation and lends independent and important support to our proposal.

Finally, our analysis is able to incorporate Richards's (2010) idea of why the relative pronoun in infinitival *wh*-RCs obligatorily pied-pipes a preposition in English.

(91) Infinitival *wh*-RCs
a. * the man *whom* to talk to
b. the man *to whom* to talk

Richards (2010) proposes that this is due to a categorial distinctness effect between the *wh*-relative pronoun and the external determiner of the RC head. Richards (2010: 35) provides the following schematic structures:

(92) Infinitival *wh*-RCs
a. * [$_{DP}$ *D* [$_{NP}$ N=RC head [$_{CP}$ [$_{DP}$ wh-*relative pronoun*] [$_{C'}$ C [$_{TP}$ …]]]]]
b. [$_{DP}$ *D* [$_{NP}$ N=RC head [$_{CP}$ [$_{PP}$ *P* [$_{DP}$ wh-*relative pronoun*]] [$_{C'}$ C [$_{TP}$ …]]]]]

According to Richards, D and N are not phase heads. Consequently, the DP relative pronoun and the external determiner D in (92a) are linearised in the same spellout domain. This yields the linearisation statement ⟨D,D⟩ (amongst others).

However, because the two D's are non-distinct, ⟨D,D⟩ is uninterpretable at the interfaces by hypothesis. This is the categorial distinctness effect and accounts for the ungrammaticality of (91a). In (92b), however, the DP relative pronoun is embedded in a PP (where P is a phase head). Consequently, the external determiner D and the DP relative pronoun are linearised in separate spellout domains so the problematic ⟨D,D⟩ statement never arises and (91b) is grammatical.

Richards (2010) highlights that his structures in (92) simply serve to illustrate his proposal; they are not integral to it. Consequently, I adapt the structures in (92) to those in (93) to be more consistent with our conclusions and assumptions.

(93) Infinitival *wh*-RCs

a. * [$_{DP}$ D [$_{TopP}$ *[$_{DP}$ RC head]* Top [$_{FocP}$ [$_{DP}$ wh-*relative pronoun]* Foc [$_{FinP}$ Fin [$_{TP}$ …]]]]]

b. [$_{DP}$ D [$_{TopP}$ *[$_{DP}$ RC head]* Top [$_{FocP}$ *[$_{PP}$ P [$_{DP}$* wh-*relative pronoun]]* Foc [$_{FinP}$ Fin [$_{TP}$ …]]]]]

Following Borsley (1997) and Bianchi (2000), I analyse the RC head as a DP phrase (rather than as an N head, as in 92). In this way, the categorial distinctness effect arises because the DP relative pronoun and the DP RC head are linearised in the same spellout domain, i.e. the categorial distinctness effect is a relation between two phrases rather than between two heads, as in (92).

Now, recall that I argued independently on the basis of the distribution of adverbial and argument fronting that infinitival *wh*-RCs are FocPs. In (93), I have shown the RC head as being in SpecTopP. This can be interpreted under the raising analysis of RCs (see especially Bianchi 1999; 2000) if one assumes that the RC head is subextracted out of the relative pronoun DP, or under the matching analysis if one assumes that the RC head can be base-generated in SpecTopP (see Douglas 2016 for discussion). What is interesting for present purposes is that, once again, the categorial distinctness effect holds between the constituents in SpecFocP and SpecTopP. According to Richards's (2010) account, this would mean that Top is not a phase head. If it were, the constituent in SpecTopP and the one in SpecFocP would be in different spellout domains and we would not expect any categorial distinctness effect, contrary to fact.[9]

Why does the RC head in finite *wh*-RCs not exhibit categorial distinctness effects with the relative pronoun?

[9]Note that, if this is correct, it would suggest that the C-domain is not a dynamic phase domain (in the sense of Bošković 2014; Harwood 2015), i.e. it cannot be the case that the highest head in the C-domain (whatever it may be) is necessarily phasal (in fact, Bošković 2014 explicitly leaves the C-domain out of his discussion of dynamic phases). If it were, we would expect the Top head in infinitival *wh*-RCs to be a phase head.

(94) Finite *wh*-RCs
  a. the man *whom* I should speak to
  b. the man *to whom* I should speak

The answer that our analysis provides is that the relative pronoun is located in SpecTopP in such cases and the RC head is higher, i.e. in SpecForceP, as schematised in (95).

(95) Finite *wh*-RCs
  a. [$_{\text{DP}}$ D [$_{\text{ForceP}}$ *[$_{DP}$ RC head]* Force [$_{\text{TopP}}$ [$_{DP}$ wh-*relative pronoun*] Top [$_{\text{FocP}}$ Foc [$_{\text{FinP}}$ Fin [$_{\text{TP}}$ …]]]]]
  b. [$_{\text{DP}}$ D [$_{\text{ForceP}}$ *[$_{DP}$ RC head]* Force [$_{\text{TopP}}$ *[$_{PP}$ P [$_{DP}$* wh-*relative pronoun]]* Top [$_{\text{FocP}}$ Foc [$_{\text{FinP}}$ Fin [$_{\text{TP}}$ …]]]]]

In other words, whilst there is a categorial distinctness effect between constituents in SpecFocP and SpecTopP, there is no such effect between constituents in SpecTopP and SpecForceP. Again, on Richards's (2010) account, this would suggest that Force *is* a phase head. As a result, the constituents in SpecForceP and SpecTopP would be in different spellout domains and no categorial distinctness effect would arise between them, i.e. if the constituent in SpecForceP is a DP, the constituent in SpecTopP can be either a PP, as in (95b), or a DP, as in (95a).

I have thus argued that infinitival *wh*-RCs are FocPs with the RC head being located in SpecTopP, and the finite *wh*-RCs are TopPs with the RC head being located in SpecForceP. This analysis allows us to give a uniform analysis of the categorial distinctness effects in the three contexts identified above: (i) between topics and foci in non-RC contexts, (ii) between relative pronouns and fronted foci in finite *wh*-RCs, and (iii) between the RC head and relative pronouns in infinitival *wh*-RCs. This brings our proposal very close to the configurations proposed by Bianchi (1999; 2000; 2004). However, whilst Bianchi proposes that the RC head moves into SpecTopP or SpecForceP, i.e. a head raising analysis of RCs, I believe that there are various reasons for adopting a matching analysis of RCs instead whereby the RC head is base-generated in SpecTopP or SpecForceP rather than moving into these positions (see Douglas 2016: Ch. 2 for details and discussion). Although it might be unorthodox to posit that the RC head in a matching analysis is base-generated in a high C-domain position, Chierchia (2016) has recently proposed that the crucial property of A-positions is that they are positions that introduce discourse markers. This applies to theta-positions and the EPP-subject position, but also to certain discourse-based positions such as topic positions. This potentially provides a rationale for why the RC head may

be base-generated in SpecTopP (self-evidently a topic position). Whether it can be extended to SpecForceP is a matter I leave for future research.

I have proposed that the restrictions on argument fronting found in finite *wh*-RCs manifest the categorial distinctness effect found more generally between the constituents in SpecTopP and SpecFocP in English. Recall that the categorial distinctness effect I have been considering effectively restricts the distribution of fronted DP arguments, i.e. I have said that two fronted arguments cannot both be DPs. What about PPs? If the effect is really one of categorial distinctness, we would predict that two fronted arguments cannot both be PPs either. However, PPs do not seem to be sensitive to the categorial distinctness effect. Recall (44), repeated as (96) below:

(96) I met a man *with whom*, *about linguistics*, I could talk all day.

In (96), the relative pronoun and *linguistics* have both pied-piped a preposition resulting in two fronted PPs in the C-domain. Totsuka (2014) concludes on the basis of such examples that there is *no* categorial distinctness effect between the relative pronoun and the fronted argument, contrary to what I have demonstrated for DPs (Totsuka does not discuss the data I have been concerned with though). However, there is a serious question about whether *about linguistics* is an argument PP as opposed to a fronted adverbial PP (see Rizzi 1997: 294, 322–325). Although it is difficult to front a lot of material simultaneously in English, it at least seems marginally possible to front the RC subject in an example like (97).

(97) ? I met a man with whom, *Mary*, about linguistics, could talk all day.

Crucially, both the focussed subject DP and the PP *about linguistics* can co-occur suggesting that they are not competing for the same position (by hypothesis, SpecFocP). This suggests that the PP *about linguistics* is lower than FocP, plausibly in SpecModP. In fact, given the difficulty of finding multiple PP arguments with any single predicate in English, it may be that the fronted "argument" PP in all examples like (96) is in fact a fronted adverbial PP.

Finally, I return to the issue of categorial distinctness effects in finite *that*-RCs, illustrated in (53) and (54), repeated below.

(98) a. I bought a dress that, *to Mary*, I might consider giving (as a present).
b. I bought a dress that, *to Mary*, could be given (as a present).
c. I bought a car that, *to children*, would give hours of entertainment.

(99) a. * I bought a dress that, *Mary*, I might consider giving to (as a present).
b. * I bought a dress that, *Mary*, could be given to (as a present).
c. * I bought a car that, *children*, would give hours of entertainment to.

This pattern can be straightforwardly assimilated to the pattern in finite *wh*-RCs if *that* is analysed as a relative pronoun rather than a complementiser, except that unlike the *wh*-relative pronouns it cannot pied-pipe a preposition (see, e.g., Kayne 2014). However, there are dialects of English where both a relative pronouns and *that* can co-occur (see Trotta 2004: 6) suggesting that *that* is not a relative pronoun and is in fact a complementiser.

If *that* is a complementiser, we can hypothesise that the fronted argument is interacting with the null relative operator in finite *that*-RCs, which (for whatever reason) is always a DP, never a PP. This is potentially problematic for Richards's (2010) approach to categorial distinctness, according to which categorial distinctness effects arise when linearisation statements involve two non-distinct categories. If one of those elements does not require linearisation, e.g. if it is unpronounced, Richards suggests that there will be no distinctness effect. For example, Richards proposes that traces (or the unpronounced copies in a movement chain) do not count for linearisation because the system can tell *pre-linearisation* that such elements will be null. If we wish to maintain Richards's system for finite *that*-RCs, the system must not be able to tell that the relative operator in finite *that*-RCs is going to be null until after the linearisation statements have been calculated. The raising analysis would have trouble accounting for this under Richards's system since the relative operator is a trace/copy, whilst the matching analysis could capture this if the relative operator becomes null post-syntactically (see Douglas 2016 for further discussion of the raising and matching analyses).

5 Conclusion

I have reached the conclusion that the different types of clausal RCs in English systematically differ in structural size. This accounts for the various fronting possibilities. Finite *wh*- and *that*-RCs are the largest: they can host fronted adverbials and fronted focussed arguments. Infinitival *wh*-RCs are the next largest: they can host fronted adverbials but not fronted arguments. Finite ∅-, infinitival *for*- and infinitival ∅-RCs are the smallest: they do not permit fronting of any kind. This is summarised in Table 15.4.

I argued that argument fronting in finite *wh*- and *that*-RCs is focalisation, not topicalisation. I suggested that topicalisation in these RCs is ruled out because relativisation and topicalisation compete for the same structural position. Similarly,

Table 15.4: Summary of RC structures

	Finite *wh*-RCs	Finite *that*-RCs	Infinitival *wh*-RCs	Finite ∅-RCs	Infinitival *for*-RCs	Infinitival ∅-RCs
Adverbial fronting	✓	✓	✓	*	*	*
Argument fronting	(✓)	(✓)	*	*	*	*
Structural size	TopP		FocP	FinP (or unsplit CP)		FinP/TP

I suggested that focalisation in infinitival *wh*-RCs is ruled out because focalisation and relativisation compete for the same structural position. I thus concluded that finite *wh*- and *that*-RCs are TopPs, whilst infinitival *wh*-RCs are FocPs. I also proposed that the other types of RC are FinPs (or unsplit CPs), i.e. they have a C-domain with a single C head, or, in the case of infinitival ∅-RCs, perhaps no C-domain at all.

I also observed that English exhibits a categorial distinctness effect in the C-domain in (at least) three environments: (i) between the relative pronoun/operator and fronted (focussed) argument in finite *wh*- and *that*-RCs; (ii) between topic and focus in non-RC contexts; and (iii) between the RC head and relative pronoun/operator in infinitival *wh*-RCs (following Richards 2010). I proposed that these are three instances of a single effect, namely the categorial distinctness effect between topic and focus in English, and that relativisation and topicalisation are formally similar (at least in finite RCs).

Abbreviations

ECM	exceptional case marking	RC	relative clause
EPP	extended projection principle	WCO	weak crossover

Acknowledgements

I wish to thank Ian Roberts, Theresa Biberauer, Luigi Rizzi, Guglielmo Cinque, David Willis, and the audience of SynCart 1 (Chiusi, 2016) for helpful feedback and comments on this work. This paper is adapted from Chapter 3 of my thesis, which was supervised by Ian Roberts at the University of Cambridge. I gratefully

acknowledge the AHRC grant no. 04271 and the ERC grant no. 269752 (*Rethinking comparative syntax*).

References

Abels, Klaus. 2012. The Italian left periphery: A view from locality. *Linguistic Inquiry* 43(2). 229–254. DOI: 10.1162/LING_a_00084.

Bak, Sung-Yun. 1984. Why not a topic in a relative clause? *Language Research* 20(2). 163–180.

Baltin, Mark R. 1982. A landing site theory of movement rules. *Linguistic Inquiry* 13(1). 1–38.

Bhatt, Rajesh. 1999. *Covert modality in non-finite contexts*. University of Pennsylvania. (Doctoral dissertation).

Bianchi, Valentina. 1999. *Consequences of antisymmetry: Headed relative clauses*. Berlin: De Gruyter. DOI: 10.1515/9783110803372.

Bianchi, Valentina. 2000. The raising analysis of relative clauses: A reply to Borsley. *Linguistic Inquiry* 31(1). 123–140. DOI: 10.1162/002438900554316.

Bianchi, Valentina. 2004. Resumptive relatives and LF chains. In Luigi Rizzi (ed.), *The structure of CP and IP* (The Cartography of syntactic Structures 2), 76–114. Oxford: Oxford University Press.

Borsley, Robert D. 1997. Relative clauses and the theory of phrase structure. *Linguistic Inquiry* 28(4). 629–647.

Bošković, Željko. 1994. Categorial status of null operator relatives and finite declarative complements. *Language Research* 30(2). 387–417.

Bošković, Željko. 1996. Selection and the categorial status of nonfinite complements. *Natural Language & Linguistic Theory* 14(2). 269–304. DOI: 10.1007/BF00133685.

Bošković, Željko. 1997. *The syntax of nonfinite complementation*. Cambridge, MA: MIT Press.

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1). 27–89. DOI: 10.1162/LING_a_00148.

Bošković, Željko. 2016. What is sent to spell-out is phases, not phasal complements. *Linguistica* 56(1). 25–66. DOI: 10.4312/linguistica.56.1.25-66.

Cheng, Lisa Lai-shen. 1991. *On the typology of wh-questions*. MIT. (Doctoral dissertation).

Chierchia, Gennaro. 2016. *Weak crossover: A "dynamic" explanation?* Paper presented at 1st SynCart Workshop, Chiusi, 14th July.

Chomsky, Noam. 1977. On wh-movement. In Peter W. Culicover, Thomas Wasow & Adrian Akmajian (eds.), *Formal syntax*, 71–132. New York: Academic Press.
Chomsky, Noam & Howard Lasnik. 1977. Filters and Control. *Linguistic Inquiry* 8(3). 425–504.
Culicover, Peter W. 1991. Topicalization, inversion and complementizers in English. In Denis Delfitto, Martin Everaert, Arnold Evers & Frits Stuurman (eds.), *Going romance and beyond*, 1–45. Utrecht: University of Utrecht.
Culicover, Peter W. & Robert D. Levine. 2001. Stylistic inversion in english: A reconsideration. *Natural Language & Linguistic Theory* 19(2). 283–310. DOI: 10.1023/A:1010646417840.
den Dikken, Marcel. 2005. A comment on the topic of topic–comment. *Lingua* 115(5). 691–710. DOI: 10.1016/j.lingua.2003.11.005.
Doherty, Cathal. 1993. *Clauses without* that*: The case for bare sentential complementation in English*. University of California, Santa Cruz. (Doctoral dissertation).
Doherty, Cathal. 2000. *Clauses without "that": The case for bare sentential complementation in English*. New York: Garland.
Douglas, Jamie. 2016. *The syntactic structures of relativisation*. University of Cambridge. (Doctoral dissertation). http://ling.auf.net/lingbuzz/003182.
Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery* (The Cartography of Syntactic Structures 8). Oxford: Oxford University Press.
Harris, Martin & Nigel Vincent. 1980. On zero relatives. *Linguistic Inquiry* 11(4). 805–807.
Harwood, William. 2015. Being progressive is just a phase: Celebrating the uniqueness of progressive aspect under a phase-based analysis. *Natural Language & Linguistic Theory* 33(2). 523–573. DOI: 10.1007/s11049-014-9267-3.
Harwood, William. 2017. Reduced relatives and extended phases: A phase-based analysis of the inflectional restrictions on English reduced relative clauses. *Studia Linguistica* 72(2). 428–471. DOI: 10.1111/stul.12069.
Henry, Alison. 1995. *Belfast English and Standard English: Dialect variation and parameter setting*. Oxford: Oxford University Press.
Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.
Huddleston, Rodney, Geoffrey K. Pullum & Peter G. Peterson. 2002. Relative constructions and unbounded dependencies. In Rodney Huddleston & Geoffrey K. Pullum (eds.), *The Cambridge grammar of the English language*, 1031–1096. Cambridge: Cambridge University Press.

Kayne, Richard S. 2014. Why isn't *this* a complementizer? In Peter Svenonius (ed.), *Functional structure from top to toe* (The Cartography of Syntactic Structures 9), 188–231. Oxford: Oxford University Press.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1). 63–99.

Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, MA: MIT Press.

Kuno, Susumu. 1976. Subject, theme and speaker's empathy: A re-examination of relativization phenomena. In Charles N. Li (ed.), *Subject and topic*, 417–444. New York: Academic Press.

Lambrecht, Knud. 1988. There was a farmer had a dog: Syntactic amalgams revisited. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*. 319–339. DOI: 10.3765/bls.v14i0.1796.

Longenbaugh, Nicholas. 2016. Difficult movement. Paper presented at GLOW 39, April 5–8, University of Göttingen.

McCawley, James D. 1998. *The syntactic phenomena of English*. Second edition. Chicago, IL: University of Chicago Press.

Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.

Radford, Andrew. 2009a. *An introduction to English sentence structure*. Cambridge: Cambridge University Press.

Radford, Andrew. 2009b. *Analysing English sentences: A minimalist approach*. Cambridge: Cambridge University Press.

Richards, Norvin. 2010. *Uttering trees*. Cambridge, MA: MIT Press.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Springer.

Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond* (The Cartography of Syntactic Structures 3), 223–251. New York: Oxford University Press.

Totsuka, Masashi. 2014. On syntactic distinctness and phase head in English relative clause. 文化 77(3/4). 74–88.

Trotta, Joe. 2004. *Comments on vacuous subject movement*. Halmstad University. (Doctoral dissertation).

Weisler, Steven. 1980. The syntax of *that*-less relatives. *Linguistic Inquiry* 11(3). 624–631.

Williams, Edwin. 2011. *Regimes of derivation in syntax and morphology*. London: Routledge.

Chapter 16

V3 in urban youth varieties of Dutch

Marieke Meelen
University of Cambridge

Khalid Mourigh
Leiden University

Lisa Lai-Shen Cheng
Leiden University

In this paper we compare new data from Dutch urban youth varieties to emerging varieties in other Germanic languages like German and Norwegian. We argue that, unlike previously thought, V3 word orders can be found in urban youth varieties of Dutch as well and present data from our new corpus. The V3 patterns in our dataset share most characteristics of the optional V3 innovations observed in other Germanic urban youth varieties: the sentence-initial constituent is a frame-setter of any category and the preverbal constituent is mainly the subject that functions as a *familiar topic*. We adopt Walkden's (2017) analysis and extend it by adding an additional FrameP so that preverbal constituents that do not function as familiar topics could be accounted for as well. Following Wolfe's cline of possible V2-languages, we argue that the Dutch urban youth varieties can best be analysed as "Force-V2 system 1" grammars with V-to-Force movement + an additional FrameP. They thus differ from Standard Dutch, which is argued to be a "Force-V2 system 2" based on the fact that only hanging or left-dislocated topics can be found in sentence-initial position of superficial V3 patterns. This data thus presents an interesting case of syntactic change in the opposite direction: from strict V2 to V2 with optional V3 orders.

Marieke Meelen, Khalid Mourigh & Lisa Lai-Shen Cheng. 2020. V3 in urban youth varieties of Dutch. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 327–355. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280657

1 Introduction

Main clauses in Modern Dutch are characterised by the verb-second (V2) constraint (cf. Zwart 1997). Just like in Modern German and Scandinavian languages, the finite verb linearly follows a variety of sentence-initial constituents, as shown in (1) for subjects, objects and adjuncts.[1]

(1) Standard Dutch
   a. Ian *vierde* zijn verjaardag gisteren.
      Ian celebrated his birthday yesterday
   b. Zijn verjaardag *vierde* Ian gisteren.
      His birthday celebrated Ian yesterday
   c. Gisteren *vierde* Ian zijn verjaardag.
      Yesterday celebrated Ian his birthday
      'Ian celebrated his birthday yesterday.'

All three options are grammatically correct in Standard Dutch, but the choice of sentence-initial constituent is pragmatically conditioned. Verb-third (V3) orders as seen in the English translation of example (1c), are not allowed in Standard Modern Dutch:

(2) Standard Dutch
   *Gisteren Ian *vierde* zijn verjaardag.
   yesterday Ian celebrated his birthday
   Intended: 'Ian celebrated his birthday yesterday.'

Recently, some varieties of Germanic V2 languages have been reported to exhibit V3 orders alongside the standard V2 patterns (see, among others, Freywald et al. 2015, Wiese 2013, Wiese & Rehbein 2016 and Walkden 2017). These new Germanic varieties have emerged in multilingual settings in large cities in various countries in Europe.[2] Various examples of these unexpected V3 or XSV orders in these

[1] Throughout this article the inflected verbs in the examples will be indicated in *italics*. Unless specified otherwise, all examples are from a small corpus of a Dutch urban youth variety compiled by Khalid Mourigh in 2013–2017, recorded in Gouda (see also §2 and the Appendix).

[2] The term "urban youth varieties" will be used for these varieties of Dutch throughout this paper, because it has the least pejorative connotation and it captures the sociolinguistic characteristics of being spoken by young people in urban, multilingual settings. Other terms for these varieties of Danish, Norwegian, Swedish and German, such as "ethnolect", "multiethnolect", "Kiezdeutsch" ('neighborhood German') or "Kebab Norwegian" are problematic because they do not characterise the exact nature of the varieties and often have strong derogatory overtones (cf. Walkden 2017; Aarsæther 2010).

languages that usually exhibit the V2 constraint have been cited by Freywald et al. (2015) and Walkden (2017):[3]

(3) a. German urban youth variety (Wiese 2009: 787)
morgen ich *geh* Arbeitsamt
tomorrow I go job.centre
'Tomorrow I will go to the job centre.'

b. Norwegian urban youth variety (Opsahl 2009: 133)
nå de *får* betale
now they get pay
'Now they have to pay.'

c. Danish urban youth variety (Quist 2008: 47)
normal man *går* på ungdomsskolen
usually one goes to youth.club
'Normally you attend the youth club.'

d. Swedish urban youth variety (Ganuza 2008: 53)
då alla *börja(de)* hata henne
then everyone started hate her
'Then everyone started hating her.'

Appel (1984), Appel & Muysken (1987: 91) and Schwartz & Sprouse (2000) have reported that adult L2 learners of Dutch produce adverb-subject-verb orders (XSV or AdvSV) as well:

(4) Dutch L2 learner (Appel 1984)
En dan hij *gaat* weg.
and then he goes away
'And then he goes away.'

[3]Since the preverbal constituent is usually the subject of the sentence, Freywald et al. (2015) refer to them as "XSV" with any type of constituent "X" preceding the subject and the verb. In our present corpus, we only find preverbal subjects as well. Walkden (2017), however, presents some examples of light adverbials in the German urban youth variety "Kiezdeutsch". The lack of light adverbials like *hier* 'here' and *da* 'there' in our present corpus is presumably the result of our small dataset rather than the result of a structural restriction. The Dutch adverbs (*hier* and *daar*) are functionally equivalent to their German counterparts and we therefore have no reason to assume urban varieties of Dutch differ in this respect from Kiezdeutsch. The Dutch urban dialect could in theory be different, however. Therefore, we continue to use the term "V3" to refer to these innovative word order patterns.

However, according to Freywald et al. (2015), there are very few violations of the V2 constraint found in three case studies of Dutch they examined: bilateral interviews with a mixed groups of young people from Lombok (Cornips 2002), interviews with four male adolescents of Surinamese, Creole descent (Cornips & De Rooij 2013) and in- and out-group conversations in the classical Labovian method with speakers from a Dutch, Moroccan–Dutch, and Turkish–Dutch background. The only three examples are the following (cited by Freywald et al. 2015: 86–87):

(5) a. Utrecht/TCULT corpus, Badir
toen we *hadden* eerst twee autos
then we had first two cars
'Then, we first had two cars (and later only one).'

b. Utrecht/TCULT: Badir
daarom ik *heb* dat probleem niet
that's why I have that problem not
'That's why I don't have that problem.'

c. Adam-Nijmegen/etnolects project: Hassan, see Lukassen (2011)
daarom Nederland *is* niet echt meer van eh
that's why the Netherlands is not really more like eh
'That's why the Netherlands is no longer more like eh …'

They conclude from this that the Dutch urban youth variety, unlike its V2 neighbours in Germany, Denmark, Norway and Sweden, "does not allow loosened grammatical restrictions in respect to the XSV order" (Freywald et al. 2015: 88).

In this article we first present new data from a Dutch urban youth variety spoken by Dutch teenagers with a Moroccan heritage in Gouda (§2 and §3). We argue that these new data show that this Dutch urban youth variety indeed exhibits violations of the strict V2 constraint. V3 orders are attested in our dataset and we suggest this is an indication that Dutch urban youth varieties show the same characteristics as their Germanic neighbours (§3). We then proceed to consider these V3 orders in their syntactic context. Although our present dataset is still quite limited, we will present a tentative synchronic analysis, elucidating this optional variation in the context of the Standard Dutch C-domain (§4.1). We then sketch a possible scenario of language change and how this relates to the diachronic analyses that have been proposed for this phenomenon in other Germanic urban vernaculars (section §4.2). Finally, we define some areas of future work, based on the need for different types of data collection and other syntactic deviations from Standard Dutch that affect the C-domain (§5).

2 Linguistic setting

The present study is based on a corpus of oral interviews conducted by one of the authors with Moroccan Dutch teenagers in Gouda. Gouda, which is a rather small city with 71,105 inhabitants, has the largest Moroccan Dutch population in the Netherlands with 6,892 members. About half of the Moroccan population in Gouda belong to the second generation, meaning that they were born in the Netherlands and have at least one parent who was born in Morocco. According to the people interviewed in Gouda, most members of the local Moroccan Dutch community originate from the region of Nador in North Morocco, more specifically from Ayt Said, making this linguistically a tight-knit group.

This means that a large percentage of its members have Riffian Berber as their heritage language (98.5% of the population of the countryside of Ayt Said speaks Tarifiyt Berber[4]). Dialectal Arabic also plays an important role as a lingua franca in general. While it is not used for everyday communication, Standard Arabic still plays an important role in religious life and in the media. People who were born and raised in the Netherlands primarily use Dutch in daily life (already in the 1980s, cf. De Ruiter 1989). With their parents they often speak Berber or (dialectal) Arabic, or they code-switch between one of these languages and Dutch. Therefore, Berber and Arabic can be considered heritage languages (cf. Montrul 2016).

The total corpus consists of roughly thirteen hours of interviews with thirty-one people (see the Appendix for a full overview of speaker codes we use in our examples, including interview settings and language backgrounds, based on Mourigh 2017). The interviews were conducted in groups of at least two people with the interviewer always present. All interviews were conducted with male teenagers except for two teenage girls who have the same ethnic background. The teenagers share a similar socio-economic and educational background. At the time of recording they either attended secondary school (VMBO) or lower vocational training (MBO). The interviews were conducted at different places in informal settings such as the hallway of a sports club, a cultural centre, close to the school and in the town centre. All interviews were conducted in Dutch with occasional code-switching to Berber or Arabic.

The interviews inevitably suffer from the observers' paradox, and even though the interviewer shares the ethnic background of the interviewees, he does not share other characteristics such as age and place of residence. The interviewer

[4]Statistics from www.hcp.ma, last accessed on 13 December 2017. Tarifiyt Berber is one of the three major Berber languages spoken in Morocco.

had the impression that many interviewees were quite comfortable. However, the lack of certain lexical elements, such as Berber and Arabic discourse markers, which are typical for Moroccan Dutch discourse indicate that their speech was somewhat influenced (Kossmann 2017). This might also be a reason for the infrequent occurrence of V3 order in the corpus. In general, even in the corpora of other Germanic urban varieties, V3 occurrences are quite rare, both in interviews and in self-recordings (cf. Ganuza 2008).

In addition to the corpus, from which most of the examples were drawn, some data originate from videoclips that Moroccan Dutch youngsters themselves put on YouTube.[5] These are not from Gouda and therefore indicate that it is a more widespread phenomenon.

3 Describing the V3 data

In this section we present the data that show deviations from the Standard Dutch V2 pattern. We describe this data in terms of the initial constituent (the "X" in XSV orders), the preverbal constituent (the subject) and, finally, the distribution of possible V3 orders. Before moving on to the aberrant V3 orders in these urban varieties, however, we must discuss the superficial V3 orders that are in fact allowed in the Standard Dutch V2 grammar.

The occurrence of such V3 orders in our urban vernacular data would not be unexpected if these sentences are acceptable in Standard Dutch. Therefore sentences like examples (6a) and (6b) with hanging topics are excluded:

(6) Standard Dutch

a. Noord-Wales, dat *is* echt een mooie plek om op vakantie te gaan.
North Wales, that is really a lovely place to on holiday to go.INF
'North Wales, that's a really lovely place to go on holiday.'

b. Die boeken, die *moet* je zorgvuldig behandelen.
those books those must you carefully treat.INF
'As for those books, you should treat those with care.'

Greco & Haegeman (2020) discuss another type of V3 order in Standard Dutch that appears in the context of circumstantial frame-setters. Frame-setting topics

[5]Data taken from videos on the following channels: https://www.youtube.com/watch?v=acFL0W3Y1ZY and https://www.youtube.com/user/Youstoub, last accessed on 13 December 2017.

are usually adjuncts in sentence-initial position. They set the scene and/or delimit the space or time in which the event described in the following comment takes place. These frame-setters can be combined with non-subject initial orders or non-declaratives, as shown in examples (7a) and (7b), respectively.

(7) Standard Dutch

a. Als je haar iets vraagt, nooit *antwoordt* ze op tijd.
if you her something ask.2SG never reply.3SG she on time
'If you ask her something, she never replies on time.'

b. Als er morgen een probleem is, MIJ *moet* je niet bellen.
if there tomorrow a problem is me must you not call
'If there is a problem tomorrow, don't call ME!'

Because these are allowed in Standard Dutch[6] as well, this paper about the Dutch youth varieties from Gouda is not concerned with these types of V3 orders. In the following sections we will present the data and describe their characteristics in terms of type of initial constituent, preverbal constituent and distribution in a wider context.

3.1 The sentence-initial constituent

There seems to be no categorial restriction on the initial constituent in the Dutch urban vernacular dataset. There are determiner phrases (DPs), prepositional phrases (PPs), adverbial phrases (APs) or entire clauses (CPs) shown in examples (8a), (8b), (8c) and (8d) respectively:

(8) a. MD-A
Een keertje ik *was* gewoon aan het fietsen
one time I was just on the cycle.INF
'One time I was just cycling.'

[6]Greco & Haegeman (2020) note that sentences with subject-initial V3 orders and *circumstantial* frame-setters are acceptable in the West-Flemish dialect of Dutch, but not in Standard Dutch.

(i) OK in West-Flemish; but * in Standard Dutch
*Als mijn tekst klaar is, ik *zal* hem opsturen.
when my text ready is I shall it send
'When my text is ready, I will send it.'

They argue, however, that these V3 orders systematically differ from the V3 orders innovated by young Germanic speakers in urban settings discussed in the present paper. We will leave this discussion for future research.

b. YouTube video Maisdokter
Op een gegeven moment hij *douwt* zo'n mais in zijn kont.
at a given moment he pushes such.a corn.cob in his butt
'At some point he pushes a corn cob in his butt.'

c. MD-I
Hier je *bent* verzekerd.
here you are insured
'Here you are insured.'

d. MD-B
Wanneer we hem slaan, hij *gaat* gelijk huilen.
when we him beat he goes straight cry.INF
'If we beat him he immediately starts to cry.'

This lack of categorial preference for the sentence-initial constituent corresponds to the V3 patterns found in urban varieties of Norwegian, Swedish and German. Walkden (2017) illustrates this with examples from Kiezdeutsch in particular, but the same seems to hold for the new V3 patterns observed in Norwegian and Swedish urban youth varieties.

3.1.1 Sentence-initial frame-setters

Although our dataset is limited, we still find such categorial variety. All these initial constituents are adjuncts indicating a specific time or location. This is exactly what has been observed in other Germanic urban youth varieties (see Freywald et al. 2015: 84 and Walkden 2017). Freywald et al. (2015) characterise this type of initial constituent as "an interpretational frame or anchor" for the immediately following proposition. This type of "frame-setter" (cf. Chafe 1976) thus provides a certain limitation in terms of time or place.[7] As Walkden (2017) points out, it is important to note that this type of frame-setter may also occur as the initial constituent in regular V2 structures in the standard varieties of Germanic V2 languages. Example (9), in Standard Dutch, would have subject-verb inversion as expected in V2 languages:[8]

[7] Freywald et al. (2015) add a "conditional" function to temporal or locational functions of these frame-setters. However, in light of the possible V3 orders with conditional frame-setters in Standard Dutch discussed above, we leave the "conditional" specification in Dutch urban vernaculars out of the present discussion.

[8] The use of the diminutive *keertje* 'small time' is actually a further characteristic of non-standard Dutch.

(9) Standard Dutch
Een keertje *was* ik gewoon aan het fietsen
one time was I just on the cycle.INF
'One time I was just cycling.'

3.1.2 Other sentence-initial constituents

Apart from these adjuncts of time and location, there are some other types of initial constituents in V3 structures in our dataset. These can be grouped into three categories, which we briefly discuss below. These examples are less straightforward, because the direct equivalent with subject-inversion in Standard Dutch does not exist. We therefore do not take these into consideration in our analysis in §4.

The first group consists of examples with *omdat* 'because', as shown in (10a) and (10b):

(10) a. MD-K
Omdat ik *vind* het niet goed.
because I find it not good
'Because I don't think it's right.'

b. MD-K
Omdat hij *is* Marokkaan natuurlijk.
because he is Moroccan obviously
'Obviously because he is Moroccan.'

These examples are difficult because *omdat* introduces a subordinate clause in Standard Dutch. Subordinate clauses have SOV order and therefore the Standard Dutch equivalent of (10a) and (10b) would have SOV order following *omdat*:

(11) Standard Dutch
a. … omdat ik het niet goed *vind*.
because I it not good find
'… because I don't think it's right.'

b. … omdat hij Marokkaan *is* natuurlijk.
because he Moroccan is obviously
'… obviously because he is Moroccan.'

In the examples from the Dutch urban youth varieties dataset, *omdat* seems to behave like another Dutch conjunction with the same meaning: *want* 'because'.

The conjunction *want* is typically followed by matrix-clause V2 syntax, as shown in example (12):

(12) Standard Dutch
Want ik *vind* het niet goed.
because I find it not good
'Because I don't think it's right.'

If the conjunction *omdat* in the Dutch urban youth varieties indeed has the syntactic specifications of Standard Dutch *want*, the superficial V3 order we observe here is not unexpected. If *want* is followed by subordinate-clause syntax, not the lack of V2 with subject-inversion, but the lack of SOV order is unexpected. According to Zwart (2011: 123–125), *omdat* can be followed by V2 in the contexts of bridge verbs like *zeggen* 'to say' as well. We therefore do not consider *omdat*-clauses in our urban varieties corpus as part of our proper V3 dataset. We will briefly discuss the implications for subordinate clauses in §5 below.

The second group of examples with superficial V3 orders in the Dutch urban youth varieties involve code-switching from Dutch to Berber and/or Arabic.

(13) a. MD-E
he, weet je, bhal jij *gaat* naar hun
hey know you bhal you go to them
'Hey, you know, *bhal* you go to them.'

b. MD-I
eentje hoor je van die: qa ik *heb* vandaag uh
one hear you of those qa I have today uh
'You hear one of those: *qa* I have today uh'

There are also examples of code-switches or Arabic/Berber interjections with V2 and the expected subject-verb in the urban youth varieties, as shown in example (14).

(14) From YousToub channel
En inshallah *haal* je goede cijfers.
and inshallah get you good grades
'And, *inshallah*, you'll get good grades.'

These sentences with Berber or Arabic discourse markers, however, cannot be compared to Standard Dutch either; we leave them out of the present analysis.

Finally, there is one category of adverbials that do not normally occur in sentence-initial position in Standard Dutch, but that do occur several times in our dataset of superficial V3 orders in the Dutch urban youth varieties:

(15) a. MD-L
zogenaamd je *hebt* geen geld meer
as-if you have no money anymore
'As if you no longer have any money (left).'

b. MD-R
… maar wel ik *begrijp* alles.
but still I understand everything
'…but I do understand everything'

The adverbs *zogenaamd* 'as-if' and *wel* 'still, nonetheless' cannot occur in sentence-initial position in Standard Dutch. In their Standard Dutch equivalents, they would follow the inflected verbs, as shown in examples (16a) and (16b), respectively:

(16) Standard Dutch

a. je *hebt* zogenaamd geen geld meer
you have as-if no money anymore
'As if you no longer have any money (left).'

b. … maar ik *begrijp* wel alles.
but I understand still everything
'… but I do understand everything'

Again, because these sentence-initial constituents with superficial V3 orders in our dataset do not have a direct equivalent, we cannot compare them to Standard Dutch V2. We will exclude these from our analysis presented in §4 below.

3.2 Preverbal constituent

The next crucial element in the superficial V3 orders is the preverbal constituent. In Standard Dutch V2 order, the preverbal constituent is the sentence-initial constituent and it can be an argument or adjunct of a wide variety of phrase types. The V3 orders in the Dutch urban youth varieties mostly exhibit arguments, or, more specifically, subject pronouns in all persons and number, as shown in examples (17a), (17b) and (17c):

(17) a. 24 maart interiew
Soms ik *gooi* iets op de grond.
sometimes I throw something on the floor
'Sometimes I throw something on the floor.'

b. MD-C
één keer we *zaten* bij big Mo film te, televisie te kijken
one time we sat.PL at big Mo film to tv to watch.INF
'Once we were watching a film, tv at big Mo's.'

c. MD-A
Toen ze *vroegen* ID.
then they asked.PL ID
'Then they asked for ID.'

The second-person singular pronoun has stressed and unstressed variants in Standard Dutch: *je* (unstressed) vs. *jij* (stressed). Both occur as the subject in our V3 dataset, as shown in examples (18a) (repeated from 8c) and (18b):

(18) MD-I

a. Hier je *bent* verzekerd.
here you are insured
'Here you are insured.'

b. Daarna jij *ging* mee.
afterwards you went along
'Afterwards you went along.'

From a cross-linguistic perspective, the occurrence of the stressed pronoun *jij* 'you' is unexpected. Freywald et al. (2015: 84) observe that preverbal constituents in urban youth varieties in Germany, Norway or Sweden are "virtually always unaccented" (see also Walkden 2017). Cross-linguistically, the preverbal element is usually the subject of the clause, but as Walkden (2017) points out, this is a "strong tendency rather than a requirement". In the Dutch urban youth varieties dataset, we also find some examples of non-pronominal subjects in preverbal position:

(19) a. MD-I
vroeger mensen *gingen* lopend
in.the.past people went.PL on.foot
'In the past people would go on foot.'

b. MD-I
daarna die, die leraar *heeft* niet meer lesgegeven
afterwards that that teacher has no longer taught
'Afterwards that, that teacher hasn't taught anymore.'

c. YouTube video Maisdokter
Op een gegeven moment iemand *zegt* tegen hem je moet naar Fez
at a certain time someone says to him you must to Fez
'At some point someone says to him: you must go to Fez.'

d. MD-I
daarna de rest *zegt* ik ga niet
afterwards the rest says I go not
'Afterwards the rest says: I'm not going.'

e. YousToub
Vaak het probleem *is* dat ze met de jaren verwachten ze meer.
often the problem is that they with the years expect.PL they more
'Often the problem is that they – as the years go by – expect more.'

According to Freywald et al. (2015), a common denominator of these preverbal constituents lies in their information-structural nature: they are all *familiar topics* that refer to a contextually given or salient discourse referent. Not all examples in the Dutch urban youth varieties data presented in (19) contain familiar topics, however. The subjects of examples (19a) and (19b) could indeed be argued to be linked to the *common ground*, either because they are generic concepts (like *mensen* 'people') or because they have been explicitly mentioned in the preceding discourse (like *die leraar* 'that teacher'). The teacher is the topic of the preceding sentences (all in Berber), in which a boy is being beaten by his teacher, but later comes back to seek revenge and hits the teacher.

The subject of example (19c), *iemand* 'someone', is technically inert and would function more as a *shift topic* than a familiar topic. The referential status of the subject in (19d), *de rest* 'the rest', can be inferred from the context, but it clearly indicates a contrast between this subject and the topic in the immediately preceding discourse. Example (19e) is a copular clause in which *het probleem* 'the problem' in preverbal position could be argued to be the predicate, with the *dat*-clause as its subject. The analysis of these types of copular clauses goes beyond the scope of the present paper, but the fact that a noun phrase like *het probleem* 'the problem' can occupy the preverbal position cannot be ignored. This phrase is certainly not a familiar topic. We will come back to these subtle information-structural differences in §4 below.

3.3 Distribution of V3 orders

The V3 orders in our data do not occur in every main clause. Just like in other Germanic urban youth varieties, the V3 orders are optional deviations from the regular V2 patterns. V3 orders can be found immediately preceding or following regular V2 sentences uttered by the same speaker in the same type of context. Example (20) immediately follows another clause with the same sentence-initial constituent *toen* 'then'. The first clause exhibits regular V2 order, whereas the second clause is V3:

(20) MD-A
Toen *gingen* we wegrennen. Toen ze *vroegen* ID.
then went.PL we run.away.INF then they asked.PL ID
'Then we ran away. Then they asked for ID.'

The V3 orders do not occur very often and when they do, they are found alongside very similar sentences with Standard Dutch V2 order. Since our current data consists of non-elicited sentences only, we cannot check the (un)grammaticality of certain types of V3 orders in different contexts. This is difficult to verify in general, because we are dealing with a non-standard variety of the language which is subject to stylistic variation. The young people who speak this variety often change to Standard Dutch in the presence of people who are not from their peer group.

Ganuza (2008: 109–130) discusses the same sociolinguistic conditions for her focus group speaking Swedish urban varieties. Walkden (2017), based on previous work on Kiezdeutsch by Wiese and Swedish urban varieties by Ganuza, notes that there are three contexts in which these types of V3 orders are not allowed. These are sentences in which the preverbal constituent is the object (rather than the subject), *wh*-interrogatives and subordinate clauses. All examples in our current urban vernacular dataset of Dutch have preverbal subjects and none of the examples are wh-interrogatives. This might be due to a limited dataset, but since these options seem to be excluded in other urban vernaculars, the same generalisation might hold for the Dutch urban vernacular. We have already briefly mentioned our examples with subordinate clauses introduced by *omdat* 'because'. Walkden (2017) notes that there are occasional examples of V3 in clauses introduced by the German *weil* 'because', but that "this is a context in which it is well known that main clause word order may occur in colloquial usage" (Walkden 2017), which is reminiscent of the above-mentioned *omdat*-clauses in Dutch we left out of our proper V3 dataset for now (see also Antomo & Steinbach 2010 and Reis 2013).

4 Analysis

Although our current dataset is still fairly limited, we will attempt to offer a preliminary synchronic analysis of these V3 orders in Dutch urban youth varieties. Until we collect more data, this analysis is necessarily preliminary, but it will help our attempts to sketch a diachronic analysis of ongoing syntactic change in Dutch.

4.1 Synchronic analysis

It is important to emphasise that the synchronic analysis of the V3 patterns should be compatible with a V2 grammar as well, because these V3 orders are only *optional* variants of the Standard Dutch V2. In other words, all speakers with innovative V3 patterns also (indeed, mostly) utter V2 sentences that are the norm in Standard Dutch. Although the V2 constraint observed in various languages shares two crucial characteristics (verb-movement to the C-layer accompanied by the merger of a phrasal constituent, cf. Holmberg 2013 and Wolfe 2015), V2 languages can differ in the way they exhibit these characteristics. Apart from a traditional distinction based on whether V2 is limited to main clauses (as in Dutch, German and Mainland Scandinavian) or appears in subordinate clauses as well (as in Icelandic or Yiddish) (cf. Holmberg 2013), languages also appear to differ in terms of their CP structure.

Recently, the typology of different types of V2 languages was further developed by Wolfe (2019) on the basis of the availability of pro-drop and optional V3 orders. In this typology of V2 languages, Wolfe (2019: 31) distinguishes three types of V2 systems named after the landing site of the verb, based on the landing site of the finite verb (Fin or Force):

Fin-V2: Frame-setter + topic + focus (Old English, Middle Low German, etc.)

Force-V2 system 1: Frame-setter + topic/focus (Later Old French, Spanish, etc.)

Force-V2 system 2: Frame-setter$_{\text{HT/LD}}$ + topic/focus (Modern Dutch and German, etc.)

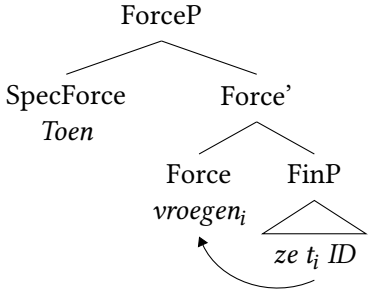
Standard Dutch is classified by Wolfe (2019) as a "Force-V2 system 2" language, because regarding V3 orders, Standard Dutch can only accommodate hanging (HT) or left-dislocated (LD) topics as a sentence-initial constituent. V3/XSV orders found in urban youth varieties are ungrammatical in the standard language.

(21) a. Standard Dutch – HT
Kaapstad, dat *is* echt een mooie plek om op vakantie te gaan.
Cape Town, that is really a lovely place to on holiday to go.INF
'Cape Town, that's a really lovely place to go on holiday.'

b. Standard Dutch – *V3, but probably OK in urban varieties
*In de zomer Kaapstad *is* echt een mooie plek om op vakantie
in the summer Cape Town is really a lovely place to on holiday
te gaan.
to go.INF
intended: 'In summer, Cape Town is really a lovely place to go on holiday.'

The Standard Modern Dutch V2 order with V-to-Force movement is shown in (22):

(22)



(23) Standard Dutch
Toen *vroegen* ze ID.
then asked.PL they ID
'Then they asked for ID.'

As described in §3.1 above, the sentence-initial constituents in the superficial V3 orders in Germanic urban youth varieties function as a frame- or scene-setter. The initial constituents are not arguments, but adjuncts with a temporal or locational meaning such as *toen* 'then', *een keer* 'one time' or *hier* 'here'. The superficial order of constituents in these sentences is thus: Frame – Subject – Verb. In line with Walkden (2017), we assume general V-to-C movement in standard modern Germanic V2 clauses in general and therefore Standard Dutch as well. If the inflected verb moves to a C-head and the subject moves to its specifier, the easiest analysis for the urban vernacular V3 sentences would involve an extra structural layer to host this frame-setting sentence-initial constituent. Independent evidence for extra structural layers in the C-domain is abundantly found in Romance languages, upon which Rizzi (1997: 283) based his split CP:

(24) [Frame… [Force… [Topic… [Focus… [Fin… [TP…]]]]]

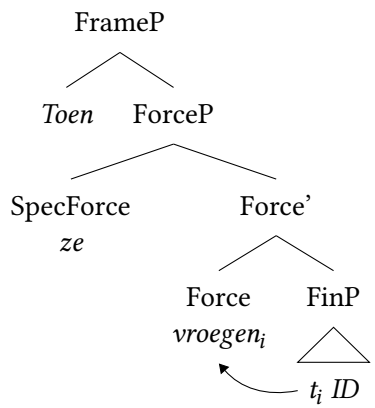
Variations on this were further developed by Benincà & Poletto (2004: 71) and by Frascarelli & Hinterhölzl (2007: 112–113), who later apply this to early Germanic (Hinterhölzl & Petrova 2009):

(25) ForceP > ShiftP > ContrP > FocP > FamP* > FinP

As Roberts (1996a) already observed, analysing V3 orders in Old English, we need to postulate at least one extra layer in the CP if we assume V-to-C movement always occurs in these V2 languages. Roberts (1996b) assumed a distinction between Fin and Focus/Force as the landing site of the finite verb in these cases. Until we have evidence for a further split, we will assume a simple split of the CP into two layers. Note that the so-called "bottle-neck effect" in strict V2 languages like Standard Dutch and German uses locality to prevent movement of more than one constituent into the C-domain (cf. among others Roberts 2004 and Mohr 2009). From this perspective a V2 language with multiple constituents in the C-domain is unexpected and needs to be explained. We follow Walkden's (2017) assumption, based on earlier work by Rizzi (1997) and Haegeman (1995), which states that certain heads may be associated with criteria requiring them to enter into a spec-head configuration with an appropriate XP. This then motivates interpretively-driven movement such as topicalisation, focalisation, wh-questions, etc. Languages with syncretised left peripheries, such as Standard Dutch, only allow one criterion to be active, resulting in the movement of one (and only one) constituent to the C-domain. With Walkden (2017), we assume that V3 orders arise when not one but two of these criteria are to be satisfied.

Since the sentence-initial constituent in Dutch urban youth varieties is always clearly a frame- or scene-setter, it seems appropriate to add an additional FrameP on top of the Standard Dutch ForceP to accommodate the V3 orders in urban youth varieties. Compare example (22) above to the innovative V3 option from our dataset of Dutch urban youth varieties with similar V-to-Force movement, but an added FrameP to host the temporal frame-setter *toen* 'then' in (26):

(26)



(27) Standard Dutch – familiar topic
Toen ze *vroegen* ID.
then they asked.PL ID
'Then they asked for ID.'

Wolfe's typology assumes a cartographic CP-structure based on Rizzi (1997) with a FrameP on top of ForceP, followed by TopP, FocP and FinP. Since urban youth varieties of Dutch allow various kinds of frame-setters (e.g. *daarna* 'afterwards', *soms* 'sometimes', etc.) and only one preverbal topic/focus, the grammar of these varieties can therefore be best described as "Force-V2 system 1" in Wolfe's typology. Speakers with optional V3 orders have access to two registers of Dutch: Standard Dutch with strict V2 ("Force-V2 system 2") and urban varieties with optional additional frame-setters ("Force-V2 system 1"). We assume that style-shifting occurs in more formal contexts, e.g. writing, speaking to non-peers, etc. Wolfe's V2 typology is ultimately a diachronic typology. In the next section, we will turn back to his typology in the light of our diachronic analysis.

4.2 Diachronic analysis

Old English was already analysed as a V2 language by Van Kemenade (1987). In 1996, Ian Roberts makes inferences based on this and work on Gothic by, among others, Kiparsky (1994) and observed that "residual V2" in Present-day English is a misleading term for the actual state of affairs. Comparing characteristics of Old English V2 and V3 orders, it appears that "Full V2" of Modern German and Dutch is better described as an innovation: a stage of "strict V2" that English has never reached. Roberts (1996b) suggests that the V2 and V3 orders in Old English can be analysed with a "split-Comp" structure allowing multiple landing sites for the verb in the left periphery.

To our knowledge, Walkden's (2017) paper on Germanic urban youth varieties (or "urban vernaculars" as he calls them) presents the only comprehensive diachronic analysis of these innovative types of V3 orders. In addition to the urban vernacular data, he draws on insights from, among others, Roberts (1996b) to develop a similar account for the situation in Old English. Walkden's analysis is based on a scenario of imperfect L2 acquisition of the standard V2 language by speakers from a different linguistic background (e.g. immigrants from Turkey, Morocco, etc. moving to Germany, or, in our case, the Netherlands). He proposes three separate stages for the development of optional V3 orders (cf. Walkden 2017):

Stage 1: L2 learners of standard Germanic V2 fail to acquire verb movement to C, resulting in SVO orders

Stage 2: L1 learners (e.g. children of first-generation immigrants) attempt to reconcile mixed input of SVO and V-to-C, resulting in a split-CP (CP1 & CP2) that allows for the observed optional V3 structures in the urban vernaculars

Stage 3: V3 structures are propagated across communities and successive generations increase their use

These diachronic developments are straightforward and they fit the overall sociolinguistic situation with first- and second-generation immigrants in the Netherlands as well. Through socio-historical circumstances, certain areas of the country had a high proportion of L2 learners. Let us go through the implications for the analysis of the Dutch urban vernacular V3 sentences stage by stage.

Stage 1 of the analysis hinges on the failure of the acquisition of verb movement to C. This is necessary for the subsequent stage in which the second generation attempts to make sense of a mixed SVO/V-to-C input. The question is whether this scenario of failure of the acquisition of V-to-C movement is likely for the Moroccan immigrants in the Netherlands. The native language of this first-generation L2 learners is Berber or Moroccan Arabic, although all of them have a good understanding of Standard Arabic as well. Both Berber and Arabic are VSO languages with optional SVO orders. Verb movement in pragmatically neutral matrix clauses in these languages is usually argued to be limited to V-to-T or V-to-AgrSP (cf. amongst others Benmamoun (1992), Jouini (2014) and Shlonsky (2000) for Arabic and Choe (1987) for Berber). In both languages, sentence-initial frame-setters can occur with following VSO orders as well. In a corpus study of child-directed Dutch, MacWhinney & Snow (1985) observed that only 23% of the input was non-subject initial. Although this is apparently enough for Dutch L1 learners to acquire the V2 constraint (see also Yang (2000: 114) for a full discussion), L2 learners might initially interpret the non-subject initial orders in a way that is compatible with the grammar of their first language. We would thus hypothesise that they do not postulate a phi-probe in the C domain resulting in V-to-C movement because they do not require this phi-feature on C to yield XVS orders in their native language. With the next generation, they use their mixed input, leading to Stage 2 in Walkden's proposal. Although at home they might also speak Berber or Moroccan Arabic, Dutch is frequently used in the Moroccan community; there are multiple dialects and languages that are not always mutually intelligible. Since our current number of examples of V3 order are still fairly limited and we have not collected any specific acquisitional data of these L2 learners yet, we leave a further exploration of this hypothesis for future research.

Assuming *Stage 1* has resulted in the failed acquisition of V-to-C movement, in *Stage 2* the next generation consisting of L1 learners of Dutch attempt to reconcile their mixed SVO/V2 input. They acquire V-to-C successfully and their language, the urban youth variety under discussion, has a V2 grammar. To reconcile this

V2 grammar with the SVO input as well, they are forced to postulate a split of the CP to accommodate additional frame-setters.

In *Stage 3* this split is then postulated to be propagated throughout the community. The V3 orders in our data are not limited to a single speaker, but found in interviews with various teenagers from Gouda. In addition to this, we found several examples of these V3 innovations in YouTube videos of young speakers with a Moroccan heritage from other parts of the country. This is a clear indication that the new split-CP grammar has spread amongst teenagers with a Moroccan background in the Netherlands at the very least. The young people with optional V3 orders seem to be aware of the fact that this grammar is associated with a specific register, as they are able to switch to a purely V2 grammar in formal contexts or simply when talking to Dutch speakers outside of their Moroccan Dutch community.[9]

4.3 V3 innovations in a diachronic typology of V2

Recall Wolfe's typology of V2 languages from §4.1, which we present in Figure 16.1. Wolfe (2019) argues that older Germanic varieties provide more options for V3 orders. Early Medieval Romance and Early Old High German allowed both topics and foci in sentence-initial position and are thus classified as a "Fin-V2" system. In later Old French and Spanish and New High German, on the other hand, only a frame-setter and either a topic or a focus constituent was found sentence-initially, making them "Force V2 system 1" languages. In both Germanic and Romance, Wolfe thus observes a change from Fin-V2 to Force-V2 (and within Force-V2 from system 1 to system 2, which ultimately happened in Modern Dutch and German).

From this perspective, the optional V3 orders in the Dutch urban varieties could indicate that this variety of Dutch is in transition (again) from a Force-V2 system 2 (back) to system 1. Would this typology be appropriate for the scenario of language contact and change proposed by Walkden (2017)? A crucial aspect of Walkden's scenario is that the CP cannot be split in the standard V2 language. The simple non-cartographic synchronic analysis with a single CP in Standard

[9] As we have only collected data from young people with a Moroccan background, at this stage we cannot comment on how widespread this phenomenon is outside the Moroccan community in the Netherlands. In addition, more data is needed on the socio-linguistic parameters associated with the possible switch in register. This, however, goes beyond the scope of the present paper and we leave this for future research.

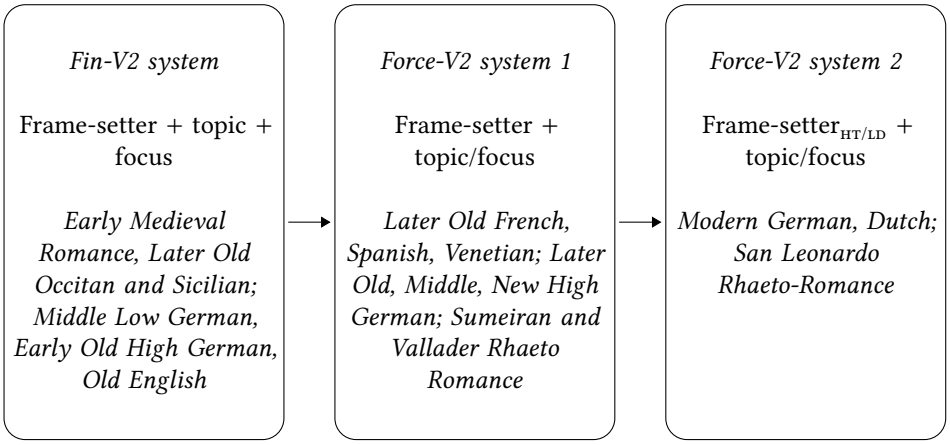


Figure 16.1: V3* in V2 languages (Wolfe 2019: 31)

Dutch splitting into a CP1 and CP2 would therefore work. In the grammar of Dutch urban youth varieties, the outer CP2 is reserved for any type of frame-setter and the inner CP1 hosts the verb and any type of preverbal constituent. These labels need no further specification, although the outer CP2 could be seen as a FrameP since it always hosts a frame- or scene-setter. This consistency provides a good argument for the mapping of information-structural features to a further-defined hierarchical structure in the left periphery, at least for FrameP and ForceP.

If we were to assume the CP of Standard Dutch is already split into further layers of ForceP, FocP, FinP etc. and we thus take a cartographic approach, Walkden's diachronic scenario can only work if the verb in Standard Dutch is in the left-most possible position. If the verb were in a lower position, the need to postulate more structure to reconcile the SVO/V2 input would not arise, so the split sketched by Walkden would not be motivated. The left-most position would be Force in a "ForceP system 2" type of language, which is indeed the position in which the verb lands according to Wolfe (2019). If Walkden's scenario is correct this implies there might be diachronic evidence in addition to Wolfe's synchronic V3 analysis to motivate V-to-Force movement in Standard Dutch. The forced split of the CP (or ForceP) Walkden describes could result in the creation of extra structure in the form of a FrameP that can host any type of frame-setter in a "Force-V2 system 1" type of grammar.

Walkden (2017), however, suggests this split CP conflates information-structural layers as follows:

- CP2 = ForceP, ShiftP, ContrP and FocP (for sentence-initial frame-setters)
- CP1 = FamP and FinP (for preverbal subjects)

CP2 does not include FrameP in this system, forcing the sentence-initial frame-setter to occur lower in the structure, in ForceP, ShiftP or ContrP. CP1 is reserved for FamP and FinP as these host the preverbal subject that are (almost) always familiar topics in the data Walkden discusses. Recall, however, that preverbal subjects in Dutch urban varieties are not always familiar topics:

(28) Standard Dutch

a. Shift topic
Op een gegeven moment iemand *zegt* tegen hem je moet naar Fez
at a certain time someone says to him you must to Fez
'At some point someone says to him: you must go to Fez.'

b. Contrastive topic
daarna de rest *zegt* ik ga niet
afterwards the rest says I go not
'Afterwards the rest says: I'm not going.'

c. Shift topic?
Vaak het probleem *is* dat ze met de jaren verwachten ze meer.
often the problem is that they with the years expect.PL they more
'Often the problem is that they – as the years go by – expect more.'

These types of contrastive or shift topics in preverbal position would be in CP2 in Walkden's split CP if we take the information-structural labels of the split CP seriously. Walkden's mechanism of change can thus only be extended to the Dutch urban varieties if the CP is split differently. We therefore propose the following split:

- CP2 = FrameP (for sentence-initial frame-setters)
- CP1 = ForceP (for preverbal subjects with any information-structural status)

To conclude, we adopt Walkden's diachronic scenario resulting in a situation in which second-generation L1 speakers of Dutch solve their ambiguous SVO/V2 input by creating additional structure in the C-domain. If we confine ourselves to an analysis of Dutch only, it would suffice to postulate a single CP in Standard Modern Dutch that is subsequently reanalysed by the speakers of urban

youth varieties as a simply binary split into CP1 and CP2. From a cross-linguistic perspective, however, it might be desirable to adopt a cartographic layering of the CP that can account for the observed differences in terms of pro-drop, optional V3 orders and the landing site of the verb, as proposed by Wolfe (2019). If we combine Walkden's diachronic scenario with Wolfe's (2019) typology of V2 grammars, the Dutch urban youth varieties are moving away from a "Force-V2 system 2" (Standard Modern Dutch) to a "Force-V2 system 1" with an additional FrameP. Although Wolfe's typology is also based on diachronic syntactic changes, both the Romance and Germanic languages he studied have moved from "Fin-V2" to "Force-V2 system 1" and, in the case of Dutch and German, all the way to "Force-V2 system 2". The innovative V3 orders in urban youth varieties present an interesting case of syntactic change in the opposite direction, i.e. from "Force-V2 system 2" to "Force-V2 system 1".[10]

5 Future work

Some issues discussed in the present paper provide interesting pathways for future work. The generalisations and analyses presented here are based on a small dataset. It would first of all be important to extend our dataset in both qualitative and quantitative ways. The quality of our current data is limited to interview settings with young people from Gouda and some videos in which Dutch teenagers with a Moroccan heritage present themselves and discuss their lives. As mentioned by Freywald et al. (2015), these methods do not necessarily get the best results, because young people change to a more formal (i.e. more Standard Dutch) register whenever an interviewer is present. In our future attempts at data collection, we will therefore aim to leave the recorder with the young people and let them speak without any interference.

From a synchronic point of view, there are some more observations in our current dataset that warrant further discussion. One pattern that is repeatedly found in these urban youth varieties, but not in Standard Dutch, is *dat*-deletion, as shown in (29a):

(29) a. MD-C
Denk je hij weet Gouda uit zijn hoofd?
think you he knows Gouda from his head
'Do you think (that) he knows Gouda by heart?'

[10] A reviewer speculates this type of change in the opposite direction might be associated with language contact and L2 acquisition, whereas change from "Force-V2 system 1" to "Force-V2 system 2" might be "the more natural 'endogenous' change". This is an interesting suggestion that we would like to explore in future research.

b. Standard Dutch
Denk je dat hij Gouda uit zijn hoofd weet?
think you that hij Gouda from his head knows
'Do you think (that) he knows Gouda by heart?'

Both the deletion of the complementiser and the lack of subordinate word order (SOV in Standard Dutch) need to be addressed in any future discussions on the C-domain of these urban youth varieties.

From a diachronic perspective, there are numerous strands for future research, especially from a cross-linguistic perspective. To mention just one in Dutch alone: a more thorough study of the process of L2 acquisition would be beneficial to provide further evidence for the scenario sketched by Walkden (2017).

6 Conclusion

In this paper we compared new data from Dutch urban youth varieties to emerging varieties in other Germanic languages like German and Norwegian. We first of all argued that, unlike previously thought, V3 word orders can indeed be found in urban youth varieties of Dutch as well. We supported this with evidence from a small dataset consisting mainly of interviews with teenagers with a Moroccan heritage living in Gouda, in the west of the Netherlands. Some further examples from Dutch-Moroccan teenagers from other parts of the country presenting themselves on YouTube and online forums suggest this phenomenon is not limited to this community in Gouda. The V3 patterns in our dataset share most characteristics of the optional V3 innovations observed in other Germanic urban youth varieties: the sentence-initial constituent is a frame-setter of any category and the preverbal constituent is mainly the subject that functions as a familiar topic.

There are, however, a couple of examples in our current dataset that do *not* function as familiar topics. We adopted Walkden's (2017) analysis and extended it by adding an additional FrameP so that preverbal constituents that do not function as familiar topics could be accounted for as well. This type of analysis fits well into Wolfe's (2019) typology of V2 languages. Following Wolfe's cline of possible V2-languages, we argued that the Dutch urban youth varieties can best be analysed as "Force-V2 system 1" grammars with V-to-Force movement + an additional FrameP. They thus differ from Standard Dutch, which is argued to be a "Force-V2 system 2" based on the fact that only hanging or left-dislocated topics can be found in sentence-initial position of superficial V3 patterns.

Abbreviations

2	second person	LD	left dislocation
3	third person	PL	plural
HT	hanging topic	SG	singular
INF	infinitive		

Acknowledgements

The research leading to these results has received funding from the European Union's seventh *framework programme* for research, technological development and demonstration under grant agreement no. 613465 and from the European Research Council advanced grant for the project *Rethinking comparative syntax* (ID 269752).

Appendix

Table 16.1 shows the dates and locations of interviews in conducted with young speakers of Moroccan Dutch in Gouda. More details about the speakers and the corpus in general can be found in Mourigh (2017).

Table 16.1: Background of speakers from Mourigh (2017)

Speaker	Interview Date	Location	Heritage language	Duration	Age
MD-A	20-11-2014	City centre	Berber	45 min	18
MD-B1	02-10-2014	Sports club	Berber	23 min	17
MD-B2	02-10-2014	Sports club	(same speaker)	8 min	17
MD-B3	16-10-2014	Sports club	(same speaker)	36 min	17
MD-C	02-10-2014	Sports club	Arabic	23 min	15
MD-E1	26-10-2014	Park	Arabic	35 min	17
MD-E2	15-06-2015	City centre	(same speaker)	60 min	17
MD-H	26-10-2014	Park	Arabic	40 min	17
MD-I	15-06-2015	City centre	Berber	60 min	21
MD-K	30-10-2014	Community centre	Berber	90 min	14
MD-L	30-10-2014	Community centre	Berber	90 min	15

References

Aarsæther, Finn. 2010. The use of multiethnic youth language in Oslo. In Pia Quist & Bente Ailin Svendsen (eds.), *Multilingual urban Scandinavia: New linguistic practices*, 111–126. Bristol: Channel View Publications.

Antomo, Mailin & Markus Steinbach. 2010. Desintegration und Interpretation: Weil-V2-Sätze an der Schnittstelle zwischen Syntax, Semantik und Pragmatik. *Zeitschrift für Sprachwissenschaft* 29(1). 1–37.

Appel, René. 1984. *Immigrant children learning Dutch: Sociolinguistic and psycholinguistic aspects of second-language acquisition*. Dordrecht: Foris.

Appel, René & Pieter Muysken. 1987. *Bilingualism and language contact*. London: Edward Arnold.

Benincà, Paola & Cecilia Poletto. 2004. Topic, focus, and V2: Defining the CP sublayers. In Luigi Rizzi (ed.), *The structure of CP and IP* (The Cartography of Syntactic Structures 2), 52–75. Oxford: Oxford University Press.

Benmamoun, Elabbas. 1992. Structural conditions on agreement. In *Proceedings of NELS*, vol. 22, 17–32.

Chafe, Wallace L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In Charles N. Li (ed.), *Subject and topic*, 27–55. New York: Academic Press.

Choe, Jae W. 1987. LF movement and Pied-Piping. *Linguistic Inquiry* 18(2). 348–353.

Cornips, Leonie. 2002. Etnisch Nederlands. In Hans Bennis, Guus Extra, Pieter Muysken & Jacomine Nortier (eds.), *Een buurt in beweging. Talen en culturen in het Utrechtse Lombok en Transvaal*, 285–302. Amsterdam: Aksant.

Cornips, Leonie & Vincent A. De Rooij. 2013. Selfing and othering through categories of race, place, and language among minority youths in Rotterdam, The Netherlands. In Peter Siemund, Ingrid Gogolin, Monika Edith Schulz & Julia Davydova (eds.), *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education* (Hamburg Studies on Linguistic Diversity 1), 129–164. John Benjamins. DOI: 10.1075/hsld.1.07cor.

De Ruiter, Adrianus Cornelis Jacob. 1989. *Young moroccans in the netherlands: An integral approach to their language situation and acquisition of Dutch*. Utrecht University. (Doctoral dissertation).

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning, and form: Generalizations across languages*, 87–116. Amsterdam: John Benjamins. DOI: 10.1075/la.100.07fra.

Freywald, Ulrike, Leonie Cornips, Natalia Ganuza, Ingvild Nistov & Toril Opsahl. 2015. Beyond verb second–a matter of novel information-structural effects? Evidence from Norwegian, Swedish, German and Dutch. In Jacomine Nortier & Bente Ailin Svendsen (eds.), *Language, youth and identity in the 21st century*, 73–92. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9781139061896.006.

Ganuza, Natalia. 2008. *Syntactic variation in the Swedish of adolescents in multilingual urban settings: subject-verb order in declaratives, questions and subordinate clauses*. Stockholm University. (Doctoral dissertation).

Greco, Ciro & Liliane Haegeman. 2020. Frame setters and microvariation of subject-initial Verb Second. In Rebecca Woods & Same Wolfe (eds.), *Rethinking verb second*, 61–89. Oxford: Oxford University Press.

Haegeman, Liliane. 1995. Root infinitives, tense, and truncated structures in Dutch. *Language Acquisition* 4(3). 205–255. DOI: 10.1207/s15327817la0403_2.

Hinterhölzl, Roland & Svetlana Petrova (eds.). 2009. Berlin: De Gruyter. DOI: 10.1515/9783110216110.

Holmberg, Anders. 2013. Verb second. In Tibor Kiss & Artemis Alexiadou (eds.), *Syntax — Theory and analysis: An international handbook*, vol. 1, 342–383. Berlin: De Gruyter. DOI: 10.1515/9783110377408.342.

Jouini, Kemel. 2014. *Parameters and micro-parameters in Arabic sentence structure*. Victoria University of Wellington. (Doctoral dissertation).

Kiparsky, Paul. 1994. Indo-European origins of Germanic syntax. In Adrian Battye & Ian Roberts (eds.), *Clause structure and language change*, 140–170. Oxford: Oxford University Press.

Kossmann, Maarten. 2017. Key and the use of Moroccan function words in Dutch internet discourse. *Nederlandse Taalkunde* 22(2). 223–248.

Lukassen, Lysette. 2011. *Zinsintroducerende elementen en woordvolgorde in het Nederlands: Interferentie effecten van het Marokkaans en het Turks*. Radboud Universiteit Nijmegen. (BA thesis).

MacWhinney, Brian & Catherine Snow. 1985. The Child Language Date Exchange. *Journal of Child Language* 12. 271–296.

Mohr, Sabine. 2009. V2 as a single-edge phenomenon. In Kleanthes K. Grohmann & Phoevos Panagiotidis (eds.), *Selected Papers from the 2006 Cyprus Syntaxfest*, 141–158. Newcastle-upon-Tyne: Cambridge Scholars Publishing.

Montrul, Silvina. 2016. *Heritage language acquisition*. Cambridge: Cambridge University Press.

Mourigh, Khalid. 2017. *Gouda Moroccan Dutch corpus*. Ms., Leiden University.

Opsahl, Toril. 2009. *"Egentlig alle kan bidra!": En samling sosiolingvistiske studier av strukturelle trekk ved norsk i multietniske ungdomsmiljøer i Oslo* [A collection of sociolinguistic studies of structural features associated with Norwegian as spoken among youth in multiethnic settings in Oslo]. University of Oslo. (Doctoral dissertation).

Quist, Pia. 2008. Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice. *International Journal of Bilingualism* 12(1-2). 43–61.

Reis, Marga. 2013. "Weil-V2"-Sätze und (k) ein Ende? Anmerkungen zur Analyse von Antomo & Steinbach (2010). *Zeitschrift für Sprachwissenschaft* 32(2). 221–262.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Springer.

Roberts, Ian. 1996a. Remarks on the Old English C-system and the diachrony of V2. *Linguistische Berichte Sonderhefte* 7. 154–167. DOI: 10.1007/978-3-322-90776-9_6.

Roberts, Ian. 1996b. Remarks on the Old English C-system and the diachrony of V2. In Ellen Brandner & Gisella Ferraresi (eds.), *Language change and generative grammar*, 154–167. Opladen: Westdeutscher Verlag.

Roberts, Ian. 2004. The C-system in Brythonic Celtic languages, V2, and the EPP. In Luigi Rizzi (ed.), *The structure of CP and IP* (The Cartography of syntactic Structures 2), 297–328. Oxford: Oxford University Press.

Schwartz, Bonnie D. & Rex Sprouse. 2000. When syntactic theories evolve: Consequences for L2 acquisition research. In John Archibald (ed.), 156–186. Oxford: Blackwell.

Shlonsky, Ur. 2000. Remarks on the Complementizer layer of Standard Arabic. In J. Lecarme, J. Lowenstamm & U. Shlonsky (eds.), *Research in Afroasiatic Grammar: Papers from the Third conference on Afroasiatic Languages, Sophia Antipolis, 1996*, vol. 202, 325–343.

Van Kemenade, Ans. 1987. *Syntactic case and morphological case in the history of English*. Dordrecht: Foris.

Walkden, George. 2017. Language contact and V3 in Germanic varieties new and old. *Journal of Comparative Germanic Linguistics* 20. 49–81. DOI: 10.1007/s10828-017-9084-2.

Wiese, Heike. 2009. Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua* 119(5). 782–806. DOI: 10.1016/j.lingua.2008.11.002.

Wiese, Heike. 2013. From feature pool to pond: The ecology of new urban vernaculars. *Working Papers in Urban Language and Literacies* 104. 1–29.

Wiese, Heike & Ines Rehbein. 2016. Coherence in new urban dialects: A case study. *Lingua* 172–173. 45–61. DOI: 10.1016/j.lingua.2015.10.016.

Wolfe, Sam. 2015. *Microvariation in Medieval Romance Syntax: A Comparative Study*. University of Cambridge. (Doctoral dissertation).

Wolfe, Sam. 2019. Redefining the typology of V2 languages: The view from Medieval Romance and beyond. *Linguistic Variation* 19(1). 16–46. DOI: 10.1075/lv.15026.wol.

Yang, Charles D. 2000. Internal and external forces in language change. *Language variation and change* 12(3). 231–250.

Zwart, Jan-Wouter. 1997. *Morphosyntax of verb movement: A minimalist approach to the syntax of Dutch*. Dordrecht: Kluwer. DOI: 10.1007/978-94-011-5880-0.

Zwart, Jan-Wouter. 2011. *The syntax of Dutch*. Cambridge: Cambridge University Press.

Chapter 17

Rethinking passives: The canonical GOAL passive in Dutch and its dialects

Liliane Haegeman
Ghent University

The main goal of this paper is empirical: it challenges the claim repeatedly found in the current generative literature (Alexiadou et al. 2014; Broekhuis & Cornips 2004; 2012) that Dutch lacks the GOAL passive. As will be shown, among other things, these claims fail to take into account the microvariation already reported in the earlier generative literature.

The paper contains a detailed discussion of the properties of GOAL passive in West Flemish, showing that, based on the standard diagnostics, the GOAL argument has acquired subject status in the passive. This conclusion thus provides a challenge for those accounts of Germanic passivization which are crucially based on the claim that English is the only West Germanic languages with a canonical GOAL passive (cf. Stein et al. 2016).

1 The typology of double object patterns

The cross-linguistic variation in passivization of double object patterns has recently been the source of renewed interest. It is sometimes claimed (most recently in Stein et al. 2016) that English is the only West Germanic language allowing for the passivization of the indirect object, illustrated in (1). The passive form in (1b) is variously referred to as the indirect object passive, the GOAL passive (Haddican & Holmberg 2012; 2015) or the RECIPIENT passive (Stein et al. 2016). I will use the label GOAL passive for convenience sake, as this term allows me to use the same term to refer to the constituent which functions as the indirect object in the active sentence and to the constituent that becomes the subject in the passive

Liliane Haegeman. 2020. Rethinking passives: The canonical GOAL passive in Dutch and its dialects. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 357–368. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280659

sentence.[1] Stein et al. claim: "the recipient passive arose in English but *not in other West Germanic languages*" (2016: slide 3, my italics). German has been reported not to have a canonical GOAL passive (2) (Anagnostopoulou (2003: 70); Alexiadou & Schäfer (2013: 9); Alexiadou et al. (2014: 10) for recent discussions). The claim that, like German, Dutch lacks a canonical GOAL passive, as shown in (3), is also common in the literature, as in, for instance, Broekhuis & Cornips (2004; 2012); Broekhuis et al. (2015); Alexiadou & Schäfer (2013: 8); Alexiadou et al. (2014: 10).

(1) Haddican & Holmberg (2012; 2015)
   a. They gave the girl the ball.
   b. *The girl* was given the ball.
   c. % *The ball* was given the girl.

(2) German (Alexiadou et al. 2014: 10)
   a. Sie hat dem Mann die Blumen geschenkt.
      she has the-DAT man the flowers given
      'She has given the man the flowers.'
   b. * *Er* wurde die Blumen geschenkt.
      he.NOM was the.ACC flowers given
      'He was given the flowers.'
   c. * *Die Blumen* wurden dem Mann geschenkt.
      the.NOM flowers were the.DAT Mann given
      'The flowers were given to the man.'

(3) Dutch (Alexiadou et al. 2014: 10)
   a. Ik heb hem het eten bezorgd.
      I have him the food delivered
      'I delivered the food to him.'
   b. * *Hij* werd het eten bezorgd (door mij).
      he was the food delivered (by me)
      'He was delivered the food by me.'
   c. *Het eten* werd hem bezorgd (door mij).
      the food was him delivered (by me)
      'The food was delivered to him by me.'

[1] I leave aside "non-canonical" passives such as the English *get* passive and the German/Dutch non- canonical *kriegen/krijgen* ('get') passives (Alexiadou & Schäfer 2013).

The goal of this paper is essentially empirical: it challenges the claim that English is the only West Germanic languages with a GOAL passive, and it challenges the specific claims made in the generative literature (Broekhuis & Cornips 2004; 2012) that Dutch lacks the GOAL passive. As I will show, among other things, such claims fail to take into account the microvariation reported in the earlier literature. The paper contains a detailed discussion of the GOAL passive in West Flemish.

2 The IO passive in West Flemish

2.1 The data: overview

As shown by the examples in (4) and (5), West Flemish (from now on WF), a dialect of Dutch and a West Germanic language, does have a GOAL passive: the definite GOAL, *Valère* in active (4a), has been promoted to become the subject of the passive sentence (4b). Similarly, the indefinite GOAL *nen student* ('a student') in active (5a) has been promoted to subject status in the passive (5b). The discussion in this section is based on my own dialect intuitions; the core intuitions are corroborated in Dhaenens (2014).[2]

(4) West Flemish

a. dan ze *Valère* die posten beloofd een
that.PL they Valère those jobs promised have
'that they promised Valère those jobs'

b. da *Valère* die posten beloofd wierd / is
that Valère those jobs promised 'became' is
'that Valère was promised those jobs'

(5) West Flemish

a. dan ze *nen student* die posten beloofd een
that.PL they a student those jobs promised have
'that they promised a student those jobs'

b. dat *(der) *nen student* die posten beloofd wierd / is
that ter a student those jobs promised 'became' is
'that a student was promised those jobs'

[2] A reviewer for this volume asks whether there are animacy effects for the double object pattern with verbs of motion, like those discussed by Haddican (2010). At first sight the effect is replicated in WF, but this issue needs further research.

Observe that the obligatory presence of expletive *(d)er* ('there') in (5b) is not a property specific to the GOAL passive. The obligatory presence of *(d)er* is fully in line with the patterns found elsewhere in (W)F: an indefinite or a quantified subject systematically requires that the sentence appear in the existential pattern with *(d)er*-insertion, as exemplified in active monotransitive (6a) or in passive monotransitive (6b).

(6) West Flemish

a. dan *(der) *drie* *studenten* dienen boek gelezen een
that.PL there three students that book read have
'that three students have read that book'

b. dan *(der) *drie* *studenten* betrapt zyn
that.PL there three students caught are
'that three students were caught'

§2.2 provides arguments to the effect that in WF GOAL passives, the GOAL argument is promoted to subject status. §2.3 shows that WF GOAL passives also comply with two specific diagnostics for Dutch passivization set out in Broekhuis & Cornips (2004; 2012), in particular with respect to the presence of an AGENT and the eventive interpretation.

2.2 Subject diagnostics for the GOAL passive

In the WF GOAL passives (4b) and (5b), the promoted GOAL acquires the syntactic properties of the WF subject, both when definite (4b) and when indefinite (5b) (for early diagnostics, cf. Haegeman 1986a,b).

2.2.1 Agreement

In the GOAL passive, the GOAL DP agrees for person and number with the finite verb and (in the relevant contexts) with the complementizer (7–8). (7a) illustrates a passive with a definite GOAL: the finite auxiliaries *wierden*/*woaren* ('were') are plural, as is the complementizer *dan* ('that'), and they thus can be seen to agree with the plural DP *de studenten* ('the students'). Neither complementizer nor auxiliary can be singular (7b–d). In (8a) agreement is triggered by the plural indefinite *drie studenten* ('three students'). Again the agreement is mandatory (8b–d). The patterns in (7) and (8) also entail that, in the passive sentences, singular agreement with the THEME *dienen bureau* ('that office') would be ungrammatical, cf. (7d) and (8d).

(7) West Flemish

a. dan *de studenten* dienen bureau beloofd wierden / woaren
that.PL the students that office promised were-PL
'that the students were promised that office'

b. * dat *de studenten* dienen bureau beloofd wierden / woaren
that.SG the students that office promised were-PL

c. * dan *de studenten* dienen bureau beloofd wierd / was
that.PL the students that office promised was-SG

d. * dat *de studenten* dienen bureau beloofd wierd / was
that.SG the students that office promised was-SG

(8) West Flemish

a. dan *(der) *drie studenten* dienen bureau beloof
that.PL there three students that office promised
wierden / woaren
were-PL
'that three students were promised that office'

b. * dat *(der) *drie studenten* dienen bureau beloofd
that.SG there three students that office promised
wierden / woaren
were-PL

c. * dan *(der) *drie studenten* dienen bureau beloofd wierd / was
that.PL there three students that office promised were-SG

d. * dat *(der) *drie studenten* dienen bureau beloofd wierd / was
that.SG there three students that office promised were-PL

2.2.2 Case

When pronominal, the GOAL DP is realised as a nominative, and, like other nominative pronouns, it allows for pronoun doubling. In (9a) the strong nominative pronoun *zie* is a doubler for the weak form *ze*. For full discussion of WF subject pronouns I refer to my earlier work (Haegeman 1990; 1992; 2004). In the Flemish regiolect, the subject of the GOAL passive can be the impersonal pronoun *men* ('one'), which is restricted to subject position of a finite clause (9b).[3]

[3]This property cannot be tested for the dialect because the impersonal pronoun *men* is not used.

(9) West Flemish
   a. da *ze (zie)* die posten beloofd wierd
      that she (she) those positions promised was
      'that she was promised these jobs'
   b. Het komt veel voor dat *men* die behandeling afgeraden
      it comes often for that one that treatment disrecommended
      wordt.
      is
      'It is quite common that one is advised against that treatment.'

2.2.3 Relativization

Like canonical definite subjects, relativized GOAL DPs are associated with relativizer *die* (10a) and with *dat/die* alternations (10b). These properties are characteristic of subject relativization in WF (10b), and they are unavailable in object relativization (10c). See Haegeman (1984; 1992).

(10) West Flemish
   a. Dat zijn de studenten *dien* die posten beloofd woaren.
      that are the students *die*-PL those jobs promised were
      'Those are the students that were promised those jobs.'
   b. Dat zijn de studenten dan-k peinzen *dien* die posten beloofd
      that are the students that-I think *die*-PL those jobs promised
      woaren.
      were
      'Those are the students that I think were promised those jobs.'
   c. Dat zijn de boeken dan-k peinzen da / *die Valère besteld eet.
      that are the books that-I think that / *die* Valère ordered has
      'Those are the books that I think that Valère has ordered.'

2.2.4 Existential patterns

When the GOAL is an indefinite nominal (5b), a numeral (11a) or a *wh*-constituent (11b), and is promoted to becoming the subject of the passive, *(d)er*-insertion is obligatory.

(11) West Flemish
   a. dan *(der) ∅ / *drie studenten* dienen post beloofd zyn
      that *ter* three students that job promised are
      'that (three) students were promised that job'

b. Kweeten niet *wien* dat *(er) dienen post beloofd is.
I know not who that there that job promised is
'I don't know who was promised that job.'

Obligatory *(d)er*-insertion is associated with indefinite or quantified subjects and not with objects.

2.2.5 Distribution

Like canonical definite subjects, the definite GOAL DP in the GOAL passive has to be linearly adjacent to the complementizer *dat* ('that')[4] in embedded clauses (12) and to the finite verb in root clauses (13). In (12a), adjuncts such as *gisteren* ('yesterday') or *verzekerst* ('probably') cannot intervene between the complementizer *dat* ('that') and the GOAL *Valère*. In (12b), the THEME *die posten* ('those jobs') cannot intervene between the complementizer *dat* ('that') and the GOAL *Valère*. In (13), the same adjacency requirement is illustrated for root clauses in which the finite verb, here the auxiliary *wierd* ('was'), has moved to C. (14) and (15) show that identical adjacency restrictions apply to definite subjects of transitive sentences.

(12) West Flemish

a. dat (*gisteren / verzekerst) *Valère* die posten beloofd wierd
that yesterday probably Valère those jobs promised was
'that Valère was (probably) promised those jobs (yesterday).'

b. * dat die posten *Valère* beloofd wierd
that those jobs Valère promised was

(13) West Flemish

a. Daarom wierd (*gisteren / verzekerst) *Valère* die posten
for that reason is yesterday probably Valère those jobs
beloofd.
promised
'For that reason, Valère was (probably) promised those jobs (yesterday).'

b. * Daarom wierd die posten *Valère* beloofd.
for that reason was those jobs Valère promised

[4] In WF the complementizer *dat* is obligatorily present in all embedded clause, frequently leading to doubly filled Comp positions.

(14) West Flemish

a. dat (*gisteren / verzekerst) *Valère* die posten beloofd eet
that yesterday probably Valère those jobs promised has
'that (probably) Valère promised those jobs (yesterday).'

b. * dat die posten *Valère* beloofd eet
that those jobs Valère promised has

(15) West Flemish

a. Daarom eet (*gisteren / verzekerst) *Valère* die posten
for that reason has yesterday probably Valère those jobs
beloofd.
promised
'For that reason, Valère probably promised those jobs (yesterday).'

b. * Daarom eet die posten *Valère* beloofd.
for that reason has those jobs Valère promised

2.2.6 Non-finite clauses

The GOAL passive is available in non-finite control clauses, in which case the GOAL will be a controlled PRO (16a). The goal subject of a passive clause may undergo raising in *te* infinitives (16b).

(16) West Flemish

a. Me [PRO] dienen anderen post beloofd te zyn, goa-se niet
with that other job promised to be goes-she not
veruzen.
move
'Having been promised that other job, she's not going to move house.'

b. Ze pleegdege zie zukken medicamenten voorengeschreven te zyn.
she used she such medications prescribed to be
'She used to be prescribed that medication.'

2.2.7 Coordination

That it is the GOAL nominal which is promoted to subjecthood in the GOAL passive is confirmed by coordination data. For instance, an active clause can coordinate with a GOAL passive clause under one shared subject (17a); a clause with a THEME passive of a transitive verb can coordinate with a GOAL passive clause under one shared subject DP (17b).

(17) a. dan die twee studenten nor us mochten en da medicament
that those two students to home might and that medication
neu niet meer voorengeschreven goan worden.
now no more prescribed go be
'that those two students were allowed to go home and now will no longer be prescribed that medication.'

b. da Valère eerst vur een interview utgenodigd is en doa toen
that Valère first for an interview invited is and there then
dienen post beloofd is
that job promised is
'that Valère was first invited for an interview and was promised the job there.'

2.3 The AGENT in the GOAL passive

As in other passive sentences, in a GOAL passive sentence, the AGENT can be overtly expressed (18).

(18) West Flemish
dan-k dienen velo aangeraden zyn *door twee collega's*
that-I this bicycle recommended am by two colleagues
'that I was recommended that bike by two colleagues.'

An implied AGENT can be modified by an adjunct: in (19), *per ongeluk* ('unintentionally') or *espres* ('intentionally') modify the understood AGENT.

(19) West Flemish
dat Valère *per ongeluk* / *espres* te vele cortisonepillen
that Valère by accident intentionally too many cortisone.pills
voorengeschreven wier
prescribed was
'that Valère was prescribed too many cortisone pills by accident / intentionally.'

2.4 Event passive

Based on the diagnostics in Broekhuis & Cornips (2004; 2012), I conclude that the WF GOAL passive can have an eventive reading both with the auxiliary *worden* ('become') and with the – probably much more common – alternative *zijn* ('be'). Temporal specifiers modifying the event time are compatible with the GOAL passive (20).

(20) West Flemish
dat Valère *gisteren* te vele cortisonepillen voorengeschreven is
that Valère yesterday too many cortisone.pills prescribed is
'that Valère was prescribed too many cortisone pills yesterday.'

2.5 Conclusion: WF has a GOAL passive

All the diagnostics discussed above converge and point clearly towards the conclusion that WF, a dialect of Dutch and a West Germanic language, has a productive GOAL passive, contrary to claims in the current generative literature.

Whether the emergence of the GOAL passive in WF can also be attributed to contact with French, as argued for English by Stein et al. (2016), is a question that needs to be addressed. It is true that the WF lexicon provides strong evidence of contact of French as shown in Haegeman (2009). An alternative hypothesis might be that the emergence of the GOAL passive is due to Ingvaeonic influence (see Dhaenens 2014). I do not further speculate on this issue here.

3 Conclusion

This paper provides empirical evidence against persistent claims in the formal literature to the effect that English is the only West Germanic language with a GOAL passive, showing that at least the West Flemish dialect of Dutch has a productive canonical GOAL passive. The WF data strongly challenge the claims in the current literature that Dutch lacks a canonical GOAL passive, since at least one Dutch dialect does display the pattern.

Abbreviations

ACC	accusative	PL	plural
DAT	dative	SG	singular
NOM	nominative	WF	West Flemish

Acknowledgements

I dedicate this paper to Ian. I have known Ian from the earliest stages of his career and I have been lucky enough to be able to work with him in Geneva. I admire the tenacity with which Ian has continued to rethink the linguistic themes that had initially preoccupied him in his early research and the way in which his research has developed into a full-fledged research programme that allows us to attain a deeper understanding of core issues of comparative syntax.

References

Alexiadou, Artemis, Elena Anagnostopoulou & Christina Sevdali. 2014. Opaque and transparent datives, and how they behave in passives. *The Journal of Comparative Germanic Linguistics* 17(1). 1–34. DOI: 10.1007/s10828-014-9064-8.

Alexiadou, Artemis & Florian Schäfer. 2013. *Non-canonical passives*. Amsterdam: John Benjamins. DOI: 10.1075/la.205.

Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Berlin: De Gruyter.

Broekhuis, Hans & Leonie Cornips. 2004. Undative constructions. *Linguistics* 32(2). 173–189. DOI: 10.1515/ling.1994.32.2.173.

Broekhuis, Hans & Leonie Cornips. 2012. *Linguistics* 50(6). 1205–1249. DOI: 10.1515/ling-2012-0039.

Broekhuis, Hans, Norbert Corver & Riet Vos. 2015. *Syntax of Dutch: Verbs and verb phrases*. Vol. 1. Amsterdam: Amsterdam University Press.

Dhaenens, Gilles. 2014. Perspectivische variatie van de actieve dubbelobjectconstructie: passieve en receptieve varianten. BA paper, Department of Linguistics, Ghent University.

Haddican, William. 2010. Theme–goal ditransitive constructions and theme passivisation in British English dialects. *Lingua* 120(10). 2424–2443. DOI: 10.1016/j.lingua.2009.11.003.

Haddican, William & Anders Holmberg. 2012. Object movement symmetries in British English dialects: Experimental evidence for a mixed case/locality approach. *Journal of Comparative Germanic Linguistics* 15(3). 189–212. DOI: 10.1007/s10828-012-9051-x.

Haddican, William & Anders Holmberg. 2015. Four kinds of object asymmetry. In Ludmila Veselovská & Marketa Janebová (eds.), *Complex visibles out there: Proceedings of the Olomouc Linguistics Colloquium 2014: Language use and linguistic structure*, 145–162. Olomouc: Palacký University.

Haegeman, Liliane. 1984. *Die* and *dat* in West Flemish relative clauses. In Hans Bennis & W.U.S. Van Lessen Kloeke (eds.), *Linguistics in the Netherlands*, 83–91. Dordrecht: Foris.

Haegeman, Liliane. 1986a. Passivisation in Flemish dialects. In Frits Beukema & Aafke Hulk (eds.), *Linguistics in the Netherlands*, 71–78. Dordrecht: Foris.

Haegeman, Liliane. 1986b. The double object construction in West Flemish. *The Linguistic Review* 5. 281–300. DOI: 10.1515/tlir.1986.5.4.281.

Haegeman, Liliane. 1990. Subject pronouns and subject clitics in West-Flemish. *The Linguistic Review* 7(4). 333–364. DOI: 10.1515/tlir.1990.7.4.333.

Haegeman, Liliane. 1992. *Theory and description in generative grammar: A case study in West Flemish*. Cambridge: Cambridge University Press.

Haegeman, Liliane. 2004. Verdubbeling van subjectpronomina in de Zuid-Nederlandse dialecten: Een reactie uit Lapscheure. *Taal en Tongval* 56. 119–159.

Haegeman, Liliane. 2009. *Linguistic variation in Belgium*. Paper presented at the Meertens Institute. http://www.gist.ugent.be/file/172.

Stein, Achim, Carola Trips & Richard Ingham. 2016. *The role of French in the rise of the recipient passive in Middle English: Can structural case be borrowed?* Paper presented at the 18th Diachronic Generative Syntax conference, 28 June–1 July 2016, Ghent University.

Chapter 18

Extraordinary second-position effects

Moreno Mitrović
ZAS Berlin and Bled Institute

Thanks to Roberts (2010), the second-position (2P) effect is given a natural explanation using narrow-syntactic utilities alone, resting on his notion of defectivity. In this paper, I review and extend a narrow-syntactic approach to some other types of 2P effects that have, as far as I know, not been studied in tandem; particularly extraordinary 2P effects involving a combination of 2P placement and left branch extraction (LBE).

1 Introduction

Thanks to Roberts (2010), the second-position (2P) effect is given a natural explanation using narrow-syntactic utilities alone, resting on his notion of defectivity. In this paper, I review and extend a narrow-syntactic approach to some other types of 2P effects that have, as far as I know, not been studied in tandem; particularly extraordinary 2P effects involving a combination of 2P placement and left branch extraction (LBE).

There is no single treatment and theory of all 2P effects: 2P typology comprises at least three classes, based on the categorial *size* properties of the 1P prima facie "hosting" element. The first is the one where the host is a maximal category – these constructions are exemplified by verb-second (V2) or LBE phenomena. The second type involves a host of minimal category and are demonstrated by V-fronted constructions (e.g., long head movement in Breton, V-topicalisation in Slavonic, etc.). Both these types are discussed on a par and given a uniform treatment in Roberts (2010). The last type features non-constituent hosts comprising of a head, say a preposition, and a maximal category, say an AP. This last type is incarnated by what Bošković (2005) calls extraordinary LBE (XLBE). It is this last

Moreno Mitrović. 2020. Extraordinary second-position effects. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 369–401. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280661

type that is most resistant to narrow-syntactic explanation and, as far as I can gather from the literature, no definitive and purely syntactic account has been proposed.

I aim to derive the last type of 2P effect using Chomsky's (2001) triadic characterisation movement that Roberts (2010: 208) restated in parametric format (1):

(1)	Move	Agree	Pied-pipe	
a.	+	+	+	A-movement
b.	+	+	−	incorporation
c.	+	−	+	Ā-movement
d.	−	+	−	Agree
e.	−	−	−	∅
f.	−	+	+	*
g.	−	−	+	*
h.	+	−	−	Ā-incorporation

If all three operations apply in tandem, A-movement obtains, while a combination of Move and Pied-piping along yield Ā-movement (with the absence of an Agree operation in Ā-processes being highly problematic). Head movement, on the other hand, can be seen as deriving from a combination of Agree and Move. While options (f) and (g) are impossible, by virtue of the axioms of Minimalist syntax (Collins & Stabler 2016), Roberts (2010) takes the last option as corresponding to predicate clefting or Ā-incorporation.[1] This paper shows that this last movement operation derives XLBE.

Roberts (2010: 421) defines intrinsic formal features (IFFs) on terminals in the clausal spine, which are provided in Table 18.1 along with corresponding IFFs in the nominal domain.

Table 18.1: Intrinsic formal features (IFFs)

in the verbal domain				in the nominal domain		
C^{min}	[*i*C]	[*i*T]	[*i*V]	P^{min}	[*i*D]	[*i*N]
T^{min}		[*i*T]	[*i*V]	D^{min}	[*i*D]	[*i*N]
v^{min}			[*i*V]	n^{min}		[*i*N]
V^{min}			∅	N^{min}		∅

[1]For further empirical evidence of Ā-incorporation, see Mitrović (2017b) and those he cites.

I assume that prepositions have no IFF other than N and D. By adopting the view that the presence of the (phasal) D head is subject to cross-linguistic parametrisation, languages lacking the D-structure will correspondingly have prepositions with only one IFF, i.e. N.

The remainder of this section is devoted to explicating some background assumptions and introducing the relevant discussion within which the analysis is couched. After a brief survey of explananda for 2P effects (§1.1), the preliminary details of the N/D parameter of Bošković (2005; 2008), which I am going to assume, are given in §1.2. Finally, in §1.3, I outline the defectivity system of Roberts (2010) that underlies the account proposed here. §1.4 provides the reader with directions I take in the following sections.

1.1 The 2P effect and its explananda

There are two general stances to explaining cliticisation phenomena. By the end of this subsection, I hope to demonstrate that one of these approaches should be preferred on both theoretical and empirical grounds.

One of the foundational questions concerning 2P cliticisation phenomena is: Where does cliticisation take place? At least two answers have been around for decades: either cliticisation configurations are established and derived in narrow syntax (NS) or, otherwise, they are epiphenomenal and reflective of post-syntactic (or more precisely phonological or prosodic) displacement and rearrangement. Let me briefly lay out a two-tiered motivation for preferring the former over the latter.

A phonological/prosodic (i.e., "anti-syntactic") motivation for second-position (2P) cliticisation is most notably and influentially characterised by the theory of prosodic inversion (PI) as advocated by Halpern (1992; 1995). As Roberts (2012: 422) notes, there are three ingredients to this theory as given in (2).

(2) i. 2P clitics are prosodically subcategorised to appear right-adjacent to a prosodic word;
ii. clitics adjoin to IP;
iii. where no element with a phonological matrix appears to the left of the IP-adjoined clitic, then PI must apply, in line with (3).

(3) CLITIC $> X > Y \longrightarrow X >$ CLITIC $> Y$

Given a relevant prosodic domain, the clitic and the rightmost element thus prosodically flip and the second-position effect obtains (3), in line with the principles in (2). Note, however, that (3) is a sketch and there are certainly works

within this approach where 2P clitics are located in positions other than IP. (For a detailed overview and a summary of all relevant arguments, I refer the reader to Bošković 2001: 75ff and citations therein.)

Let me now review some arguments that undermine the nature of such principles.[2] Firstly, with respect to (2i), the 2P order may be derived using more general syntactic principles, as I will demonstrate. Additionally, categorising an element as, and assigning it a descriptively arbitrary label of, a clitic is extraneous insofar as the "clitic effect" may arise from the configuration of the clitic with respect to other elements, especially its "host". Secondly, and in connection to (2ii), it is not only stipulative but also counter-theoretical to assume that clitics adjoin to IP. On the one hand, the current minimalist model of phasal syntax demonstrably takes the C^0, and not the T^0, head to be a phase head and, as such, the locus of clitic-clustering should be on phase heads, i.e. C^0 and v^0 (I demonstrate the conceptual and empirical connection between cliticisation target sites and the phasal nature of such sites below but see Roberts (2010; 2012) for a detailed account and motivation). An additionally problematic conception of (2ii) concerns the nature of "adjunction" which cannot be maintained in line with the standard assumptions of syntax. This proviso of PI predicts all clitics to either be base generated at IP-level or internally moved to an IP-level adjunct position. Consider empirical instances of DP-level conjunction clitics in Indo-European (e.g. Latin *-que*, or Hittite *-a*) or, say, object clitics in Romance or South Slavonic in relation to this proviso. The amount of stipulation that would ensue if I assume there exists movement of a DP conjunction in the former example or object DP in the latter in order to render the syntactic conditions for PI to apply, in line with (2ii), would be too great for a theory of syntax to remain consistent.

On a more general level, the existence of a structure-tampering operation, such as PI as formulated above, breaches the basic tenets of the minimalist linguistic theory or, at least, cannot be defined in accordance with the general minimalist assumptions. Since the Merge operation derives syntactic structures and the nature of movement operations, it has to be confined to the core syntactic module of grammar. I thus cannot maintain this theoretical principle and expect to find displacement operations, derived by Merge, outside the modular confines of syntax.

A less general and more damaging evidence against PI is empirical. I briefly provide an argument coming from Ser-Bo-Croatian LBEs. Bošković (2009), among others, convincingly shows that PI cannot account for the following morphosyntactically conditioned violations of the left branch condition (LBC). While

[2] In doing so, I also adopt the rationale of Roberts (2012: 422).

non-extracted DPs containing both forenames and last names allow the forename to be unmarked for case, a left-branch extracted forename must obligatorily be case-marked; in the case of (4), as an accusative.

(4) Ser-Bo-Croatian

a. i. Lav-a Tolstoj-a sam čitao
Leo-ACC Tolstoy-ACC AUX.1SG read.PPL.SG.M
'I'm reading Leo Tolstoy.'

ii. Lav Tolstoj-a sam čitao
Leo-NOM/∅ Tolstoy-ACC AUX.1SG read.PPL.SG.M
'I'm reading Leo Tolstoy.'

b. i. Lav-a sam Tolstoj-a čitao
Leo-ACC AUX.1SG Tolstoy-ACC read.PPL.SG.M
'I was reading Leo Tolstoy.'

ii. * Lav sam Tolstoj-a čitao
Leo-NOM/∅ AUX.1SG Tolstoy-ACC read.PPL.SG.M
'I was reading Leo Tolstoy.'

If some post-syntactic algorithm did in fact derive PI, it is nearly impossible to account for the empirical facts stated above without having the phonological-prosodic module of grammar be sensitive to narrow morphosyntactic properties or features such as case marking.

Also consider the fact that it is not clitics alone that may interrupt a complex DP, such as the "Leo Tolstoy"-type compounds names above. As Bošković (2009) observes, a non-clitic item, such as a full finite lexical verb *čitam* 'read.1SG.PRS', may also break up the name (5). In line with Roberts (2012), I assume that the first-name D^{max} Ā-moves to the position of Spec(Force^{max}) with the full verb remaining in T^{min}. Note further the obligatory case-marking on the extracted forename DP.

(5) Ser-Bo-Croatian
Lava čitam Tolstoja
Leo.ACC read.1SG.PRS Tolstoy.ACC
'I'm reading Leo Tolstoy.'

Furthermore, the following is also well-formed, which lends empirical support to Roberts's (2010) motivation that Ā-movement of minimal categories should exist. The continued range of cases of clitic interruptions of the first-last-name DP should amplify empirically this argument.

(6) Ser-Bo-Croatian

a. (?) Lava sam čitao Tolstoja
Leo.ACC AUX.1SG read.PPT.SG.M Tolstoy.ACC
'I (have) read Leo Tolstoy.'

b. (?) Lava čitao sam Tolstoja
Leo.ACC read.PPT.SG.M AUX.1SG Tolstoy.ACC
'I (have) read Leo Tolstoy.'

(7) Ser-Bo-Croatian

a. Lava mi je Tolstoja dao da čitam
Leo me.DAT is Tolstoy gave that read.1SG.PRS
'He gave me Leo Tolstoy to read.'

b. Lava sam joj Tolstoja dao da čita
Leo am her.DAT Tolstoy gave that read.2SG.PRS
'I gave her Leo Tolstoy to read.'

c. Lav si je Tolstoj (sam) doručak pravio
Leo self.DAT is Tolstoy (himself) breakfast made
'Leo Tolstoy (himself) made himself breakfast.'

Note that some speakers concede that (6b) is degraded without a pause following *Lava*. The requirement for the pause is captured prosodically by a generalisation that Ser-Bo-Croatian 2P clitics must be second within their intonational phrase (Bošković 2001: 65, n. 120). The account I provide is consistent with this generalisation as I advocate a view that NS movement coincides with intonational phrasing.

If the theory of PI cannot account for the contemporary LBE phenomena found in Ser-Bo-Croatian, I inductively find it untenable to entertain this theory as general explanandum applicable to a cross-linguistic patterns of cliticisation which also display LBC violations. On grounds of both theoretical and empirical motivation, I thus pursue a NS aetiology of cliticisation, also for reasons of more general parsimony, as noted by Roberts (2010: 73–74); namely I choose, and logically prefer, not to accord extra-syntactic factors too prominent a role in order to maintain the approach in full generality. It is thus, ceteris paribus, more theoretically consistent to adhere to the central syntactic account and derive a maximally possible account of the distribution of facts from that.

More specifically, since a NS account of cliticisation does not suffer from the two drawbacks stated above, I am lead to maintain this assumption in the analysis.

1.2 The N/D parameter

With background notions in place, I discuss in the remainder of the paper how the relation between the N/D parameter and the system of defectivity can be married in an analysis of XLBE.

Assuming that D^{max} constitutes a phase, Bošković (2005) provides an account of why some languages allow and others disallow LBE.[3] Given that D^{min} is a phase head, it prohibits movement of its complement with only its edge being accessible as per the PIC. His first assumption is that languages like Ser-Bo-Croatian lack the D-layer in their nominal spine and, due to this, lack a nominal phase, making their interior accessible. His second assumption is that adjunction structures come in two parametric options: either the adjective takes an NP complement (AP-over-NP) or the AP is adjoined to NP (NP-over-AP).

Consider a scenario of AP-extraction in English which is barred due to the presence of the phasal D. In order for AP to extract, it must pass through D's edge, i.e. Spec(D^{max}). This, however, is an anti-local move and thus prohibited by the independently motivated principles of grammar. Thus, the combination of the PIC and anti-Locality bans LBE in D-containing language like English.

By contrast, Bošković (2005; 2008) contends that Ser-Bo-Croatian is a D-less language in which nominals are not phasal, hence the PIC is inapplicable. Consequently, there is no need for anti-local moves of the AP since the AP may immediately and directly extract to the final position. This is the line of reasoning I will adopt on both empirical and theoretical grounds.

1.3 Defectivity

The second and more foundational is the assumption surrounding triggers of head-movement. Roberts's (2010) system predicts incorporation to take place where an Agree relation holds between a probe and a goal such that the formal features of the goal form a proper subset of the features specified on the probe. This constitutes the goal as defective and such goals incorporate. The concept of defectivity thus regulates movement of the minimal category.

(8) DEFECTIVITY (Roberts 2010)
A goal G is defective iff G's formal features are a proper subset of those of G's probe P.

Thus, in more formal terms, a set of formal features (F) on a minimal category that enters an Agree relation as a Probe (P) will incorporate the Goal (G) iff (9) is met.

[3] See Bošković (2013) for a more recent and phase-based discussion of LBE.

(9) $F_G \subset F_P$

For instance, Romance pronominal objects clitics are taken to correspond to $\varphi^{min/max}$, lacking a D feature. The v^{min}, bearing an IFF [*i*V] (Table 18.1), probes for valuing its [*u*φ]. Upon valuation, $F_G \subset F_P$ holds and the object $\varphi^{min/max}$ incorporates into v^{min}. As Roberts (2012: 391) further notes, "[t]his means that the Match relation holding in virtue of Agree causes the host to become a featural copy of the probing features of the host." The chain reducing algorithm that applies post-syntactically, and which ensures economical assignment of phonological indices, will treat the host-probe and the defective clitic-goal as a single feature bundle. Thus, for a chain

$$\langle [G+P], t_G \rangle$$

the algorithm will pronounce the head of the chain only, giving the effect of movement.

By contrast to Romance, Slavonic clitics are not *v*-oriented but cluster in the C-domain. Roberts (2010) derives the C-orientation by positing that Slavonic clitics are not $\varphi^{min/max}$ elements (since they would be *v* incorporating otherwise) but $D^{min/max}$. Since v^{min} has no uninterpretable D-feature, these clitics can thus escape incorporating into *v*.[4] By virtue of C's bearing an uninterpretable D-feature, pronominal $D^{min/max}$ elements (as well as D-bearing auxiliaries sitting in T^{min}) cliticise onto C.

In conclusion to this section, consider the apparent contradiction that arises in our assuming the systems of Roberts (2010) and Bošković (2005). For Roberts (2010), it is critical that pronominal clitics in Ser-Bo-Croatian be $D^{min/max}$. For Bošković (2005), on the other hand, Ser-Bo-Croatian has no D category. I propose to reconcile the two approaches, in their assumptions and conclusions, by treating Ser-Bo-Croatian pronominal clitics not as D elements but as making up $N^{min/max}$. To maintain the defectivity approach of Roberts (2010), I take the C^{min}, conversely, as being specified with a [*u*N].

This view of subsuming the N/D parameter alongside a defectivity-based system of explananda which require me to adjust some of the basic assumptions and tenets of Roberts (2010). As preliminarily discussed in the following subsection, this is a fully compatible view which expands the explanatory adequacy of the defectivity approach and helps resolve XLBE.

[4]On the escape system, see Roberts (2012: 391–392) and references there.

1.4 Desiderata and roadmap

In the previous two subsections, two seemingly orthogonal ideas were laid out: a parametric and a presumably universal one. The former concerns the choice between encoding arguments as N- or D-elements. The latter concerns defectivity conditions defined on an Agree operation between objects bearing formal features which, when met, legislate incorporation of the goal into the probe.

The two views, while appealing to different derivational devices and conditions, are seemingly incompatible as one assumes that clitics are D-elements (Roberts 2010) while another opposes this view (Bošković 2009).

The primary desideratum is to derive a narrow-syntactic analysis of the word-first 2P effect by suggesting that the effect derives from constituent-only consideration, as opposed to (linearity-based) word-level "counting" which phonological explananda suppose.

Secondarily, I will restate the N/D parameter in terms of the defectivity technology that applies to a pair in an Agree relation, rather than general structural edge- or barrier-based restrictions on extraction domains. This will show that the N/D split theory is compatible with the defectivity approach to head movement.

The scope of this paper is largely restricted to achieving the first desideratum, with the second one requiring apparent abandonment of the assumptions made in the previous subsection, especially in connection to defectivity. §5, however, outlines a resolution for the question of how the defectivity approach may be integrated with the N/D parameter.

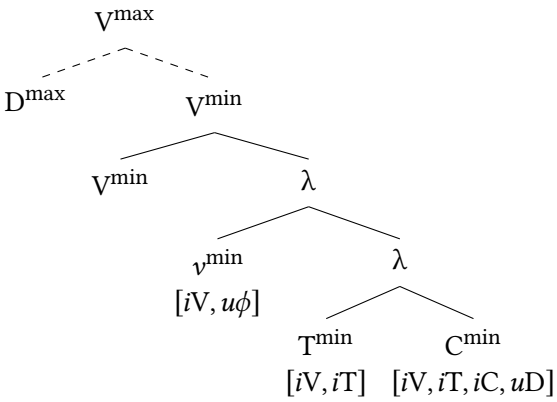
In §2 I outline a technical assumption which will allow me to combine the N/D and cliticisation parameters. In §3, a second position typology is presented with the empirical core of XLBE, which is analysed in §4. §5 provides a programmatic post hoc outlook on rectifying the counterintuitive assumption on the internal structure of clitics in South Slavonic. I essentially appeal to a parametric recasting of the nature of the relevant IFF in pronominal clitics which would yield the two core taxonomies, C- and *v*-oriented clitics, while retaining the view that South Slavonic pronominal clitics are not D-elements, in line with the tenets of Bošković (2001; 2004; 2005; 2009). The following section first provides another crucial piece of technology I rely on in order to derive a narrow-syntactic analysis of XLBE.

2 The unrolling spine: Shimada (2007)

While my account rests on the notion of defectivity as underlying narrow-syntactic incorporation as per Roberts (2010), I add another theoretical ingredient.

I follow Shimada (2007) in assuming that the clausal spine in fact results from a successive unrolling or excorporation of a head verbal complex that contains the entire clausal extended projection (cf. Saito 2012). I assume that the label every branching non-root node in the head-complex lacks the label (λ). I define on the clausal terminals their IFFs along with the [*u*φ] and [*u*D] at phasal levels of v^{min} and C^{min}, respectively (in line with Roberts 2010).

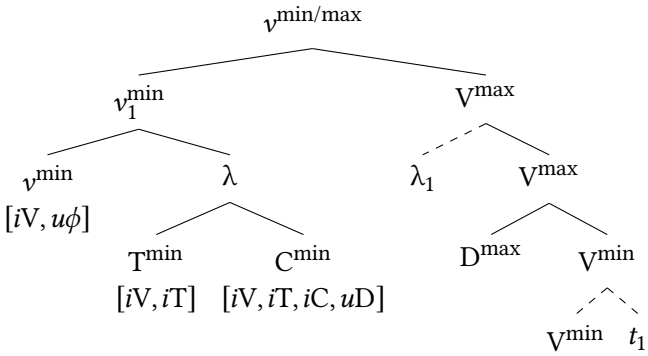
(10)



Note that prior to excorporation of Compl(V^{min}) in (10), there is only one pair of terminals satisfying the defectivity condition on incorporation: T^{min} and C^{min}. However, the *linear correspondence axiom* (LCA) prohibits such movement, making incorporation inapplicable at this stage.

Once the V has combined with an argument, say D^{max} (which has undergone spine-unrolling), its complement, headed by v^{max}, excorporates to the root for two reasons: semantically, there is a type-mismatch (hence the λ) and, perhaps more importantly for our syntactic purposes, Complement(V^{min}) is lacking a label. Once it excorporates, the c-selecting head, v^{min} projects the label (11).

(11)



Given the strong cycle, V^{min}-incorporation takes place as well as External Merge of the argument, checking [$u\varphi$] on v^{min}. In the next derivational step, the remaining λ-complex containing T^{min} and C^{min} excorporates for the same reasons I gave earlier. The result, after subject raising (SBJ) and final excorporation of C^{min} from the T-complex, viz. the structure in Figure 18.1.

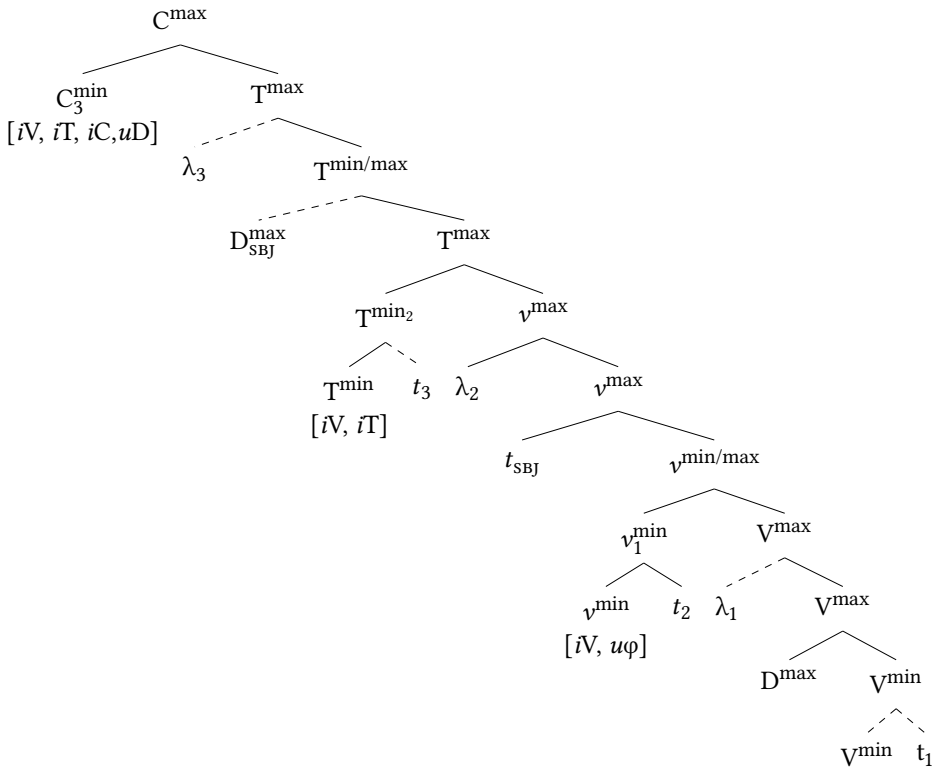


Figure 18.1: A clause-unfolding analysis utilising successive excorporation (Shimada 2007)

The resulting derivation is identical to the standardly assumed one, hence standard operations, including A- and Ā-processes, apply. I will tacitly assume in the remainder of the paper that the spine unrolls along the lines just sketched and, therefore, use a traditional and simplistic drawing of the trees. In §4, the details of the assumptions concerning the excorporational onset of derivations will become clear.

3 Deriving the phrase-/head-first 2P effect

In this section, I provide a derivational account of constituent-first 2P effects. In §3.1, I sketch an account of Wackernagel effects found across old IE conjunction structures which feature a minimal category as the host of enclisis. I turn to hosts of maximal categories in §3.2, and, lastly, to a phenomenon which seems to alternate between phrase/head-first in Slavonic in §3.3.

Note however, that the empirical locus of paper lies in XLBE (§3.3). While other phenomena, including v2 and V-topicalisation may well be analysed using the same principles of the derivation I adopt and propose, these fall outside of the scope of the present paper.[5]

3.1 X-first

Word-first constructions are a wide-spread phenomenon in old IE coordination structures and were first described by Wackernagel (1892). I cite below three examples from Old Irish (12), Gothic (13) and Old Avestan (14). [6]

(12) Old Irish; *Laws*, 4.179 (Thurneysen 2003)
... ba **ċ** ri Temrach
COP and king Tara.GEN
'And he was king of Tara.'

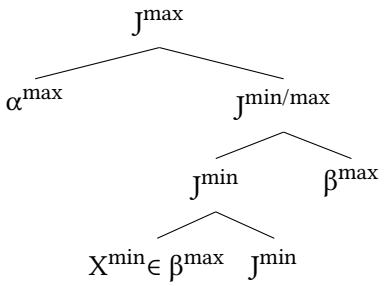
(13) Gothic; *Codex Argenteus*, Jn. 18:33
wopida Iesu qaþ **uh** imma.
called.PRT.3SG J.ACC said.PRT.3SG and him.M.DAT.SG
'(Then Pilate entered into the judgment hall again, and) called Jesus, and said unto him.'

(14) Old Avestan; *Yasna Haptanghāiti*, 29.10
yūžəəm aēibiiō ahurā aogō
you.2.SG.NOM them.PL.DAT lord.M.SG.VOC strength.N.SG.ACC
dātā aṣā xṣaϑrəm **cā**
give.2.PL.AOR.IMP truth.N.SG.INST power.N.SG.ACC and
'O Lord, may you give strength to them through truth and that power [...]'

[5]For an analysis of v2, compatible with the spine-unrolling tenets, see Shimada (2007: Ch. 2). For an analysis of V-topicalisation, see Ćavar & Wilder (1994), Mitrović (2017a), among others.
[6]For a detailed view, see Mitrović (2014; 2021), and references therein.

The common pattern that emerges in these coordinate constructions is that there is exactly one word preceding the conjunction maker. Assuming a J(unction) structure, I take this one-word precedence to derive from head-movement from within the internal (second) conjunct:

(15)



Coordination structures of this type are semantically unmarked across all old IE languages. Since incorporation into the coordinator is consistently blind to the category of the incorporee, Ā-incorporation would appear as the best candidate for an explanandum. This would require positing some Ā-feature such as [edge feature (EF)] on J^{min}, making it phasal in nature. Assuming that it lacks a categorial label (see Chomsky 2013, inter alia), J^{min} has some IFF and an uninterpretable categorial feature which is checked via c-selection. Note that its bearing an uninterpretable feature makes J^{min} potentially phasal in nature.[7]

An alternative view to Ā-incorporation would be to adopt an Agree-based account of incorporation. Assume J has no [EF] specified, but does have a category feature without a value, as per standard assumptions. Once valued, every accessible minimal category in Compl(J^{max}) is a defective goal and the closest one undergoes incorporation. (For a synchronic and diachronic account of the syntax of coordination in IE, see Mitrović 2014; 2018; 2021.)

Similar 2P effect with a minimal category can be observed in Slavonic. Unlike the Wackernagel data above, it is the pronominal clitics that undergo movement by virtue of their being defective goals. In Slavonic, pronominal clitics are treated as $D^{min/max}$ which are probed by a [*u*D]-carrying C (more precisely, Fin^{min}). Once incorporated, the C's [EF], specified presumably on $Force^{min}$, is checked via Ā-movement to its edge (see Roberts 2012: 386–399 and citations there for details).

[7] Mitrović (2014) provides semantic arguments for information-related properties of 2P in IE, lending support to the Ā-incorporation analysis.

3.2 XP-first

The phrase-first 2P effect is elegantly parallel to the head-first 2P effect. One difference is that in XP-first constructions, the phasal [EF] is checked by phrasal movement.

The Germanic v2-type falls into this category and differs minimally from the Slavonic type in that, as Roberts (2012: 401) writes, while Slavonic 2P "require[s] fronting of just one element – *either* a head *or* an XP – the latter require fronting of *both* a *head* and an XP."

3.3 XP/X-first

What follows is the core of this section: there are configurations which seemingly alternate between X-first and XP-first. The constructions in question concern Ser-Bo-Croatian subject conjunctions (SCS).

The empirical focus of this section lies on the following pair of data:

(16) Ser-Bo-Croatian
[Ja i Mujo] smo otišli na pivo.
I and M will.PL go.PTCP on beer
'Mujo and I are going for a beer.'

(17) Ser-Bo-Croatian
[Ja smo i Mujo] otišli na pivo.
I will.PL and M go.PTCP on beer
'Mujo and I are going for a beer.'

While (16) shows a plain vanilla subject conjunction structure, the availability of (17) does not readily follow, prima facie, from Roberts's (2010) tenets. With regards to the conjunction subject, the plural auxiliary verb *ćemo*, once raised from Aux^{min} to T^{min}, is in 2P with respect to the *maximal* category linearly to its left. What (17) shows, however, is that the Aux may be placed in a 2P with respect to the *minimal* category – I refer to this construction as SECOND-WORD (2W) effect. This very oscillation between word- and constituent-second configurations raises the core question on how a narrow-syntactic explanandum for seemingly string-related, and linearity-based, behaviour may obtain.[8]

On independent empirical grounds, then, we are led once more to reconsider the 2P effect with regards to the structural size of the first-position host.

[8]For independent arguments against the view that second-position effects derive from phonological processes, see Bošković (2001: 11–36, 75–93), Roberts (2010: Ch. 3), and further references therein.

While nominal clitics in Ser-Bo-Croatian are D^{min} elements that obligatorily incorporate into (some) C^{min} by virtue of defectivity, there is no defective relation constituted by an Agree chain between a clausal head and the verb, or Aux. Roberts (2012: 391) takes the auxiliary clitics to also bear D-features, just like nominal clitics, and assumes they are first-merged in T^{min}. Hence they are specified with [*i*D, *i*T]. Since Fin also bears [*i*T], auxiliaries are further assumed to incorporate to Fin^{min}, presumably after its [*u*ϕ/D] is valued. By contrast, full main verbs do not raise to Fin since they lack the relevant [*i*T] feature. If the Aux/T moved, accordingly, to Fin, wrong word order would ensue, assuming the subject conjunction is in Spec(TP). I exploit this seemingly wrong prediction to derive the 2w effect.

We take a slight excursus to discuss Ser-Bo-Croatian auxiliary clitics. While auxiliaries are in T^{min}, by being first-merged there Roberts (2012) or moving there from, say, Aux^{min}, there is one auxiliary clitic, *je* 'is.3SG', displaying different distribution. I take this auxiliary to be first-merged in C, specifically as the Fin category.[9]

(18)

φ	SG	PL
1	sam	smo
2	si	ste
3	*je*	su

To maintain the special syntactic status of *je* as a C-occupying clitic with its morphology, I take its form to be an allomorphic default. Hence, at C-level, its ϕ/D-features are not only irrelevant but non-existent:

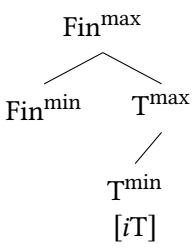
(19) a. /je/ ⇔ Aux

b. /sam/ ⇔ Aux / [1SG]

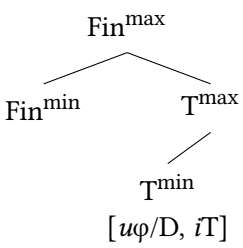
c. /smo/ ⇔ Aux / [1PL]

d. …

(20) [3SG] Aux:



(21) non-[3SG] Aux:

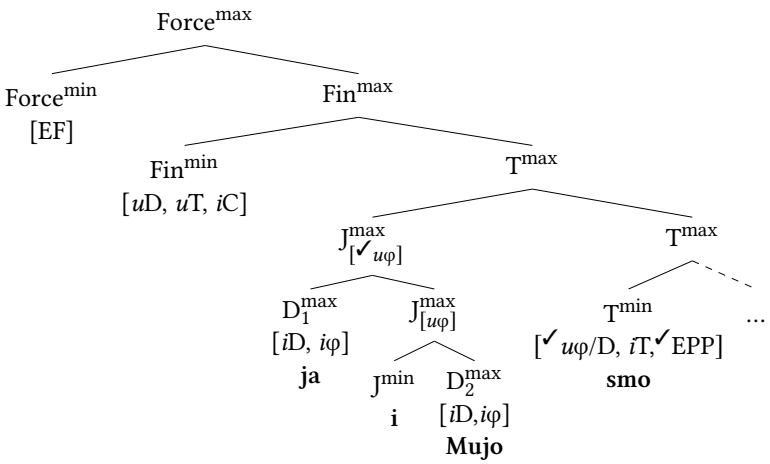


[9] Bošković (2004) in fact provides evidence that *je* is generated in the same position in the syntax as other auxiliaries.

This leads me to assume that Fin, where *je* is first-merged, does not carry a probing feature [*u*φ] but, as Roberts (2010; 2012) contends on independent grounds, the probe [*u*D].

A standard 2P clitic construction with a conjoined subject is then the one in which Aux is *in situ* in T^{min}.[10]

(22)



Note that the [1SG.NOM] pronoun *ja* is not a clitic but truly a D^{max}. This is confirmed by the fact that *ja* may coordinate and a pronominal clitic like *me* 'me.ACC' may not, since only maximal categories coordinate (Kayne 1994).

As for the position of the Aux/T^{min}, I take it to raise to Fin^{min}, as per Roberts (2012: 396) and references therein. Full main verbs or long/non-clitic auxiliaries, are taken to originate as V^{min} and raise to T^{min}, presumably via v^{min} and any other relevant aspect/mood head on the way to T^{min}. Once there, however, full verbs and full auxiliaries are not assumed to be able to raise to Fin^{min} as Fin^{min} lacks the V-feature specified on the complex T^{min}. As such, they are fronted by virtue of [EF] on $Force^{min}$. This, then, constitutes an instance of Ā-movement of a minimal category to the Spec(ForceP) position, as Roberts (2012: 396) contends.[11]

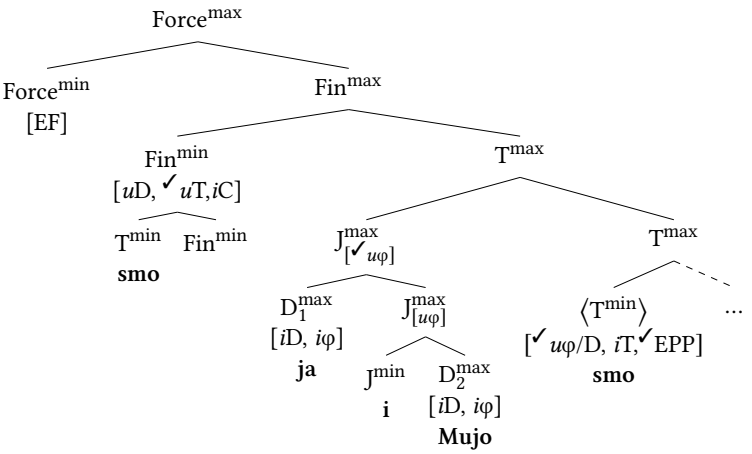
The set of probing features [*u*D, *u*T] on Fin^{min} in (22) are valued with the raising or incorporation of T^{min} which carries the corresponding values for [*u*D, *u*T]

[10] Since the system resting on defectivity we are adopting requires valued uninterpretable features to not undergo deletion upon valuation, I represent checked [*u*F]s with a superscripted ✓ next to the [*u*F]. Equally parsimoniously, if [*u*F] do not delete once checked, neither should discourse-related [EF] or [EPP] delete by the same token.

[11] Another view would be to maintain head-to-head movement and assume that Force's EF may be checked by incorporation of T^{min}, as Roberts (2012) proposes for European Portuguese. If this is desirable, then incorporation is extendable to Ā-processes, as well as prima facie potentially non-defective goals.

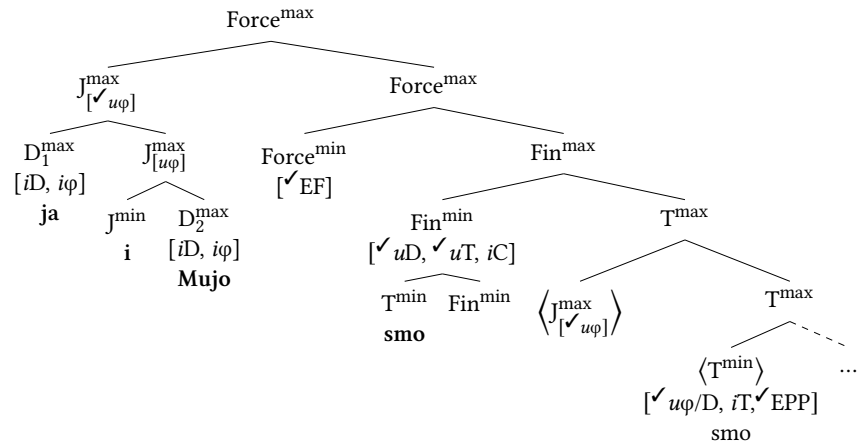
and which constitutes a defective goal with regard to Fin^{min} which, aside from the two uninterpretable features, bears some intrinsic C-feature.

(23)



Upon raising to Fin^{min}, the subject, independently of its internal (non/conjunctional) structure, moves to Spec(ForceP) to check the relevant [EF]. The subject may well move to, say, Spec(TopP) and check the clausal [EF] there; nothing hinges on the precise location of the subject.

(24)



The derivational step involved movement of the maximal category for purposes of [EF]-valuation. How do I then derive the 2w configuration using the exact set of narrow-syntactic devices?

The most obvious option, given the analysis thus far, is to focus methodologically on the derivational steps motivated thus far and maintain as much as possible for the 2w configuration. In this view, I solely restrict or modify the application of a rule that operates anyway. Since a coordinate structure (CS) should not introduce any special restrictions on phrase structure, it is untenable on conceptual grounds to assume that a presence of a subject CS would tamper with the rules operating independently of it. What I would like to maintain, ceteris paribus, is the raising of the defective T^{min} as probed by Fin^{min}'s [*u*D, *u*T], and the raising of the subject to check locally the [EF].

Two narrow-syntactic options make themselves available and amenable to an analysis that bears out the desired word order. The first is methodologically parsimonious insofar as it maintains both of the movement steps. One entails *movement out* of a CS, violating Ross's (1967) coordinate structure constraint (CSC).[12] Another option violated anti-locality involving movement *movement into* the CS. In what follows, I consider each of the analyses in turn concluding with a note on theoretical risk management and appeal to some wider economy considerations. Let me repeat the relevant 2w configuration I focus on: in the two subexamples, I make reference to the base/trace option underlying the 2w configuration by assuming that either the D^{max} conjunct moves from the CS in (25a) or that the T-auxiliary moves into the CS and cliticises onto, or incorporates into, J^{min}.

(25) Ser-Bo-Croatian
[Ja smo i Mujo] otišli na pivo.
I will.PL and M. go.PTCP on beer
'Mujo and I are going for a beer.'

a. D-movement from the CS:
Ja_1 [t_1 smo i Mujo] otišli na pivo.
I will.PL and M. go.PTCP on beer

b. Aux/T-movement into the CS:
[Ja smo_1 i Mujo] t_1 otišli na pivo.
I will.PL and M. go.PTCP on beer

Let us start with the latter idea exemplified by (25b) involving the movement of Aux in T^{min} to J^{min}. While incorporation into the conjunction maker, for which I use the category J^{min}, is a well-attested phenomenon across old Indo-

[12] For other analyses of CSC violations in Ser-Bo-Croatian, see also Stjepanović (2014), Oda (2017), or Bošković (2017).

European languages,[13] movement of a head (T^{min}) into its own specifier, i.e., J^{max} in Spec(T^{max}), is both anti-local[14] and is ruled out by extension. The idea that a Probe and a Goal constitute two separate syntactic objects seems to be an axiomatic foundation of the Agree-based Minimalism I assume. Attraction, resulting from Agree, is, as Roberts (2012: 397) succinctly notes, an irreflexive relation. Even if such strong evidence is suppressed, it remains untenable to motivate movement of T^{min} into J^{min} which by feature-absorption acquires the label [D], since (con)junction inherently lacks categorial features. Therefore, if the categorial label of J^{max} in Spec(T^{max}) is [D], setting aside the anti-locality and extension issues, it is still untenable to motivate incorporation of T^{min} into what may essentially be D^{min}. Such a D/J^{min} object lacks neither the ϕ/D-features which T^{min} could (even more) locally check – hence any variant of A-movement is dispelled. It is also unnatural to ascribe the CS subject with any [EF] which could be checked by movement of T^{min}. Lastly, the formal feature specifications on T^{min} do not in any way constitute a proper subset of the features on D/J^{min}, hence the defectivity of T^{min} and its subsequent incorporation cannot be motivated.

By unsuccessfully exhausting the theoretical space that the first analysis of T-to-J movement would entail, we are led to abandon this view and turn to the second view.

The second analysis appeals to the Ā-movement of the maximal D category *ja* 'I' from within the coordinate J^{max} to the clausal subject position, maintaining both T-raising and subject movement. This approach in fact parallels, and falls within, the well-observed pattern of left branch condition (LBC) violations, a.k.a. left branch extraction (LBE), see Figure 18.2.

Ignore temporarily the fact that this analysis rests on a violation of CSC. Once ignored, the question concerns the computational preference, or indeed availability, of the conjunct D^{max} for extraction. In this regard, I appeal to the A-over-A condition as formulated in Rackowski & Richards (2005) and applied in Roberts (2010).

What derives the 2w configuration is Rackowski & Richards's (2005) definition of the CLOSEST AVAILABLE GOAL (26):

[13] Such constructions derive from the well-known Wackernagel's (1892) law and give rise to the 2P effect. For an extensive overview of this phenomenon, see Mitrović (2014) and references therein.

[14] For overwhelming evidence that movement of a head into its own specifier is anti-local, see Saito & Murasugi (1999); Abels (2003); Grohmann (2003); Doggett (2004); Bošković (2005); Boeckx (2007), among others. As a reviewer reminds me, the ban on movement that is too short was first stated in Bošković (1994).

(26) A goal α is the closest one to a given probe if there is no distinct goal β such that for some X (X a head or maximal projection), X c-commands α but does not c-command β. (Rackowski & Richards 2005: 579)

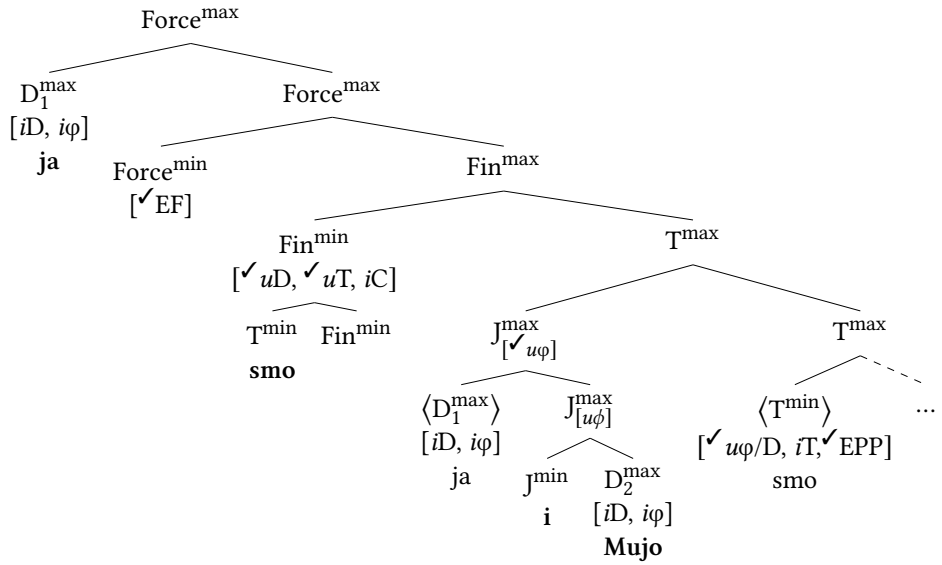


Figure 18.2: Deriving clitic placement using Ā-incorporation in the clausal edge

4 XLBE and non-constituent-first

Roberts (2010; 2012) has convincingly demonstrated not only that an exclusively syntactic approach to cliticisation phenomena is possible but that such an account is elegantly couched within some primitive theorems of syntax. If all cliticisation phenomena find a natural explanation, then it seems objectively odd, and subjectively disturbing, that one type of 2P effect should be afforded an extra-syntactic explanation. In fact, as it turns out, such an explanation is intractable. Hence, if narrow syntax cannot generate the XLBE string, which post-syntactic operations cannot derive (to which I turn), then the phenomenon of non-constituent-first (XLBE) constructions is even more intriguing.

What I aim to explain is the derivational nature of the strings such as the following, involving movement of a non-constituent.

(27) Ser-Bo-Croatian
U veliku on uđe sobu.
in big.LOC he.NOM entered.AOR room.LOC
'He went into a big room.' (Bošković 2005: 30n78)

As Bošković (2005: 30) notes, "under no approach to the internal structure of PP and the traditional NP do the preposition and the following adjective form a constituent to the exclusion of the noun modified by the adjective." This seeming fact potentially devastates an exclusively syntactic approach to XLBE. To maintain such an approach, for reasons of generality just given, one must logically invalidate Bošković's assertion. What I will develop is an approach that utilises the unrolling view of the spine that allows for a constituency structure of the preposition and the adjective. In concert with Roberts's (2010) approach to defectivity, a perfectly syntactic view of XLBE will be demonstrated. Before proceeding, I review the failed analyses. In doing so, I follow Bošković (2005: 30ff.) and cite two syntactic approaches first, and then a post-syntactic analysis.

The first possible analysis is syntactic. One way of deriving constituency of P and A is to posit remnant movement, as Franks & Progovac (1994) assume, namely movement of the NP to the edge of PP, followed by PP-fronting.

(28) $[_{\text{PP}}$ U veliku t_i $]_j$ on uđe t_j sobu_i. (Bošković 2005: 30, n. 79)

Bošković (2005) gives evidence against the remnant PP analysis. If the phrasal movement of the noun is what the remnant PP analysis rests on, it is predicted that the noun would be able to move on to the clausal edge, which is not the case.

(29) Ser-Bo-Croatian
* Sobu on uđe u veliku.
room he entered in big

The remnant PP analysis supposes PP extraction which precedes remnant fronting. Among other arguments, Bošković (2005) shows that, given the evidence from adjunct extraction (30), the analysis predicts movement of the noun *studenata* out of an adjunct, which should be barred on independent grounds.

(30) Ser-Bo-Croatian (Bošković 2005: 32)
Zbog čijih je došao studenata?
because-of whose is arrived students
'He arrived because of whose students?'

The second syntactic approach is that of Borsley & Jaworska (1988), who assume XLBE instantiates ordinary adjectival LBE. By invoking a restructuring operation, Borsley & Jaworska (1988) analyse XLBE as involving P-adjunction to the adjective. In a similar vein, both Corver (1992) and Franks & Progovac (1994) assume XLBE is derived from lowering, resulting in procliticisation of the preposition. Recall that the system we are assuming, most notably the LCA, prohibits rightward movement, qua lowering, and is both methodologically and conceptually reluctant to making reference to phonological operations if we are not forced to so independently. Note, however, that the preposition indeed shows phonological and prosodic evidence of proclisis (Talić 2013; 2015). Our account should, therefore, provide means for these post-syntactic facts to obtain without positing post-syntactic movement. I revisit this at the end of the section.

The third final possible alternative that Bošković (2005) entertains is to assume post-syntactic processes of *scattered deletion* or *copy and delete* (CD) that manipulate the linear configuration of the PP containing a modified noun and pronounce, in one segment, the P and the A strings in a moved constituent, while pronouncing the N in the base/trace position. This approach is sketched in (31).

(31) [U veliku ~~sobu~~] on uđe [~~u veliku~~ sobu] (Bošković 2005: 32n85)

A serious impediment to the CD account is the fact that it cannot predict the elements that may and may not undergo "deletion", since it is not the case that "anything" goes, as long as it is split. (See Bošković 2005 et seq. for more arguments against the CD account.)

(32) Ser-Bo-Croatian
* [Pravo u veliku ~~sobu~~] on uđe [~~pravo u veliku~~ sobu]
straight in big he entered room

Now let us turn to explicating the proposal. Given that the structural spine is taken to enter the derivation in the form of a head-complex, I take the following unfolding steps in the derivational course of a PP.[15]

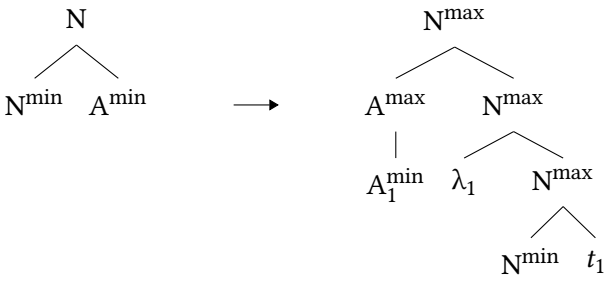
Bošković's (2005) phase-based account of LBE rests on Ser-Bo-Croatian being an NP-over-AP language (33a), unlike English which is AP-over-NP (33b).[16] I take the sole derivational difference between the NP-over-AP versus AP-over-NP structure to lie in the resulting label.[17]

[15]Since adjectives in Slavonic display morphological definiteness (via so-called short/long form), I take them to bear an IFF [*i*DEF].

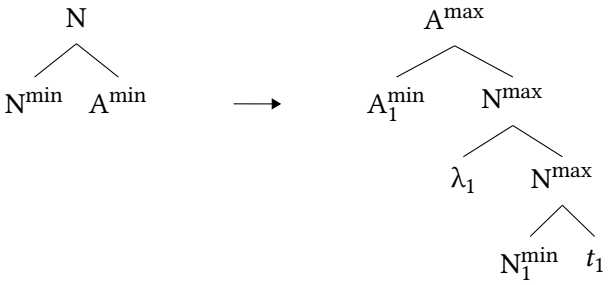
[16]The NP-over-AP vs. AP-over-NP difference/parameter is also entertained as an alternative to the phase account in Bošković (2005).

[17]For a conceptually parallel approach, see Donati & Cecchetto (2011).

(33) a. NP-over-AP:

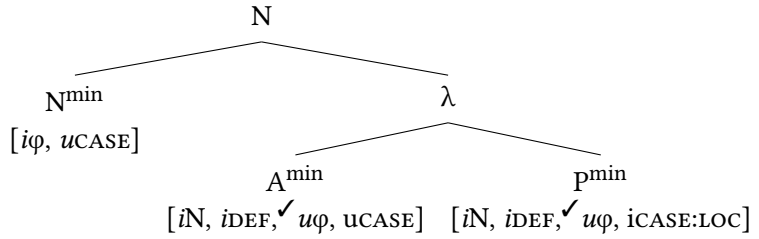


b. AP-over-NP:



In what follows, I provide a stepwise derivation of the PP and derive the availability of XLBE in line with the assumptions with which I started. At the onset, the c-commanding relations are in place for N^{min} to check the [*u*φ] probes on A^{min} and P^{min}.

(34)



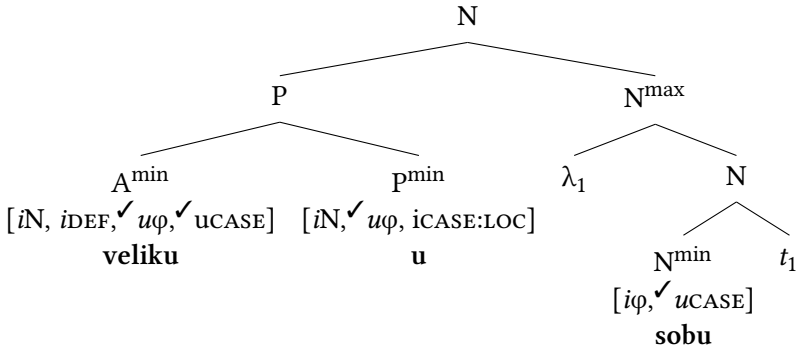
Note that the present proposal actually strengthens Bošković's (2005) proposal regarding the NP-over-AP structure, which amounts to stating that the A category is too weak to label in Ser-Bo-Croatian, a theoretical possibility argued for in Chomsky (2013).

Following the tenets laid out in §2, while N^{min} projects, its complement excorporates, as shown in (35). Since APs in Ser-Bo-Croatian do not project a label, P

projects upon excorporation (nothing hinges on this, as far as I can tell, but cf. the adjunction possibility discussed below).

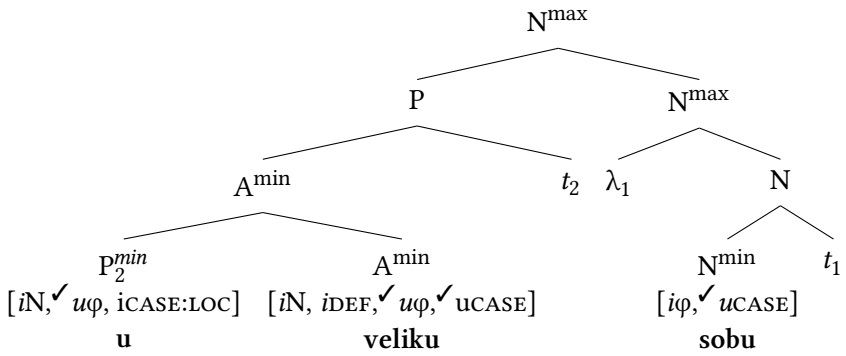
Upon raising, the case-features are checked as the c-commanding relation is established between the case-probe P and the case-seeking N^{min} and A^{min}.

(35)



By virtue of the DEF feature on the A^{min}, P^{min} under sisterhood constitutes a defective goal which gives rise to incorporation under defectivity.[18]

(36)



Upon final movement, the adjective is a maximal category via a mechanism of reprojection or Self Merge, see Figure 18.3 (I remain agnostic or rather apathetic with regards to this issue).

Note that even if I were to adopt a view according to which the A-adjunction is external to the unrolling of the nominal spine, I would arrive at a critically similar configuration. Since A^{max} adjoining the N-complex would not project, due to the nature of the NP-over-AP status of Ser-Bo-Croatian, P^{min}, contained in

[18]The fact that XLBE material is in focus testifies to the definiteness of the AP. Unlike ordinary LBE, XLBE obligatorily displays a definiteness effect.

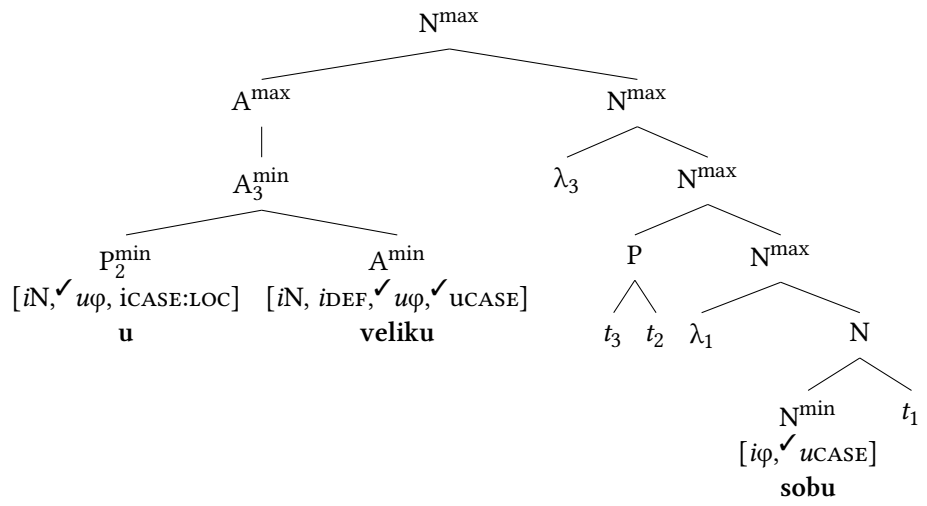
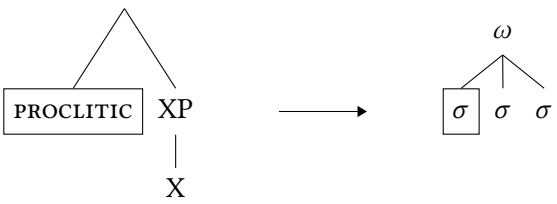


Figure 18.3: Successive excorporation as derivation of XLBE effects

COMPL(N^{min}), would excorporate to the root, ceteris paribus. A^{min} would have its [*u*φ] features checked via c-selection of N and its [*u*CASE] feature valued presumably via the chain ⟨N^{min}[uCASE:], N^{min}[iCASE:LOC]⟩. In case A^{min} is specified with a [DEF] feature, the features constitute a superset of those on P^{min} which would, in absence of [DEF] on A^{min}, otherwise excorporate to the root. This way, P is a defective goal that would undergo A-incorporation.

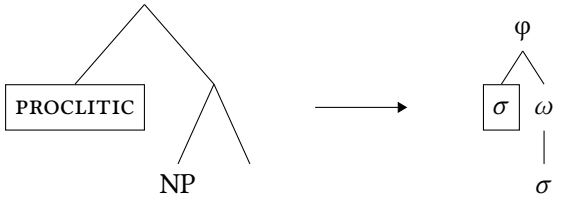
The preposition *u* has the prosodic properties of a proclitic, as mentioned earlier. Due to this, Talić (2013; 2015) provides a morphosyntactic account that is predicated on the assumption that proclitics, like prefixes, incorporate into the prosodic word ω of their host (37).

(37) (Talić 2015: ex. 7)



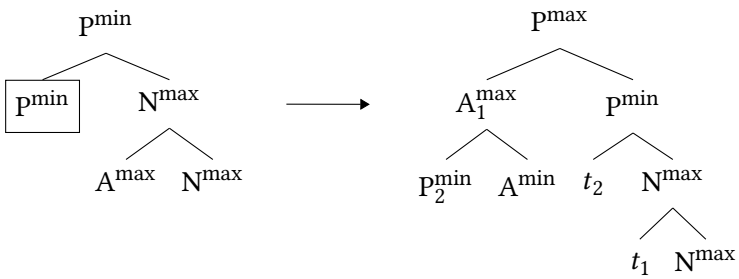
However, the clitic cannot interact with accent when syntactically attached to a branching host. In this case, the latter forms a prosodic phrase (φ) to which the proclitic may only attach.

(38) (Talić 2015: ex. 8)



Therefore, for the correct prosody to obtain, the syntactic configuration in (37) is required. Since under no approach can I derive such base-generated constituency (recall the drawbacks), Talić (2015) assumes that such orders are syntactically derived. In (39), I show her approach as demonstrated by her example (15) (ignoring the possibility of secondary AP and converting the phrase marker into BPS).

(39)



Such a syntactic approach assumes adjunct raising to Spec(root), viz. $\langle A_1^{max}, t_1 \rangle$, and subsequent incorporation of the preposition. This approach is architecturally rather similar to the approach I developed, with one crucial exception. The chain $\langle P_2^{min}, t_2 \rangle$ can be seen as breaching the anti-locality condition by moving the head into its own specifier.[19] The author, however, adopts the lines of reasoning from Matushansky (2006), i.a., which are, on independent grounds, divorced from the system of Roberts (2010; 2012) I am building on.

Also note that the relation between the prosodic constituency property and the availability of XLBE is not one of entailment. While the preposition *u* I have been citing in our data does have proclitic properties and is monosyllabic (its syllabic ω-weight: $\omega_\sigma(P^{min}) = 1$) there are other, prosodically non-simplex prepositions that feature in XLBE:

[19] See footnote 14.

(40) $\omega_\sigma(\mathrm{P}^{min}) = 2$ (Ser-Bo-Croatian)
Prema velikoj je zgradi otišao.
toward big.LOC is building.LOC went
'He went towards a big building.'

(41) $\omega_\sigma(\mathrm{P}^{min}) = 3$ (Ser-Bo-Croatian)
Povodom / uprkos teških je uslova ipak uspio.
in line despite difficult.GEN is circumstances.GEN still succeeded
'Due to difficult circumstances, he still succeeded.'

Thus, independently of the prosodic mappings, the anti-local configurations in (39) look as if, ceteris paribus, they should represent a standard derivation of Ser-Bo-Croatian PP grammar. Instead, I proposed a non-violating derivation that maintains the approach in full format, with little stipulation, and no reference to extra-syntactic modules.[20]

5 Phase-parameters of defective goalhood

Following Chomsky (2008) in assuming that only phase heads trigger movement, Roberts (2010) concludes that phase heads must, thereby, constitute the only cliticisation sites. For the clause, such phase heads are only C and v and may adduce from this idea of landing sites, or incorporation loci, a dichotomous typology of pronominal cliticisation: D-level arguments obligatorily cliticise onto C^0, while φ-level pronouns target v^0, as outlined in previous sections.

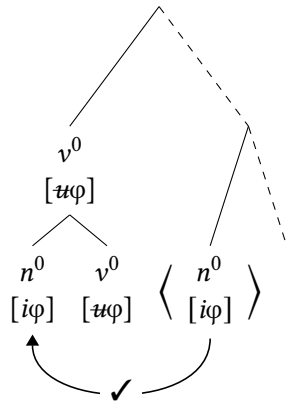
It is a fundamental requirement of the defectivity system that Roberts (2010) develops that lexical categorial features not constitute formal features on which the notion of defectivity is defined.

Assume a configuration in which v^0 combines with a φ-bearing nominal element, n^0. According to the theory, the minimal noun, bearing $[i\phi]$, incorporates[21] into v^0 after valuation of $[u\varphi]$ on the latter. This is demonstrated in (42). Assume, on the other hand, that lexical categorial features constitute legitimately formal features: since [N] ≠ [V], the condition on defectivity is not met in (43) and incorporation does not obtain. This is the problem I propose to resolve.

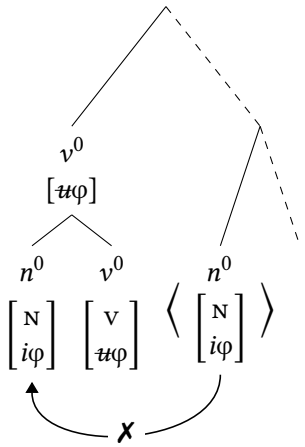
[20] The end result is similar to one Bošković (2005) achieves, being the only other account which achieves the required constituency here, but the road to it is very different.

[21] Or, rather, the feature valuation gives the effect of incorporation given that the chain reduction algorithm pronounces the copy at the head (effectively "in" v^0, by virtue of its feature makeup).

(42) $F_{v^0} \subset F_{n^0}$



(43) $F_{v^0} \not\subset F_{n^0}$



For the principle of defectivity to be operational in its full generality, it is necessary to develop the conditions under which both nominal and verbal categorial (formal) features are subsets of a larger feature-class which would legitimise (43).

In this regard, I adopt the tenets that the lexical categorial features are located in the categorisation formatives which combine with categoriless roots. These are the standard assumptions of Distributed Morphology.

Furthermore, it has been independently motivated that categorisers constitute the First phase. I propose to treat categorisers as phasers more explicitly. In this regard, I treat categorisers as "first-phasers", with the nominal or verbal lexical category as their attribute.

(44) a. $v^0 =_{\text{DEF}} [\pi : \text{V}]$
b. $n^0 =_{\text{DEF}} [\pi : \text{N}]$

What satisfies the defectivity condition in (43) is that both the probe and the goal bear the feature $[\pi]$, regardless of its (nominal or verbal) attribute.

This alone derives the non-arbitrariness of the defectivity system, as developed in Roberts (2010), which recognises and addresses only two types of defective goals insofar as pronominal cliticisation is concerned.

(45) a. C-orientation:
i. The relevant category of the defective goal α: D/N
ii. The category of the relevant probe β: C
iii. Agree between phase-phase objects yielding incorporation via chain $\langle \alpha_{[+\pi]}, \alpha_{[+\pi]} \rangle$

b. *v*-orientation:
   i. The relevant category of the defective goal α: φ
   ii. The category of the relevant probe β: *v*
   iii. Agree between phase-non-phase objects yielding incorporation via chain $\langle \alpha_{[+\pi]}, \alpha_{[-\pi]} \rangle$

My account leaves the analysis of Romance pronominal cliticisation, which Roberts (2010) treats as involving a defective φ goal and overall *v*-orientation, untouched. What we are allowing for is that the minimal D-less noun may count as a minimal phase and, thus, as a defective goal by virtue of categorisation constituting a first phase.

Let me wrap up this section on a diachronic note and the question of the historical sources of the D category in Slavonic as compared to, say, Romance.

(46) a. Romance pronominal clitics are φ-categories.
b. South Slavonic pronominal clitics are N-categories.

Some varieties of South Slavonic (including Macedonian, Bulgarian, and, to some extent, Slovenian) have developed an overtly full-fledged D-category which historically derives from demonstratives, in contrast to Romance, where it derives from pronouns. Given the approach I just outlined, the N/D parameter is therefore independent from the C-orientation parameter for cliticisation.

6 Discussion & conclusion

Let me take stock of the specific results this paper provides. The particular goal was to derive a NS constituency-compliant analysis of XLBE and X2P. To achieve this, I assumed an unrolling excorporation mechanism, according to which all functional layers of the clause (and, inversely and similarly, any other functional structure) originate as a complex head and proceed to unroll and excorporate as each argument is introduced in the structure. XLBE/X2P effects derive, as I have shown, from the featural subset relation, which either holds or does not hold at the point when the functional structure excorporates form the nominal category. In the last section, I showed how the defectivity-driven approach to cliticisation is consistent with the N/D parametric theory which assumes that some languages lack the functional D-layer. Assuming categorisation is an attributive property of the first phase, I have posited, on conceptually natural grounds, that phasality be recast as a feature with categorial attributes. With this twist, the subset relation

between N and C categories can be established, and the N-clitics consistently treated as C-orienting in South Slavonic.

The analysis I provided derives from basic properties of phrase-structure building, coupled with the notion of defective goals and a derivational onset as involving a head-complex (Shimada 2007). As it turns out, XLBE is perfectly amenable to an exclusively syntactic account of its configuration, thanks to Roberts's (2010) defectivity. A side product of such an approach was also a desirable account of 2P phenomena found in Bosnian CSs, which feature the seeming movement of the plural auxiliary into the first conjuncts.

Such an approach may be a stepping stone to understanding the interaction of pragmatics with speech act and vocative driven (X)LBE phenomena, as the following one, which I leave for future research.

(47) $\left[_{wishP}\right.$ Sretan$_i$ ti, Ian-e, t_i rođendan! $\left.\right]$
happy.M.SG you.DAT Ian-VOC birthday.M.SG

'Happy Birthday, Ian!'

Abbreviations

1	first person	IFF	intrinsic formal feature
2	second person	LBC	left branch condition
3	third person	LBE	left branch extraction
ACC	accusative	LCA	linear correspondence axiom
AUX	auxiliary	LOC	locative
BPS	bare phrase structure	M	masculine
COP	copula	NOM	nominative
CS	coordinate structure	PI	prosodic inversion
CSC	coordinate structure constraint	PL	plural
		PRS	present
DAT	dative	PRT	preterite
DEF	definite	PTCP	participle
EF	edge feature	SBJ	subject
EPP	extended projection principle	SG	singular
GEN	genitive	VOC	vocative

References

Abels, Klaus. 2003. *Successive cyclicity, anti-locality, and adposition stranding*. University of Connecticut, Storrs. (Doctoral dissertation).

Boeckx, Cedric. 2007. Some notes on bounding. *Language Research* 43(1). 35–52.
Borsley, Robert D. & E. Jaworska. 1988. A note on prepositions and case marking in Polish. *Linguistic Inquiry* 19(4). 685–691.
Bošković, Željko. 1994. D-structure, θ-criterion, and movement into θ-positions. *Linguistic Analysis* 24. 247–286.
Bošković, Željko. 2001. *On the nature of the syntax-phonology interface: Cliticization and related phenomena.* Amsterdam: Elsevier.
Bošković, Željko. 2004. Clitic placement in South Slavic. *Journal of Slavic Linguistics* 12. 37–90.
Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59(1). 1–45. DOI: 10.1111/j.1467-9582.2005.00118.x.
Bošković, Željko. 2008. What will you have, DP or NP? In Emily Elfner & Martin Walkow (eds.), *NELS 37: Proceedings of the 37th annual meeting of the North East Linguistic Society*, 101–114. Amherst, MA: GLSA Publications.
Bošković, Željko. 2009. On Leo Tolstoy, its structure, case, left-branch extraction, and prosodic inversion. In Steven Franks, Vrinda Chidambaram & Brian Joseph (eds.), *A linguist's linguist: Studies in South Slavic linguistics in honor of E. Wayles Browne*, 99–122. Bloomington, IN: Slavica.
Bošković, Željko. 2013. Phases beyond clauses. In L. Schürcks, A. Giannakidou, U. Etxeberria & P. Kosta (eds.), *Nominal constructions in slavic and beyond* (Studies in Generative Grammar 16), 75–128. Berlin: De Gruyter.
Bošković, Željko. 2017. On the coordinate structure constraint, islandhood, phases, and rescue by PF deletion. Ms., University of Connecticut. https://boskovic.linguistics.uconn.edu/wp-content/uploads/sites/2801/2019/05/CSCFranksShortNoAcknow.pdf.
Ćavar, Damir & Chris Wilder. 1994. Long head movement? verb-movement and cliticization in Croatian. *Lingua* 93(1). 1–58. DOI: 10.1016/0024-3841(94)90352-2.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press.
Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49. DOI: 10.1016/j.lingua.2012.12.003.
Collins, Chris & Edward Stabler. 2016. A formalization of minimalist syntax. *Syntax* 19(1). 43–78. DOI: 10.1111/synt.12117.

Corver, Norbert. 1992. Left branch extractions. In Kimberley Broderick (ed.), *NELS 22: Proceedings of the 22nd annual meeting of the North East Linguistic Society*, 67–84. Amherst, MA: GLSA Publications.

Doggett, Teal Bissell. 2004. *All things being unequal: Locality in movement*. MIT. (Doctoral dissertation).

Donati, Caterina & Carlo Cecchetto. 2011. Relabeling heads: A unified account for relativization structures. *Linguistic Inquiry* 42(4). 519–560. DOI: 10.1162/LING_a_00060.

Franks, Steven & Liljana Progovac. 1994. On the placement of Serbo-Croatian clitics. *Indiana Linguistic Studies* 7. 69–78.

Grohmann, Kleanthes K. 2003. Successive cyclicity under (anti-)local considerations. *Syntax* 6(3). 260–312. DOI: 10.1111/j.1467-9612.2003.00063.x.

Halpern, Aaron L. 1992. *Topics in the placement and morphology of clitics*. Stanford University. (Doctoral dissertation).

Halpern, Aaron L. 1995. *On the placement and morphology of clitics*. Stanford, CA: CSLI Publications.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 37(1). 69–109. DOI: 10.1162/002438906775321184.

Mitrović, Moreno. 2014. *Morphosyntactic atoms of propositional logic: A philological programme*. University of Cambridge. (Doctoral dissertation).

Mitrović, Moreno. 2017a. Phasal economy & incorporation. Ms. Bled Institute & University of Cyprus. Submitted to *Linguistic Inquiry*.

Mitrović, Moreno. 2017b. Phasal economy and incorporation. Ms., under revision. http://mitrovic.co/output.

Mitrović, Moreno. 2018. Configurational change in Indo-European coordinate constructions. In Ana Maria Martins & Adriana Cardoso (eds.), *Word order change*, chap. 2, 19–44. Oxford: Oxford University Press.

Mitrović, Moreno. 2021. *Superparticles*. Dordrecht: Springer.

Oda, Hiromune. 2017. Two types of the coordinate structure constraint and rescue by PF deletion. In Andrew Lamont & Katerina Tetzloff (eds.), *NELS 47: Proceedings of the 47th annual meeting of the North East Linguistic Society*, 343–356. Amherst, MA: GLSA Publications.

Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. 36(4). 565–599. DOI: 10.1162/002438905774464368.

Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press.

Roberts, Ian. 2012. Phases, head-movement, and second-position effects. In Ángel J. Gallego (ed.), *Phases: Developing the framework*, 385–440. De Gruyter.

Ross, John Robert. 1967. *Constraints on variables in syntax*. MIT. (Doctoral dissertation).

Saito, Mamoru. 2012. Case checking/valuation in Japanese: Move, Agree or Merge? *Nanzan Linguistics* 8. 109–127.

Saito, Mamoru & Keiko Murasugi. 1999. Subject predication within IP and DP: Essays in memory of Osvaldo Jaeggli. In Kyle Johnson & Ian Roberts (eds.), *Beyond principles and parameters*, 167–188. Dordrecht: Kluwer.

Shimada, Junri. 2007. Head movement, binding theory, and phrase structure. Ms., MIT.

Stjepanović, Sandra. 2014. Left Branch Extraction and the Coordinate Structure Constraint. In Jyoti Iyer & Leland Kusmer (eds.), *NELS 44: Proceedings of the 44th annual meeting of the North East Linguistic Society*, 157–170. Amherst, MA: GLSA Publications.

Talić, Aida. 2013. Extraordinary complement extraction: PP-complements and inherently case-marked complements. *Studies in Polish Linguistics* 8. 127–150.

Talić, Aida. 2015. Syntactic mobility of the host and accent shift to proclitics. In Thuy Bui & Deniz Özyıldız (eds.), *NELS 45: Proceedings of the 45th annual meeting of the North East Linguistic Society*. Amherst, MA: GLSA Publications.

Thurneysen, Rudolf. 2003. *A grammar of Old Irish*. Dublin: Dublin Institute for Advanced Studies.

Wackernagel, Jacob. 1892. Über ein Gesetz der indo-germanischen Wortstellung. *Indogermanische Forschungen* 1(1). 333–436. DOI: 10.1515/9783110242430.333.

Chapter 19

Person splits in Romance: Implications for parameter theory

M. Rita Manzini
University of Florence

Leonardo M. Savoia
University of Florence

This contribution addresses person splits in which 1/2P and 3P, or 1P and 2P systematically differ from one another with respect to the core grammar properties of case and agreement, giving raise to parametric variation. We consider two case studies from Romance varieties. The first one concerns 1/2P object clitics which, in Italian like in other Romance languages, have a simplified morphology with respect to 3P clitics, namely a single gender- and case-neutral object form, as opposed to the accusative vs. dative distinction, and the gender distinctions found in 3P. Moreover, 1/2P clitics only optionally trigger perfect participle (*v*) agreement, otherwise obligatory with 3P accusative clitics. We argue that these behaviors correspond to a core syntax phenomenon, whereby 1/2P clitics trigger DOM, which in the Romance languages takes the form of obliquization. The fact that 1/2P clitics are DOM obliques explains their specialized behavior in comparison with 3P clitics. The second case study has to do with partial pro-drop patterns in Northern Italian dialects involving the 1P vs. 2P split, interacting with the Externalization process and the Recoverability principle. We show that the (micro)parameters regulating the distribution of subject clitics are best seen as a reflex of macrocategories of grammar. Finally, we compare our approach with the literature on these phenomena (Cardinaletti & Repetti 2008; Calabrese 2008) and with the ReCoS parametric theory of Ian Roberts and his collaborators, discussing their different explanatory capabilities and results.

M. Rita Manzini & Leonardo M. Savoia. 2020. Person splits in Romance: Implications for parameter theory. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 403–434. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280663

1 Introduction

Our focus in this contribution is person splits, by which we mean interactions between pronouns and syntactic rules and relations such as Agree, Case, etc. in which 1/2P and 3P, or 1P and 2P, are seen to systematically differ from one another. We provide two case studies from Romance varieties.[1] In §2 we argue that partial pro-drop patterns in Northern Italian dialects involve the 1P vs. 2P split, interacting with the Externalization process and the Recoverability principle. Though the possible parametric values individuate a microvariation set (including only subject clitics), the parameters are best identified with the categorial splits themselves (such as 1/2P vs. 3P etc.), which involve macrocategories of grammar.

In this section, we concentrate on object clitics in Standard Italian, henceforth Italian. 1/2P object clitics have a simplified morphology (a single object form, gender neutral) with respect to 3P clitics (encompassing the accusative vs. dative distinction and gender distinctions). They also only optionally trigger perfect participle (*v*) agreement. We argue that these behaviours do not involve low-level morphological readjustments – but correspond to core syntactic phenomena. In this respect, we reject not just descriptive accounts, but also accounts that require an independent morphological component within formal models.

Several properties distinguish 1/2P clitics from 3P clitics in Romance, which for ease of exposition we will illustrate with just one language, namely Italian. Leaving aside the locative/instrumental *ci*, the genitive *ne* and the middle-reflexive *si*, the inventory of Italian clitics is as in Table 19.1. What is immediately evident from the table is that 3P clitics are differentiated by gender (masculine/feminine) and by case (accusative/dative) – but 1/2P are insensitive to either distinction.

The classical approach to asymmetries like those in (1) is to postulate a single underlying phi-features and case system, namely a system rich enough to be able to account for 3P – and to assume that morphological mechanisms (perhaps impoverishment and underspecification, in the way of Distributed Morphology) are responsible for the surface syncretisms observed in 1/2P. However, there is a third phenomenon with respect to which 1/2P and 3P differ, which does not directly involve the morphology of the clitics, but rather their syntactic behavior. As shown by Kayne (1989), in Italian (and French, etc.) perfect participles Agree with D(P) complements moved to their left, hence with accusative clitics. Dative

[1]Though our focus is on Northern Italian dialects (§2) and on Standard Italian (§1), the title refers to Romance varieties, in that the database of Manzini & Savoia (2005), which we use in particular in §2, includes Occitan, Franco-Provençal and Ladin (Rhaeto-Romance) dialects, spoken within the borders of Italy and Switzerland.

Table 19.1: Italian accusative and dative clitics

	ACC.M	ACC.F		DAT.M	DAT.F
1SG			*mi*		
2SG			*ti*		
3SG	*lo*	*la*		*gli*	*le*
1PL			*ci*		
2PL			*vi*		
3PL	*li*	*le*		*(loro)*	*(loro)*

clitics do not Agree, even if they are associated with gender features in normative Italian. We may assume that this is due to the fact that they are embedded under an oblique case. The relevant contrasts with 3P clitics are illustrated in (1).

(1) Italian

a. Lo / la / li / le ha aiutat-o / aiutat-a / aiutat-i / aiutat-e
him her them-M them-F he.has helped-M.SG talked-F.SG talked-M.PL talked-F.PL
'He helped him / her / them'

b. * la / li / le ha aiutat-o
her them-M them-F he.has helped-M.SG
'He helped her / them'

c. Gli / le ha parlat-o / *parlat-a
to.him to.her he.has talked-M.SG talked-F.SG
'He talked to him / her'

d. Ha loro parlat-o / *parlat-i / *parlat-e
he.has to.him/to.her talked-M.SG talked-M.PL talked-F.PL
'He talked to him / her'

Surprisingly, notionally accusative 1/2P clitics may not Agree in either gender or number, as in (2a), paralleling the dative clitic in (2c). Agreement of the 1/2P clitic with the perfect participle, as seen in (2b), remains possible, but it is optional. Free alternations of this type are standardly seen as pointing to the existence of two slightly different grammars. In the first one, 1/2P clitics Agree with the perfect participle; in the alternative grammar they do not. If two slightly

different languages are involved in the free alternation of agreeing and non-agreeing participles in (3), we expect there to be languages where only agreement is allowed and languages where only invariable participial forms are. Indeed there are many Italian varieties where 1/2P never trigger agreement (contrary to 3P forms), as documented by Manzini & Savoia (2005: §5.1.2).

(2) Italian

a. Mi / ti / ci / vi ha aiutato
me you us you.PL he.has helped-M.SG
'He helped me / you / us'

b. Mi / ti ha aiutata
me you he.has helped-F.SG
'He helped me / you'

c. Ci / vi ha aiutati / aiutate
us you.PL he.has helped-M.PL helped-F.PL
'He helped us / you'

d. Mi / ti / ci / vi ha parlato / *parlata / *parlati / *parlate
to.me to.you to.us to.you.PL he.has talked-M.SG talked-F.SG talked-M.PL talked-F.PL
'He talked to me / you / us'

It is true that, as we have noticed at the beginning, 1/2P pronouns lack nominal class features, but they have overt number properties. Therefore, relating optionality in agreement to the lack of (overt) morphological features is not immediately possible. What is more, under a morphological analysis, we would expect 1/2P to always display optional agreement, while agreement is clearly obligatory in subject contexts, as in (3). The same incidentally is true in Northern Italian dialects where 1/2P subjects are obligatorily realized as clitics. This forces the view that the optionality of 1/2P object agreement depends not on the lexical content of the 1/2P forms, but rather on their structure of embedding.

(3) a. (Io_F) sono arrivata / *arrivato
I.F am arrived-F.SG arrived-M.SG
'I have arrived'

b. (Noi) siamo arrivati / arrivate / *arrivato
we are arrived-M.PL arrived-F.PL arrived-M.SG
'We have arrived'

The alternative option, taken by Manzini & Savoia (2005) and Kayne (2010), is embedding the analysis of clitics firmly within core syntax, including their apparently idiosyncratic syncretisms. As Kayne (2010: 144) argues, "syncretism of the sort under consideration is nothing other than a particular kind of syntactic ambiguity". Specifically, addressing the 1st pronoun plural *ci* (syncretic with locative) he proposes that "it is not that *ci* has multiple possible values. Rather, *ci*, the same *ci*, is compatible in Italian with a certain range of syntactic contexts, … a silent PLACE, … a silent 1PL", where silent constituents are constituents grammatically represented but not pronounced. Manzini & Savoia (2005), Manzini (2012), and Manzini & Franco (2016) provide partial discussions of the range of empirical data that interests us here, which we will pursue in a more systematic manner in what follows.

1.1 Clitics and Case

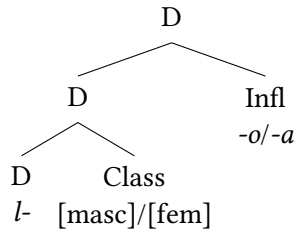
We pointed to three respects in which 1/2P objects differ from 3P objects. Two of them involve relational notions, namely case and agreement. Before we turn to them, let us consider the different phi-features make-up displayed by the two series of pronouns. The absence of nominal class endings (gender) on 1/2P clitics is a pan-Romance characteristic. In fact, according to Siewierska (2004: 194), "gender oppositions are characteristic of third rather than first or second person. Of the 133 languages in the sample (33%) which have gender in their independent person forms, 129 (97%) have gender in the third person as opposed to 24 (18%) in the second and three in the first (3%)".[2] Furthermore 1/2P forms are differentiated for number via their lexical basis. Thus even in Romance languages in which number is factored away from nominal class and lexicalized by a specialized *-s* ending, it is impossible to have 1st plural formed by adding *-s* to 1st singular. This is not necessarily a consequence of the absence of gender inflections. For instance, Sardinian varieties which present a dative singular form not inflected for gender, of the type *li* 'to him/her', also regularly pluralize it as *li-s* 'to them' (Manzini & Savoia 2005).

By contrast, the generalization holds that in Romance languages 3P clitics have an internal structure comparable to that of lexical nouns. Simplifying somewhat, the consensus in the literature is that at least two functional projections are needed for Ns – corresponding roughly to gender and number. In homage to the cross-linguistic comparison with Bantu languages, the lower category is often labelled Class, the higher category is Num (Picallo 2008), i.e. [[$\sqrt{}$ Class] Num].

[2] We thank Ludovico Franco for research and discussion on this point.

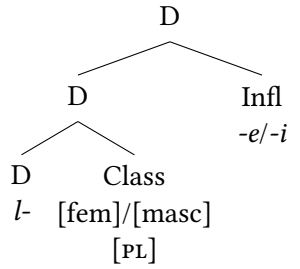
Extra complexity arises in Indo-European languages from the fact that there is no one-to-one mapping between the content of Class, which enters agreement with determiners and modifiers of N, and the inflections immediately following the root. We tentatively assign the inflectional vowel of Italian to an Infl position – which embeds both the root and the Class node. Transposed to the analysis of singular 3P clitics, this yields structures like (4).

(4)



Languages like Spanish have an independent lexicalization for the plural, namely *-s*; in Italian however pluralization is obtained by a change of the inflectional vowel. We may suppose that the plural 3P clitics, namely *li/le*, have the structure in (5), where the plural property is associated with the Class node. Note that this is in keeping with current ideas about Num not being a quantifier – but rather a divisibility predicate (Borer 2005).

(5)



The morphological structures in (4–5) map to a compositional semantics, essentially as outlined by Kratzer (2009: 221):

> the alleged "3rd person" features are in fact gender features, a variety of descriptive feature ... If [a descriptive feature] is to grow into a pronoun, it has to combine with a feature [def] that turns it into a definite description. If [def] is the familiar feature that can also be pronounced as a definite determiner in certain configurations, it should head its own functional projection, hence be a D. It would then not originate in the same feature set as descriptive features, which are nominal, hence Ns.

In this perspective, the pan-Romance (near-universal) fact that 1/2P forms are not associated with gender morphology, far from being a morphological syncretism or other quirk of pronunciation, corresponds to a potentially interesting (morpho)syntactic generalization – namely that 1/2P are pure deictic forms, deprived of predicative restrictions, even as elementary as Class (gender, countability).

A notable characteristic of Italian 1/2P clitics, apart from the lack of nominal class inflections, is the absence of case differentiations or, if one wishes, the accusative/dative syncretism – which is also replicated by many languages (e.g. French, Spanish, Albanian), though not by all (e.g. Romanian, Greek). In fact, in Italian (2), the *m-i, t-i* 1/2P person forms have the same *-i* inflection as the 3P dative *gl-i*. This inflection contrasts with that of the accusative in (1), corresponding to gender morphology (*-o, -a, -i, -e*).[3] Now, obliquization and specifically dativization of highly ranked referents normally characterized differential object marking (DOM) in Indo-European languages (Manzini & Franco 2016). Specifically in Romance, DOM marking of lexical DPs generally takes the form of the preposition *a* 'to' (in Ibero-Romance, in Southern Italian dialects).

At the basis of DOM is the fact that in many languages, case assignment depends on the referential content of the argument DPs. This is often described in terms of an animacy hierarchy. The classical discussion by Dixon (1979: 85–86) is based on the "potentiality of agency" scale, i.e 1st person < 2nd person < 3rd person < proper name < human < animate < inanimate. According to Dixon,

> it is plainly most natural and economical to "mark" a participant when it is in an unaccustomed role… A number of languages have split case-marking systems exactly on this principle: an ergative case is used with NPs from the right-hand end, up to some point in the middle of the hierarchy, and an accusative case from that point on, over to the extreme left of the hierarchy… Though the phenomenon is often referred to under the heading of split ergativity, it is evident that in the typological continuum it touches what we may call split accusativity.

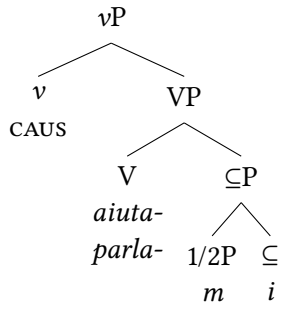
Similarly, using a different terminology, Aissen (2003: 473) states that "the factors that favor differential subject marking will be the mirror image of those that favor DOM".

[3] *-i* is the Latin inflection of the dative singular (in all declension classes excepting the II), also syncretic with the genitive (in the I class). Note further that though in Table 19.1, we have illustrated normative Italian, in colloquial Italian there is a single dative form for masculine and feminine, singular and plural, corresponding to *gli* (*l-* definiteness base + *-i* inflection).

The overt dative morphology of DOM objects suggests that these forms are not directly embedded as the internal argument of the event. Rather, their embedding requires the presence of a case layer, the dative, dedicated to the expression of possessors. We follow Belvin & den Dikken (1997: 170) in characterizing the possession relation in terms of zonal inclusion, i.e. "[e]ntities have various zones associated with them, such that an object or eventuality may be included in a zone associated with an entity without being physically contained in that entity". Following Manzini (2012), we label the dative case, carrying the relational inclusion content, as ⊆.

In these terms, the structure of embedding of *mi/ti* in (2) remains constant despite the fact that two different structures of embedding are implied by the predicates *aiutare* 'help' and *parlare* 'speak (to)' with 3P clitics in (1). In the structure in (6) we propose that the two arguments of ⊆ are the 1/2P clitic and – we assume – the event itself, adopting and adapting in this respect an idea of the applicative literature (Pylkkänen 2008).

(6)



Intuitively, transitive predicates can be paraphrased by an elementary predicate associated with an eventive name. Thus *aiutare* 'help' alternates with *dare aiuto a* 'give help to'. Hale & Keyser (1993), Chomsky (1995) formalize this intuition about the complex nature of transitive predicates by assuming that they result from the incorporation of an elementary state/event into a transitivizing (typically causative) predicate. Within such a conceptual framework it becomes clearer what we mean when we say that in (6), ⊆ takes as its arguments the 1/2P pronoun and an elementary state/event. In other words, (6) can be informally rendered as 'He caused me to have help/talk'. We claim that the 1/2P pronoun in (6) is introduced as a possessor, taking in its "zonal inclusion" domain an elementary event – for instance *aiuto* 'help'. By contrast, 3P complements of *aiutare* 'help' (or rather 'cause help') are embedded in a canonical transitive (causative) structure comprising a nominative agent and an accusative theme. The fact 3P arguments of *parlare* 'talk (to)' require the ⊆ embedding must be considered a

lexically governed alternation (subject to considerable cross-linguistic variation, see Svenonius 2002).

Manzini & Franco (2016) discuss potential problems for the present analysis in some detail. Specifically, the 1/2P argument of *aiutare* 'help' raises to the nominative position in the passive, while that of *parlare* 'talk (to)' does not, as in (7a) vs. (7b). The contrast in passivization is traditionally explained by the assumption that underlying cases are identical for 1/2P and 3P, though 1/2P are morphologically syncretic between dative and accusative. Thus the accusative object of *aiutare* 'help' can be passivized independently of whether it is 1/2P or 3P, while that of *parlare* 'talk (to)' cannot. Therefore the possible way to passivize *parlare* 'talk (to)' is an impersonal passive, as in (7b′).

(7) Italian

a. Sono stato aiutato
I.am been helped
'I was helped'

b. * Sono stato parlato
I.am been spoken
'I was spoken to'

b′. Mi è stato parlato (di te)
to.me it.is been talked of you
'It was talked to me (about you)'

Manzini & Franco (2016) propose a different explanation. They argue that the dative case with *parlare* 'talk (to)' is inherent, in the sense of Chomsky (1986), i.e. it is selected by the verb. Under passive, inherent dative case must be preserved, yielding an impersonal passive, as in (7b′) but barring raising to nominative position as in (7b). On the contrary, the dative case with *aiutare* 'help' and 1/2P objects is structural, since it depends not on the selection properties of the verb, but on the DOM configuration. Passive voids the context for the application of DOM, since the internal argument is raised out of its VP-internal position to [Spec, IP]. Therefore, no dative need be present in the derivation and sentences like (7a) are well-formed.

Before turning to agreement, it is worth mentioning that independent evidence for the presence of 1/2P vs. 3P splits in Romance DOM comes also from full pronouns – though it can only be briefly reviewed here. The standardly recognized manifestation of DOM in the Romance languages is the so-called prepositional accusative, whereby in a large number of Romance varieties (Ibero-Romance, Central and Southern Italian dialects, Romansh, Corsican, Sardinian,

Romanian) highly ranked objects are introduced by a preposition (with or without clitic doubling), most often *a*. The best known and most frequently attested pattern has DOM associated with definite/animate DPs, as in Standard Spanish (see Aissen 2003 for a typological survey, von Heusinger & Kaiser 2011 for a corpus study). However, as illustrated Manzini & Savoia (2005: §4.9), D'Alessandro (2015), other splits along the descriptive animacy/definiteness hierarchies are attested by Italian varieties. What is relevant for present purposes is that in some Center-South Italian varieties only 1/2P internal arguments require DOM, as in (8a). 3P pronouns and kinship terms (essentially functioning as proper names) undergo ordinary (bare) embedding, as in (8b).[4]

(8) Italian Colledimacine (Abruzzi, Manzini & Savoia 2005: 505)

a. a camatə a mme / a nnu
he.has called DOM me DOM us
'He called me / us'

b. a camatə frattə tiə / kwiʎʎə
he.has called brother yours him
'He called him / my brother'

Importantly, though the evidence from Italian 1/2P clitics reviewed would traditionally be treated in terms of morphological syncretism, there is no question that facts like (8) are syntactic.

1.2 Clitics and Agree

Let us then turn to agreement. Consider first 3P clitics. Under Chomsky's (2000; 2001) model of Agree, we may say that transitive verbs (i.e. verbs with an external argument and a *v* structural layer) include a probe on *v*, which attracts the closest

[4]Other varieties displaying the same pattern are Cagnano Amiterno (Abruzzi) and Borbona (Lazio); optionality of DOM in the 3P characterizes a few more dialects in the corpus, specifically Avigliano Umbro (Umbria), Torricella Peligna (Abruzzi), Canosa Sannita (Abruzzi). In fact, in contexts involving 1/2P pronouns, or in any event pronouns, DOM and clitic doubling can also surface in Northern Italian. In (i) we reproduce an example from Trieste (an anonymous reviewer suggests data from the dialectologically close variety of Padua).

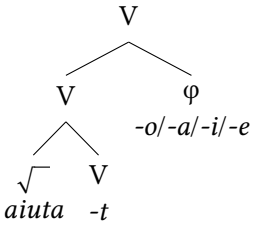
(i) Trieste, Venezia Giulia (Ursini 1988: 548)
el te ga bastonado a ti
he you has beaten DOM you
'He beat you up'

argument (by Minimal Search), namely the object of V. Agree (i.e. Match/Identity) then goes through, yielding (9a); for the sake of exposition we have assumed that the clitic has a base position inside the VP. Otherwise, the perfect participle turns up inflected with the invariable masculine singular ending, as in (9b). The traditional assumption in this respect is that some sort of morphological default repairs the lack of syntactic agreement.

(9) a. [$_{v\mathrm{P}}$ *aiutata* [$_{\mathrm{D}}$ *la*]]
b. [$_{v\mathrm{P}}$ *parlato* [$_{\subseteq}$ [$_{\mathrm{D}}$ *gli/le*]]]

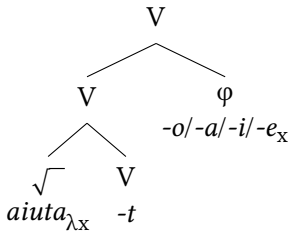
For ease of exposition, we have assumed that the perfect participle is an unanalyzed unit, associated with a probe in the form of a feature matrix, essentially as in Chomsky (1995). In reality, the perfect participle consists of a lexical base (inclusive of a so-called thematic, or inflectional class, vowel, which will be disregarded here), followed by a perfect ending *-t*, followed in turn by a suffix containing gender and number information (*-o, -a, -i, -e*), as in (10). The φ constituent is presumably to be identified with the agreement probe.

(10)



Classical theories of null subjects hold the view that the finite inflection of languages like Italian is pronominal-like (Rizzi 1982), hence it represents a lexicalization of the subject. In fact, in some models the *pro* empty category is dispensed with altogether (Borer 1986 for an early statement, Manzini & Savoia 2005; 2007). Suppose we generalize this idea to all agreement inflections. The perfect participle inflection, seen in Italian (9), will be construed as an elementary lexicalization of the internal argument within the morphological structure of the verb, as schematized in (10). Classical theories of pro-drop hold the view that the finite inflection of languages like Italian is pronominal-like (Rizzi 1982); in fact, some models treat it as satisfying the EPP, so that the pro empty category becomes redundant (Borer 1986; Manzini & Savoia 2005; 2007). Suppose we generalize this idea to all agreement inflections. The perfect participle inflection will then be construed as an elementary lexicalization of the internal argument, as schematized in (11).

(11)



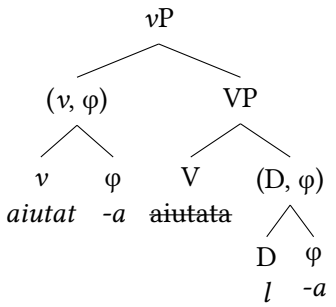
In (11), the φ constituent endowed with gender and number (i.e. nominal class) specifications needs a 1/2P or D closure in order to achieve referential status. This can only be obtained via the application of Agree. According to Chomsky (2000: 122) "the simplest assumptions for the probe–goal system" are formulated as in (12). Matching, namely feature identity according to (12a), "is a relation that holds of a probe P and a goal G. Not every matching pair induces Agree. To do so, G must (at least) be in the domain D(P) of P", defined as in (12b). Furthermore, "a matching feature G is closest to P if there is no G' in D(P) matching P such that G is in D(G')" as in (12c).

(12) Chomsky (2000: 122)

a. Matching is feature identity.
b. D(P) is the sister of P.
c. Locality reduces to closest c-command.

Our proposal (see also Manzini & Savoia 2005; 2007; 2011) holds on to these "simplest assumptions", but revises their standard implementation, in keeping with the need to interweave morphological and syntactic analysis. Specifically, we may expand the schematic structure in (9a) as in (13). We translate the classical idea that φ features percolate to the head level *v* by assuming that labelling creates a (*v*, φ) projection. At this point Agree proceeds along the lines in (12) creating a pair ordered by c-command and obeying locality, normally taken to be (*aiutata, la*). We may equally, and more perspicuously, pare the Agree sequence down to (*-a, -a*).

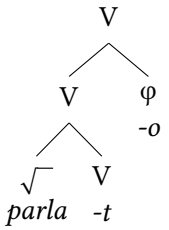
(13)



We know that in Chomsky's (2000; 2001) conception, Agree is a matter of deleting the uninterpretable features of the probe, with the result that a single copy of an agreement pair survives, namely the interpretable copy of the goal. But this is simply a technical implementation. One may keep closer to the morphological reality of agreement and assume that agreement is a matter of feature unification. Thus the agreement pair for (13) unifies the feminine features instantiated by the *-a* inflections of *v* and D. As a result, the D features morphologically instantiated by *l-* provide the necessary and sufficient referential closure for the internal argument of *aiutare* 'help'. In this perspective, the satisfaction of Full Interpretation at the conceptual-intentional (CI) interface depends on the fact that the operation of Agree creates an equivalence set, interpreted as a single argument with multiple occurrences (what Manzini & Savoia 2007 call agreement chains).

Let us then consider the 3P non-agreeing pattern in (9b). The internal structure of the perfect participle is as already indicated in (9), except that *parlare* 'speak (to)' does not introduce an internal argument. Rather, it selects the dative preposition or case, i.e. an element with (⊆) relational content, introducing a possessor. As a consequence, the φ node is externalized by the invariable *-o* ending, as in (14); the latter could be the realization of an empty φ node, i.e. what is traditionally called a default.[5]

(14)

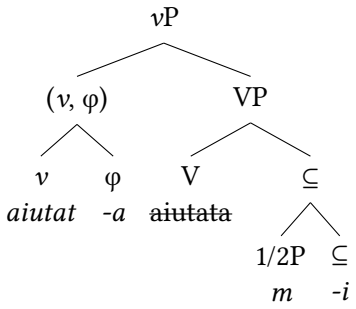


At this point, we are in a position to consider the crucial 1/2P data. Specifically, with *aiutare* 'help' two alternatives are possible. In present terms, the first alternative consists in the partial saturation of the internal argument of the participle by a gender and number inflection, as in (15). The φ probe can be matched with the 1/2P content as a goal, creating an agreement pair. The operation requires that the 1/2P constituent is visible despite the presence of ⊆ oblique morphology; in other words the ⊆ case morphology must be transparent. We already suggested in the discussion surrounding (13) that the right way to think about

[5]In a less stipulative way, in the absence of an internal argument, we could take the φ node to realize the abstract event argument. Note that in Romance languages where productive neuter gender is available (Central Italian dialects, Manzini & Savoia 2005; 2017), the latter is associated with mass and eventive contents and also with invariable perfect participles.

agreement pairs is not in terms of feature deletion (à la Chomsky), but rather of feature unification. Hence the descriptive gender and number properties of the *-a* inflection are unified with the 1/2P deictic properties of the clitic *m-/t-* under non-distinctness. More conventionally, we may add to the structure of the *m-/t-* clitic an abstract φ node, and assume that the content of this abstract φ node gets identified with that of the participle; the deictic content of 1/2P provides the required referential closure.

(15)



Next, consider the non-agreeing 1/2P structure in (16). With *parlare* 'talk (to)', as already reviewed in relation to the 3P clitic in (13), ⊆ is selected by the verb, and an agreement probe cannot be generated; rather the φ slot of the participle is empty, i.e. a default (but see footnote 5). With *aiutare* 'help' the agreement probe may be generated and satisfied along the lines of (15). We now propose that the agreement probe may equally not be generated, since the structure includes an oblique ⊆ object, albeit a structural (non-selected) one as in (17).

(16) $[_{vP}$ *aiutato/parlato* $[_{\subseteq}$ $[_{1/2P}$ *m-/t-*] $[_{\subseteq}$ *i*]]]

Let us summarize so far. We propose that a verb like *parlare* 'talk (to)' selecting an inherent ⊆ oblique, never generates a φ probe on the participle. A verb like *aiutare* 'help' generates a φ probe, when it is construed with an internal argument. However if DOM changes the internal argument to an ⊆ oblique, two possibilities are available. The first one is that the φ probe is generated on the participle and matched to the DOM object – in other words the latter is treated like a direct object and unlike an inherent oblique. Alternatively, the structural oblique is treated like an inherent (i.e. selected) oblique, resulting in empty/default agreement.

An analysis along these lines is supported by the observation that agreement is optional also with 3P clitics, if they are associated with structural oblique case,

i.e. oblique case which is not inherently assigned by the verb. Thus the *ne* genitive clitic in (17) licences agreement in the plural (masculine or feminine); however the invariant (masculine singular) form of the perfect participle is equally allowed in the relevant idiolects. We assume that genitive represents an instantiation of the same predicative content (⊆) as dative – except that dative predicates possession/inclusion between two arguments of a VP, while genitive predicates possession/inclusion between a D(P) and a modifier it embeds. In (17) the genitive *ne* clitic refers to a larger set including the two (*due*) objects I bought. On this basis, an agreement alternation is as expected depending on whether the (⊆) argument is treated along the lines of (14) or (15).

(17) Italian
Ne ho comprat-i / comprat-e / comprat-o due
of.them I.have bought-M.PL bought-F.PL bought-M.SG two
'I have bought two of them'

The facts that we have considered so far involve an extremely limited portion of the lexicon of just one language, essentially Italian clitics. Yet we have sought to explain them in terms of syntactic macrocategories, such as the Participant/non-Participant Person split and specifically its interaction with DOM phenomena. We must therefore briefly pause to consider whether these proposals are tenable with respect to available crosslinguistic evidence.

Importantly, the optionality of agreement with 1/2P clitics in Italian simply replicates at a smaller scale a well-known independent parameter affecting DOM obliques. The Indo-Aryan languages are a case in point. On the one hand, these languages present agreement of the perfect participle with the internal argument, for instance in Hindi (18a), where the internal argument is absolutive (and the external argument ergative). On the other hand the relevant languages are characterized by DOM, generally opposing animates to inanimates, realized by means of a postposition, which in Hindi is *-ko*, as in (18b). What is relevant here is that the DOM object does not Agree with the perfect participle, which shows up in the default masculine singular.

(18) Hindi (Mohanan 1995: 83; Ahmed 2006: 3)
a. Anil-ne kitaabẽ becĩ
Anil-ERG book.F.PL sell.PFV-F.PL
'Anil sold (the) books.'
b. Anjum-ne saddaf-ko dekhaa
Anjum.F.SG-ERG Saddaf.F.SG-DOM see.PFV.M.SG
'Anjum saw Saddaf.'

Though the Hindi pattern is robustly attested, in some Indo-Aryan languages DOM objects, also realized by an oblique postposition Agree with the perfect participle exactly as absolutive objects do. Thus in Marwari, a Rajasthani language the perfect participle "always agrees with O whether it is [DOM] marked or not" according to Verbeke (2013: 234). Crucially "agreement with an IO or an experiencer, marked with the same postposition is out of the question" (Verbeke 2013: 234). In (19) we illustrate just agreement of the perfect participle with DOM objects (*-nai*).

(19) Rajasthani (Khokhlova 2002)
RaawaN giitaa-nai maarii hai
Rawan.M Gita.F-DOM beat.PFV.F be.PRS.3SG
'Rawan has beaten Gita'

Recall that our thesis is that it is not possible to explain the case and agreement patterns of 1/2P clitics in Italian in terms of morphological idiosyncrasies. Rather, 1/2P clitics are targeted by DOM, hence they are externalized by oblique case. This in turn yields two possible grammars for agreement, one in which agreement probes characterize bare objects and DOM objects – and an alternative grammar in which agreement probes are restricted to bare objects. The data from Indo-Aryan languages are introduced here to confirm that these two options characterize DOM (of the Indo-European type) quite generally.

Thus, given any language in which we have evidence for both object agreement and DOM (on a person split basis, on an animacy basis), we expect optionality of DOM agreement (Italian) or obligatoriness of DOM agreement (Rajasthani/Marwari) or impossibility of DOM agreement (Hindi). These predictions are quite weak, but the data do not seem to warrant any stronger analysis; in other words we only predict that we will not find agreement with DOM objects to the exclusion of bare objects – which is correct.[6]

In conclusion, Italian (and Romance) object pronouns (clitic and full) provide evidence for the presence of 1/2P vs. 3P splits. Some of the facts we observed could in principle be handled in terms of morphological idiosyncrasies. Here we

[6]We do not have data on how DOM interacts with perfect participle agreement in varieties like Colledimacine in (8) or Trieste in fn 4. In any event, the analysis in the text excludes only the possibility that 1/2P agrees while 3P does not; this state of affairs is not attested in any Italian dialect, to the best of our knowledge. Note also that we do not make predictions on languages with no DOM. In principle we do not expect any asymmetries (for instance between 1/2P and 3P) in (object) agreement – but there may be reasons independent of DOM why such asymmetries are found.

argued instead that their lack of gender/number inflections may points to a genuine difference in constituent structure with 3P pronouns, which are effectively definite Ds. More to the point, the so-called accusative/dative syncretism in Italian 1/2P clitics and their optional activation of perfect participle agreement are connected with the DOM treatment of 1/2P clitics in the core syntax.

2 1P vs. 2P: Northern Italian subject clitics

In this section we address the issue of whether the Romance languages display evidence for a 1P vs. 2P split. To this end we consider subject clitics in Northern Italian varieties and specifically patterns of partial pro-drop. The microparametric variation involved (in the sense of Kayne 2000) will ultimately lead us to discuss recent proposals as to the nature of parameters and specifically their relation to macrocategorial splits such as 1P vs. 2P or, going back to §1, 1/2P vs. 3P.

2.1 Partial pro-drop in Northern Italian dialects

Manzini & Savoia (2005: §2.3) provide subject proclitic paradigms for 187 Northern Italian varieties (as counted by Calabrese 2008). Many of these dialects are characterized by partial pro-drop, namely the presence of no lexicalization for certain forms of the paradigm. The interest of the phenomenon is that only a minority of the logically possible patterns are actually attested. To begin with, 3P clitics (or a subset of them) are lexicalized in the quasi totality of Northern Italian dialects. Because of this, we illustrate first variation in the P(erson) paradigm, keeping the presence of D (i.e. 3P) forms constant.

The logical possibilities for combining four person denotations with two choices for lexicalization (P vs. zero) are sixteen. In the absence of further constraints, we expect to find all of them. However Manzini & Savoia (2005), Manzini (2015) tabulate only six possible proclitic patters, as shown in (20). This result remains constant if instead of considering null subjects slots, we consider slots taken by syncretic clitics lacking specialized P morphology.

(20)

		1st	2nd	3rd	4th	5th	6th
1.	*Prali*	—	P	D	P	P	D
2.	*Corte/Sief*	—	P	D	—	—	D
3.	*Càsola*	—	P	D	—	P	D
4.	*	—	P	D	P	—	D
5.	*	P	—	D	P	P	D
6.	*	P	—	D	—	—	D
7.	*	P	—	D	—	P	D
8.	*	P	—	D	P	—	D
9.	*French*	P	P	D	P	P	D
10.	*Sillano*	P	P	D	—	—	D
11.	*	P	P	D	—	P	D
12.	*	P	P	D	P	—	D
13.	*Livo*	—	—	D	—	—	D
14.	*	—	—	D	P	P	D
15.	*	—	—	D	—	P	D
16.	*	—	—	D	P	—	D

French in line 9 is the best-known Romance language that lexicalizes all P and D subject clitics. A language like Livo in line 13 further implies a 1/2P vs. 3P split. Apart from French and Livo, the other existing languages of (20) externalize subject clitics along a finer fault line, that between speaker and hearer. This may result in the externalization of just hearer reference, as in line 3 (Càsola); however, the lexicalization of just speaker is unattested. In order to account for the speaker/hearer asymmetry, Manzini & Savoia (2011), Manzini (2015) formulate the split between speaker and hearer (1P vs. 2P) as in (21), in terms of the salience of speaker reference.

(21) Speaker reference is (pragmatically) salient

(21), interacting with a universal rule/principle of grammar, namely Recoverability (22), explains why Càsola in line 3 of (20) is a possible language, while its mirror image in line 8 is impossible. Recoverability is standardly conceived as a principle constraining the deletion operation. Equivalently one may construe it as a constraint on the enrichment of L(ogical)F(orm), as in (22); in either case its content remains constant, i.e. that of licensing lack of Externalization. The salience of 1P in (21) makes it (pragmatically) recoverable, in the sense of (22), independently of any other syntactic or semantic condition being satisfied – licensing its lack of externalization. This is not the case for 2P, which must therefore be lexicalized. Therefore (21) crossed with Recoverability yields the prevalence of

2P lexicalizations over 1P ones in (20). To be more precise, rows 1-3 are allowed because 1P is not lexicalized and 2P is; rows 5 to 8 are excluded because 1P is lexicalized and not 2P; rows 4, 12 and 16 are excluded because this latter pattern holds in the plural.

(22) Recoverability
Recover non-externalized LF content (referential etc.)

Nevertheless, there are patterns in (20) which are excluded even though 2P is lexicalized, including rows 11 and 15. Descriptively, what seems to be relevant is that the speaker vs. hearer split is defined in the plural but not in the singular. We may therefore assume that (21) either applies to the singular, i.e. to speaker proper, or it cannot apply at all, as in (23). In other words, it is possible for it to be defined in the singular of a given language, and not in the plural– but not vice versa. A point to which we will return is that (23) is a statement about a value of a given categorial split (singular vs. plural) blocking another categorial split, namely the salience or prominence of speaker (vs. other referents).

(23) (21) is not defined in the plural.

Recall next that (20) records the attested variation in P lexicalization in languages where 3P (D) is invariably lexicalized. It is implicit in the way data are tabulated that the lexicalization of 3P is assumed to define an independent parameter. Thus in (20) there are varieties, for instance Livo, where the D series is lexicalized, but there is no exponent for P, defining a categorial split along the lines of (24), i.e. the 1/2P vs. 3P split also dealt with in §1.

(24) P (Participant) vs. D (Definiteness) referent

One may then expect the reverse situation to (20) to be attested, where 3P pronouns are not lexicalized, while on the contrary P pronouns are. Specifically, we may expect six languages to be generated, where 3P is zero and P slots vary along the lines discussed for (20) – i.e. lexicalization only of 2P is possible, and plural is not more differentiated than singular. If D is not lexicalized and P is not either we obviously have a classical pro-drop language like Italian (pattern 13). Pattern 2, with 2P as the sole lexicalized Participant form is also found. Pattern 9 is possible in turn – but it should be noted that in the dialect of Faeto (and the similar dialect of Celle, cf. footnote 5), the 3P form is undifferentiated/syncretic,

rather than zero.[7] These facts are depicted in (25), where pattern numbers refer back to corresponding patterns in (20). Evidently, our analysis overgenerates three patterns, namely 1, 3, 10. However, the sample of dialects missing 3P is very small (cf. footnote 7). This means that the conclusions we can infer from it are not necessarily significant when it comes to overgeneration. In any event, the analysis does not undergenerate.

(25)

		1st	2nd	3rd	4th	5th	6th
1.	*	—	P	—	P	P	—
2.	*Tetti*	—	P	—	—	—	—
3.	*	—	P	—	—	P	—
9.	*Faeto*	P	P	—	P	P	—
10.	*	P	P	—	—	—	—
13.	*(Italian)*	—	—	—	—	—	—

We should also consider the possibility that 3P singular splits from 3P plural. The lexicalization of the 3P plural to the exclusion of the 3P singular is not attested; this may be due to the fact that the plural cannot be more highly differentiated (via lexicalization) than the singular. In other words, the proposal we put forth in (23), saying that the 1P vs. 2P split may not be instantiated in the plural, should really be generalized to the possibility that any given split may be instantiated in the singular and the plural, along the lines in (26), but not vice versa. Thus, since 3P singular will have nominal class properties, along the lines of §1, we may conclude that it is possible to have them represented in the singular and not in the plural (pro-dropped) but not vice versa.

(26) Categorial split x is not defined in the plural.

By combining a lexicalized 3P singular, a zero 3P plural and the attested P configurations in (20), we may expect six patterns, as in (27). Only two of them are found, namely pattern 13, where only the 3P singular is lexicalized, and pattern 2 where 2P singular and 3P singular are lexicalized.[8] We observe that in all possible patterns the plural is consistently zero, suggesting that patterns 1, 3

[7] Besides Tetti (Dronero, in the Occitan Val Maira) other varieties that display the pattern in line 2 are Sarre (Franco-Provençal), and Bonifacio (at the southern tip of Corsica). Celle San Vito and Faeto, exemplifying the pattern in line 9, are Franco-Provençal varieties of Southern Italy (Franco-Provençal colonies).

[8] Besides Olivetta (West Ligurian, on the Occitan borders), other varieties that display the pattern in line 2 are Olivetta San Michele (Western Liguria, on the Occitan borders), Varese Ligure (Liguria), Calasetta (Ligurian dialect of Sardinia) and Como (Lombardy). Acceglio (in the Occitan Val Maira) is the only representative for the pattern in line 13 present in the corpus.

and 9 ought to be excluded because of the presence of plural P forms. Again the relevant idea seems to be that the plural cannot be more highly differentiated than the singular, excluding a person split in the plural (zero 3P vs. lexicalized 1/2P) where there is none in the singular. This would mean that our approach overgenerates only pattern 10 – though the disclaimer about the small number of dialects with the desired 3P configuration (cf. footnote 8) applies here as well. Importantly, the approach does not undergenerate.

(27)

		1st	2nd	3rd	4th	5th	6th
1.	*	—	P	D	P	P	—
2.	*Olivetta*	—	P	D	—	—	—
3.	*	—	P	D	—	P	—
9.	*	P	P	D	P	P	—
10.	*	P	P	D	—	—	—
13.	*Acceglio*	—	—	D	—	—	—

Moving away from the finer empirical details and on to the overall theoretical picture, we assume that a rule of Externalization, in the sense of Berwick & Chomsky (2011) pairs a CI content with a sensory-motor (SM) content, as in (28). Parameter values are the SM choices that (28) brings into effect, by interacting with C-I categorial splits such as Participant vs. Definite/Demonstrative, 1P vs. 2P, singular vs. plural. Similarly the 1P vs. 2P categorial split may interact with Recoverability, determining a fundamental asymmetry in Externalization. If so, the parameters are effectively the categorical splits themselves.

(28) Externalization
Pair a CI content x with a SM content y

Activating a yes value of a parameter implies activating the categorial split – otherwise the split remains inactive, corresponding to the zero value of the parameter. Generalizing from statements like (23), (26) one may further surmise a schema for the interaction between parameters, as in (29). In other words, when parameters cross, one of them may remain undefined for one value of the other. Thus the Speaker vs. other referents parameter (or categorial split) may remain undefined for value plural of the singular vs. plural parameter.

(29) Parameter (i.e. categorial split) A is not defined for value 0/1 of parameter (i.e. categorial split) B

In the next section we try to clarify our conception of the relation between categorial splits and parametrization, by comparing it to the notion of parameter

proposed within the *Rethinking comparative syntax* (ReCoS) project. Before doing so, we will briefly turn to alternative analyses of the Northern Italian partial pro-drop patterns, in terms either of cartographic hierarchies or of a Distributed Morphology-type component.

2.2 Competing views of parametrization

The data tabulated in (20) have attracted at least two types of analyses, besides the one defended here. Cardinaletti & Repetti (2008) argue that Person implicational hierarchies of the type proposed by typological work translate into structural hierarchies of Person positions. As the empirical basis of their work, they adopt Renzi & Vanelli's (1983) generalizations, which are based on a relatively restricted set of 30 dialects. These generalizations yield an implicational hierarchy 2^{nd} singular < 3^{rd} singular < 3^{rd} plural. Thus a language may lexicalize only 2^{nd} singular; it may lexicalize 2^{nd} singular and 3^{rd} singular, or it may lexicalize 2^{nd} singular, 3^{rd} singular and 3^{rd} plural – but other possibilities are excluded. Cardinaletti & Repetti map this implicational hierarchy to the structural configuration in (30). They propose that in (30) the 2SG position is licenced by verb movement to it. In turn, both the 3SG and the 2SG positions are licenced by verb movement to the 3SG, and so on. This means that no position can be licences unless 2SG is; 3SG can be licences only if 2SG is; and so on.

(30) [3PL [3SG [2SG

Cardinaletti & Repetti's (2008) proposal is typical of a range of cartographic responses to microparametric variation, under which a relatively simple computational component is maintained, while the underlying structures on which it operates are finely articulated. This response is empirically inadequate for the Northern Italian subject clitic data. The larger database of Manzini & Savoia (2005) brings out a few systematic counterexamples to Renzi & Vanelli (1983) and hence to Cardinaletti & Repetti; notably in varieties like Livo in (20), 3P subject clitics are realized, but not the 2P clitic.

A different approach is taken by Calabrese (2008), who concludes that the correct level of analysis at which to account for the intricate microvariation illustrated by Northern Italian subject clitics is not syntax but morphology. Recall that in introducing (20) we have noticed that the absence of subject clitics for a given set of forms is attested if and only if syncretic realizations are attested for the same set. It is therefore syncretisms, rather than partial pro-drop, that Calabrese sets out to account for. Calabrese's analysis is again based on a person hierarchy, namely 2SG < 3SG < 3PL < 1SG < 2PL < 1PL. For Calabrese, this

hierarchy corresponds to a set of constraints, each of which blocks the realization of the relevant forms, as in (31). For instance, the activation of constraint (31f) means that the feature cluster [+speak, +augm], i.e. 1st plural, is excluded. This in turn triggers morphological readjustment, in order to allow for lexicalization, yielding syncretism. Alternatively, the activation of a constraint can lead to obliteration, i.e. lack of the relevant lexicalization, hence to partial pro-drop.

(31) In the context [[AgrS ____] +V

a.	*[+part, −speak, −augm]	2SG
b.	*[−part, −augm]	3SG
c.	*[−part, +augm]	3PL
d.	*[+speak, −augm]	1SG
e.	*[+part, −speak, +augm]	2PL
f.	*[+speak, +augm]	1PL

Despite the wealth of detail present in Calabrese's analysis, the initial step of the hierarchy, i.e. 2P > 3P is violated by all languages where only 3P is lexicalized, like Livo in (20). Furthermore, Calabrese also notes that his system does not deal with the proclitics of a language where only the 1st singular is missing and all other forms are specialized – such as Prali in (20). From a theoretical point of view, the morphological repairs that Calabrese assumes to be at work require Late Insertion, in the sense of Distributed Morphology; these postulates violate minimalist principles such as Inclusiveness and no backtracking. It is possible that these minimalist principles hold in syntax and not in morphology for some reason, but the result is in any case an enrichment of the grammar.

It is also interesting to note that for Calabrese (2008) the conceptual basis for lexicalizing 2P but not 1P in Northern Italian subject proclitic paradigms is that marked forms such as 1P "shy" away from lexicalization. Technically, in his filter hierarchy in (31), the more marked a form is, the less likely it is that the constraint blocking it will be deactivated. Therefore, it is it the marked status of 1P that determines its lack of lexicalization. The present approach is the reverse – it is the inexpensive status of 1P in terms of Recoverability that determines its lack of lexicalization. Importantly, under this latter approach there is no special 2 < 1 markedness hierarchy for Italian dialect proclitics, but only the prominent status of speaker reference, corresponding to the classical 1 < 2 animacy ranking.

In conclusion, both the cartographic approach of Cardinaletti & Repetti (2008) and the morphological approach of Calabrese undergenerate in one crucial respect – i.e. they do not provide for the existence of languages with 3P (i.e. D)

clitics and no P clitic. Similarly, Calabrese's approach undergenerates with respect to pattern (21), line 1; the approach in Cardinaletti & Repetti does not really address 1P, so that the issue remains indeterminate. The crucial assumption in Manzini & Savoia (2005), Manzini (2015) that allows the correct results to be obtained in this respect is that the P vs. D split in (24) is independent of the 1P vs. 2P split in (21) – and in fact the singular vs. plural split is independent of both.[9] Vice versa the model overgenerates, at least as far as our empirical basis goes. The order of magnitude of overgeneration is 4 patterns over 64 (2^6), namely one in (27) and three in (25). The large majority of non-existing patterns is correctly excluded (49 of them) and more importantly all existing patterns are correctly generated (11 altogether) – i.e. the model does not undergenerate.

The absence of undergeneration (and the presence of some overgeneration) correlates with the fact that the present model is weaker than its competitors. Empirically, we have just argued that this represents an advantage – but the same conclusion holds from a theoretical point of view, since both cartographic hierarchies and a morphological filtering component are expensive devices and best avoided (see also Chomsky et al. 2019).

Let us then turn to the notion of parameter. According to Berwick & Chomsky (2011), parameters are not an external addition to the faculty of language, but are coevolved with it. In other words, parameters simply correspond to degrees of freedom open within Universal Grammar (UG), specifically in what concerns Externalization. As a consequence, the idea that parameter values are associated with lexical items (the so-called Borer–Chomsky conjecture, Baker 2008) takes on better defined contours – since the lexicon is the main locus of externalization, pairing CI and SM content.

Studies like the present one further argue that it is at best descriptively useful to refer to micro- and macro-variation – the former affecting very closely related languages and/or a small extension of the lexicon/grammar, while the latter covers comparison between different families and a considerable extension of their grammar. However, there is no sense in which one can define an opposition between macroparameters and microparameters. Manzini & Savoia (2011), discussing auxiliary selection (*be* vs. *have*) in Italian varieties, have this to say:

> The distinction between microparametric and macroparametric approaches to variation has been so often discussed that the contours of the debate

[9] There is further dimension of variation, discussed by all of the works quoted – namely the fact that enclitic paradigms differ from proclitic ones. Enclitic paradigms are largely irrelevant for the issue at hand, since it appears that essentially all of the logically possible patterns in (20) are instantiated (Manzini & Savoia 2005; Manzini 2015).

> have become somewhat blurred. It is evident that, to the extent that the primitives manipulated by variation are macrocategories like transitivity or voice, we could describe our approach as macroparametric – though the fact that the unit of variation can be as small as a single lexical item qualifies it as microparametric

Transposing this discussion to the case study in §2.1, Speaker, Plural, Participant, etc. are macrocategories capable of influencing the global forms of a grammar; at the same time, they can be seen to determine the microvariation in subject clitic systems in (20). Going back to §1, the same holds for DOM, which may determine macroalignment phenomena but also microphenomena restricted to the sole clitic domain.

In the recent ReCoS model (Roberts & Holmberg 2010; Biberauer & Roberts 2012; 2015; Sheehan 2014; Biberauer et al. 2014), microparameters and macroparameters simply represent different levels of application of a given parameter. The internal organization of parametric space is determined by general processing/economy principles, specifically feature economy (FE, Roberts & Roussou 2003) and input generalization (IG, Roberts 2007). These "general cognitive optimisation strategies" determine the general form of parameter hierarchies by interacting with the schema $Qhh \in P\ [F(h)]$ regarding "generalised quantification over formal features". In this schema *h* stands for head(s) belonging to set P, of which feature(s) F are predicated. Universal negative, universal and existential quantification over h are ranked in this order by feature economy and input generalization. The passage from larger to smaller sets of restrictor heads yields the descending hierarchy of macroparameters, mesoparameters, microparamenters (Biberauer et al. 2014 and references quoted there).

Biberauer et al. (2014) exemplify their model with several different hierarchies. Here, since we have discussed null subjects and subject clitics, we exemplify their null arguments hierarchy (cf. Roberts & Holmberg 2010: 49), which we reproduce in Figure 19.1.

The macroparametric region of the schema in Figure 19.1 corresponds to Figure 19.1a–c. In Figure 19.1a, lack of attestation for a particular type of features, here uninterpretable phi-features, counts as the least marked value in the parametric hierarchy, namely radical pro-drop languages (languages of the Chinese/Japanese type). In Figure 19.1b, the universal value of the parameter, corresponding to pronominal argument languages, in the sense of Jelinek (1984), already implies the restriction of the domain of application of the quantificational statement to certain categories, namely functional heads. Figure 19.1c, which posits

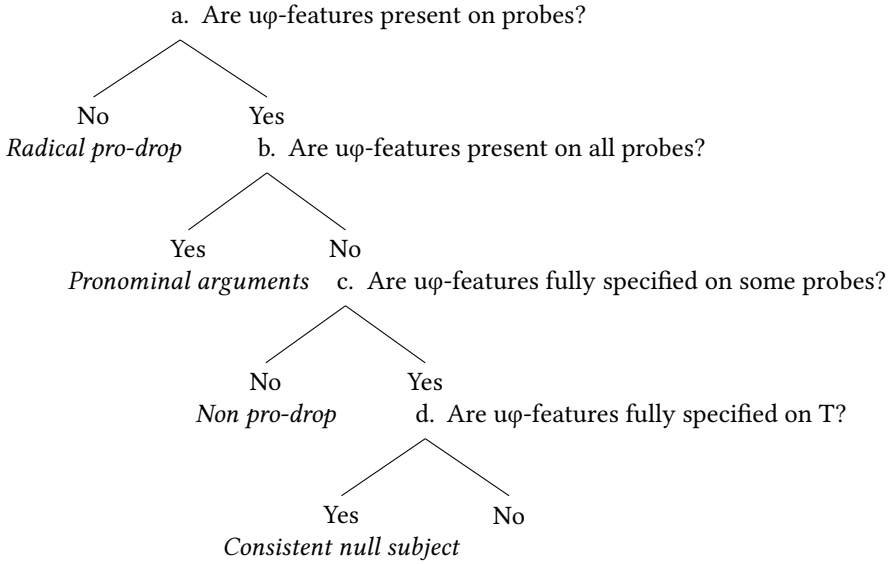


Figure 19.1: Null arguments hierarchy

the existence of uninterpretable phi-features sets on some functional heads, triggers the next set of statements (mesoparameters), concerning the association of uninterpretable phi-features with all T heads Figure 19.1d, and presumably further down with some T heads, and then on to microparameters etc.

Note that from mesoparameters down, what drives the construction of the hierarchy is a progressive domain restriction. We already mentioned that this is relevant for the head set *h* of which feature F is predicated; for instance, in the macroparametric steps (Figure 19.1a–c), the uninterpretable phi-features property is evaluated in relation to functional heads, while in the mesoparametric steps from Figure 19.1d down it is evaluated in relation to T heads. But if so, parameters are structured by something altogether more elementary than quantificational schemas and processing/economy principles, namely the existence of a Boolean superset/subset organization in the categorial domain. In the specific case at hand, this conclusion is strengthened by the observation that in the passage from Figure 19.1b to c, the query switches from "is present" to "is fully specified". This means that restrictions down the scale apply not only to the head set *h*, but also to the property F in the quantificational schema.

Informally, the basic aim behind the ReCoS approach is the integration of the microparametric scale with the macroparametric one. This seems eminently

compatible with the views expressed by Manzini & Savoia (2011) and here on microvariation and macrocategories (macroparameters). There are, however, differences between the position articulated by ReCoS and that expressed by Manzini & Savoia (2011) and endorsed here. The ReCoS model sees macroparameters and microparameters as applications of the same property in progressively smaller domains. Indeed much of the discussion of the ReCoS model is devoted to the progression down such hierarchies, like Figure 19.1. Manzini & Savoia (2011) take a weaker position, under which no such hierarchy holds, or at least not necessarily. In their terms, categorial splits between 1/2P (Participant) and 3P (Demonstrative/Definite), between Speaker and Hearer, and so on may become externalized in small areas of the lexicon (Northern Italian subject clitics) or may have systemic consequences (ergativity splits) – but this difference has no theoretical import.

In fact, Manzini & Savoia (2011) make a stronger point, namely that "macrophenomena can be decomposed into the same elementary conceptual components that determine local lexical variation – and in fact the latter is the true matrix of perceived macroparameters". In other words, let us keep to the idea that (micro)parameters are binary choices (categorial splits), applying to minimal units such as a single category or in the limit a single lexical item. Manzini & Savoia propose that macroparameters may have a purely logical existence, as extrapolations from microparameters (e.g. if category x has property P, x a functional category, then all functional categories have property P). This second point goes against the grain of the ReCoS models, as can be seen more clearly if we translate the two approaches in terms of acquisition or markedness

Suppose with Manzini & Savoia that the learner fixes lexical choices such as those concerning partial pro-drop in Northern Italian dialects locally. In their terms, this "local lexical variation" is "the true matrix of … macroparameters". This means that the differential treatment of 1/2P vs. 3P (or 1P vs. 2P etc.) in the lexicalization of subject clitics triggers the activation of the relevant categorial splits in the grammar of the language – leading the child to look out for these splits in other areas of the lexicon/grammar. In this sense, the microparametric (i.e. lexical) setting has a macroparametric (i.e. systemic) consequence in the acquisition process. Vice versa in the ReCoS model, if we understand it correctly, the learning path is strictly downwards, proceeding from macroparametric default to actual microparametric settings.

Similarly, for Biberauer et al. (2014) languages that are highest in the hierarchy in Figure 19.1, i.e. Chinese-style "radical pro-drop" languages or Jelinek's (1984) pronominal argument languages, are least marked. But it does not seem to be true that unmarked status corresponds to relative frequency of these languages or

other similar independent criteria for default status. In fact, the choice of treating all 1/2P clitics alike by lexicalizing all of them, or by not lexicalizing any of them (as opposed to 3P clitics) is certainly possible in Northern Italian dialects, but unpopular. More than half of the dialects in the corpus present a pattern whereby 1P singular and 1/2P plural are associated either with subject clitic drop (39/187) or with an uninflected subject clitic (65/187). In other words, on statistical grounds alone, one can legitimately conclude that the supposedly more marked mixed bag choice is in fact the default one.

3 Conclusions

In this contribution, we have argued for the existence of 1/2P vs. 3P splits, and 1P vs. 2P splits in important areas of the lexicon/syntax of Romance languages. On the one hand 1/2P vs. 3P splits (or 1P vs. 2P) interact with core grammar properties of case and agreement. On the other hand, in so far as certain split may or may not be activated, they yield parametric variation.

In the first part of the article, we noted that in many Romance languages, including Italian, 1/2P object clitics have a simplified morphology with respect to 3P clitics, namely a single gender- and case-neutral object form, as opposed to the accusative vs. dative distinction, and the gender distinctions found in 3P. 1/2P clitics also only optionally trigger perfect participle (v) agreement, which is obligatory with 3P accusative clitics. We have argued that these behaviours do not involve low-level morphological readjustments, but correspond to core syntax phenomena. Specifically, 1/2P clitics trigger DOM, which in the Romance (and Indo-European) languages takes the form of obliquization. Therefore, the special behaviours of 1/2P clitics with respect to 3P clitics (specifically the optionality of agreement) are to be imputed to the fact that the former are DOM obliques.

Our second case study is partial pro-drop patterns in Northern Italian dialects – which in our terms involves the 1P vs. 2P split, interacting with the Externalization process and the Recoverability principle. Though the possible parametric values individuate a microvariation set (including only subject clitics), the parameters are best identified with the categorial splits themselves (such as 1/2P vs. 3P etc.), which involve macrocategories of grammar.

Abbreviations

1	first person	F	feminine
2	second person	FE	feature economy
3	third person	IG	input generalisation
ACC	accusative	M	masculine
CAUS	causative	PFV	perfective
CI	conceptual-intentional	PL	plural
DAT	dative	PRS	present
DOM	differential object marking	SG	singular
EPP	extended projection principle	SM	sensory-motor
ERG	ergative	UG	Universal Grammar

References

Ahmed, Tafseer. 2006. Spatial, temporal and structural uses of Urdu *ko*. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG2006 conference*, 1–13. Stanford, CA: CSLI Publications.

Aissen, Judith. 2003. Differential object marking: iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483. DOI: 10.1023/A:1024109008573.

Baker, Mark C. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–373. Amsterdam: John Benjamins. DOI: 10.1075/la.132.16bak.

Belvin, Robert & Marcel den Dikken. 1997. *There*, happens, *to*, *be*, *have*. *Lingua* 101(3–4). 151–183. DOI: 10.1016/S0024-3841(96)00049-6.

Berwick, Robert C. & Noam Chomsky. 2011. The biolinguistic program: The current state of its evolution. In Anna Maria Di Sciullo & Cedric Boeckx (eds.), *The biolinguistic enterprise: New perspectives on the evolution and nature of the human language faculty*, 19–41. Oxford: Oxford University Press.

Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2014. Complexity in comparative syntax: The view from modern parametric theory. In Frederick J. Newmeyer & Laurel B. Preston (eds.), *Measuring grammatical complexity*, 103–127. Oxford: Oxford University Press.

Biberauer, Theresa & Ian Roberts. 2012. Towards a parameter hierarchy for auxiliaries: Diachronic considerations. *Cambridge Occasional Papers in Linguistics* 6. 267–294.

Biberauer, Theresa & Ian Roberts. 2015. Clausal hierarchies. In Ur Shlonsky (ed.), *Beyond functional sequence* (The cartography of syntactic structures 10), 295–313. Oxford: Oxford University Press.

Borer, Hagit. 1986. I-subjects. *Linguistic Inquiry* 17(3). 375–416.
Borer, Hagit. 2005. *The normal course of events* (Structuring sense 2). Oxford: Oxford University Press.
Calabrese, Andrea. 2008. On the shyness of the first person: Investigations on markedness and underspecification in morphology. Ms., University of Connecticut, Storrs.
Cardinaletti, Anna & Lori Repetti. 2008. The phonology and syntax of preverbal and postverbal subject clitics in Northern Italian dialects. *Linguistic Inquiry* 39(4). 523–563. DOI: 10.1162/ling.2008.39.4.523.
Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
Chomsky, Noam. 1995. *The Minimalist program*. Cambridge, MA: MIT Press.
Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions and challenges. *Catalan Journal of Linguistics*. DOI: 10.5565/rev/catjl.288.
D'Alessandro, Roberta. 2015. Merging probes: A typology of person splits and person-driven differential object marking. Ms., Leiden University.
Dixon, R.M.W. 1979. Ergativity. *Language* 55(1). 59–138. DOI: 10.2307/412519.
Hale, Kenneth L. & Samuel Jay Keyser (eds.). 1993. Cambridge, MA: MIT Press.
Jelinek, Eloise. 1984. Empty categories and non-configurational languages. *Natural Language & Linguistic Theory* 2(1). 39–76. DOI: 10.1007/BF00233713.
Kayne, Richard S. 1989. Facets of Romance past participle agreement. In Paola Benincà (ed.), *Dialect variation and the theory of grammar*, 85–103. Dordrecht: Foris.
Kayne, Richard S. 2000. *Parameters and universals*. Oxford: Oxford University Press.
Kayne, Richard S. 2010. *Comparisons and contrasts*. Oxford: Oxford University Press.
Khokhlova, Ludmila. 2002. Syntactic peculiarities of Rajasthani. Paper presented at the 17th European Conference on Modern South Asian Studies, Heidelberg, September 2002.
Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237. DOI: 10.1162/ling.2009.40.2.187.

Manzini, M. Rita. 2012. From Romance clitics to case: Split accusativity and the person case constraint. In Irene Franco, Sara Lusini & Andrés Saab (eds.), *Romance languages and linguistic theory 2010*, 1–20. Amsterdam: John Benjamins. DOI: 10.1075/rllt.4.01man.

Manzini, M. Rita. 2015. On the substantive primitives of morphosyntax and their parametrization: Northern Italian subject clitics. In Marc van Oostendorp & Henk C. van Riemsdijk (eds.), *Representing structure in phonology and syntax*, 167–194. Berlin: De Gruyter.

Manzini, M. Rita & Ludovico Franco. 2016. Goal and DOM datives. *Natural Language & Linguistic Theory* 34(1). 197–240. DOI: 10.1007/s11049-015-9303-y.

Manzini, M. Rita & Leonardo M. Savoia. 2005. *I dialetti italiani e romanci: Morfosintassi generativa*. 3 vols. Alessandria: Edizioni dell'Orso.

Manzini, M. Rita & Leonardo M. Savoia. 2007. *A unification of morphology and syntax: Studies in Romance and Albanian dialects*. London: Routledge.

Manzini, M. Rita & Leonardo M. Savoia. 2011. *Grammatical categories: Variation in Romance languages*. Cambridge: Cambridge University Press.

Manzini, M. Rita & Leonardo M. Savoia. 2017. N morphology and its interpretation: The neuter in Italian and Albanian varieties. In Anna Bloch-Rozmej & Anna Bondaruk (eds.), *Constraints on structure and derivation in syntax, phonology and morphology*, 213–236. Frankfurt am Main: Peter Lang.

Mohanan, Tara. 1995. Wordhood and lexicality: Noun incorporation in Hindi. *Natural Language and Linguistic Theory* 13(1). 75–134. DOI: 10.1007/BF00992779.

Picallo, Carme. 2008. Gender and number in Romance. *Lingue e Linguaggio* VII. 47–66.

Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge MA: MIT Press.

Renzi, Lorenzo & Laura Vanelli. 1983. *I pronomi soggetto in alcune varietà romanze*. Paola Benincà, Manlio Cortelazzo, Aldo Prosdocimi, Laura Vanelli & Alberto Zamboni (eds.). Pisa: Pacini. 120–145.

Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Roberts, Ian. 2007. *Diachronic syntax*. Oxford: Oxford University Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.

Sheehan, Michelle. 2014. Towards a parameter hierarchy for alignment. In Robert E. Santana-LaBarge (ed.), *Proceedings of the 31st West Coast Conference on Formal Linguistics*, 399–408. Somerville, MA: Cascadilla Proceedings Project.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
Svenonius, Peter. 2002. Icelandic case and the structure of events. *Journal of Comparative Germanic Linguistics* 5(1–3). 197–225. DOI: 10.1023/A:1021252206904.
Ursini, Flavia. 1988. Varietà venete in Friuli – Venezia Giulia. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der Romanistischen Linguistik*, vol. IV: Italienisch, Korsisch, Sardisch, 538–550. Tübingen: Niemeyer.
Verbeke, Saartje. 2013. *Alignment and ergativity in new Indo-Aryan languages*. Berlin: De Gruyter.
von Heusinger, Klaus & Georg A. Kaiser. 2011. Affectedness and differential object marking in Spanish. *Morphology* 21(3-4). 593–617. DOI: 10.1007/s11525-010-9177-y.

Chapter 20

High and low phases in Norwegian nominals: Evidence from ellipsis, psychologically distal demonstratives and psychologically proximal possessives

Kari Kinn
University of Bergen

This squib discusses the idea of a high and a low phase in Norwegian nominals. I argue that ellipsis phenomena and syntactic constructions yielding speaker perspective meanings corroborate the proposal that nominals may have a biphasal structure.

1 Introduction

This squib picks up on an idea most recently proposed by e.g. Cornilescu & Nicolae (2011), Simpson & Syed (2016), Simpson (2017), Syed & Simpson (2017) and Roberts (2017: 161), namely that the extended nominal projection may consist of two phases. If on the right track, this proposal gives us a new type of evidence for parallel structure in nominals and clauses (e.g. Abney 1987; Szabolcsi 1994).[1]

While Cornilescu & Nicolae (2011) and the studies by Simpson and Syed focus on Romanian and Bangla, I will discuss the idea of a high and a low nominal phase in Norwegian. Previously, Julien (2005) has made a case for biphasal nominals in Scandinavian on the basis of case-licensing and definiteness phenomena

[1]On phases in the clausal domain, see Chomsky (2000) and much subsequent work.

Kari Kinn. 2020. High and low phases in Norwegian nominals: Evidence from ellipsis, psychologically distal demonstratives and psychologically proximal possessives. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 435–450. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280665

in certain possessive constructions.[2] I will introduce two types of data that are new in the context of Norwegian: first, like Simpson (2017) and Syed & Simpson (2017), I will look at ellipsis. Then I will consider speaker-perspective meanings, which I, drawing on work by e.g. Sigurðsson (2014), take to be derived via syntactic operations at the phase edges.[3] The speaker-perspective meanings to be considered are (i) psychologically distal demonstratives (e.g. Johannessen 2008) and (ii) a possessive construction that I describe as psychologically proximal.

I assume the following structure of the extended nominal domain in Norwegian, as proposed by Julien (2005):

(1) $[_{\text{QP}}\ldots\ [_{\text{DemP}}\ldots\ [_{\text{DP}}\ldots\ [_{\text{CardP}}\ldots\ [_{\alpha\text{P}}\ldots\ [_{n\text{P}}\ldots\ [_{\text{NumP}}\ldots\ [_{\text{NP}}\ldots]]]]]]]]$

In this hierarchy, QP hosts strong quantifiers, DemP demonstratives, CardP numerals/weak quantifiers, and αP adjectives (adjectives are sitting in the specifier of the α head). DP and *n*P both contribute to definiteness; the definite suffix originates in *n*P; D mostly probes and attracts lower material, or, in the case of modified nouns, can be lexicalised by a pre-adjectival definite determiner which comes in addition to the definite suffix (so-called *double definiteness*). example (2a) illustrates the order of different elements in the nominal phrase (quantifier – demonstrative – numeral – adjective – noun with definite suffix); example (2b) shows double definiteness with a pre-adjectival definite determiner.

(2) Norwegian

a. alle disse tre gode bøk-ene
all these three good book-PL.DEF
'all these three good books'

b. den nye bok-a
the new book-DEF
'the new book'

On Julien's (2005: 12) analysis, DP, *n*P, NumP and NP are present in every DP, whereas CardP and αP are only merged when they contain lexical material. I take it that this also applies to QP and DemP.

[2] Julien argues for a low phase in addition to the more standardly assumed high phase; see Julien (2005: 4–5, 73, 202, 219) for details.

[3] Cornilescu & Nicolae (2011: 40) mention speaker-perspective meanings ("judgements by the speaker") as a characteristic of the higher nominal phase, but not of the lower one. Their arguments for a biphasal structure are based on the properties of prenominal adjectives and the so-called adjectival article construction. The main data discussed in Simpson & Syed (2016) are blocking effects on nominal-internal movement. Roberts (2017) proposes a biphasal structure in a discussion of the final-over-final condition in DP.

2 Ellipsis

Like Simpson (2017), I adopt Bošković's (2014) proposal that ellipsis is constrained by phases; more precisely, ellipsis can affect either (i) the phase itself, or (ii) the complement of the phase head (see Bošković's paper and references there for cross-linguistic evidence). On this approach, ellipsis of complements of non-phase heads is disallowed (Bošković 2014: 42). For illustration, compare (3a) and (3b) (from Bošković 2014: 56; ellipsis is marked by strikethrough):

(3) a. Betsy must have been being hassled by the police, and Peter must have been ~~being hassled …~~

b. *Betsy must have been being hassled by the police, and Peter must have been being ~~hassled …~~

In (3a), the complement of a phase head is elided (the phase head is Asp1, spelt out by *been*; see Bošković 2014: 62 for the full syntactic structure). In (3b), on the other hand, not only *been*, but also *being* is stranded; this would involve ellipsis of the complement of a non-phase head, which is not acceptable.

Some languages seem to disallow ellipsis for independent reasons even under the appropriate phasal conditions (Bošković 2014: 48); thus, ellipsis being impossible does not necessarily exclude the presence of a phase. However, according to Bošković's analysis, the possibility of ellipsis can be taken as an indication of phasehood.

2.1 Ellipsis in the higher phase

Ellipsis data suggest the presence of a phase in the higher nominal domain in Norwegian. It is, for example, possible to strand a prenominal possessive pronoun while the rest of the nominal phrase is elided, as illustrated in example (4) (the relevant nominals are in italics):

(4) Norwegian

a. Han er min beste venn, og jeg er *hans* ~~*beste venn*~~.
he is my best friend and I am his ~~best friend~~
'He is my best friend, and I am his.'

b. Jeg kom i min fineste kjole, og Anne kom i *sin* ~~*fineste kjole*~~
I came in my nicest dress and Anne came in her.REFL ~~nicest dress~~
'I was wearing my nicest dress, and Anne was wearing hers.'

I follow Julien (2005: 207, 210), who argues that prenominal possessive pronouns are first-merged in Spec-NP and move to Spec-DP (via intermediate positions). What we have in example (4) then, is ellipsis of everything below D (αP, *n*P, NumP and NP). The most obvious analysis that presents itself is that D is a phase head whose complement is elided. The analysis is illustrated (somewhat simplified) in (5):

(5) hans ~~beste venn~~
[DP [~~αP~~ [~~*n*P~~ [~~NumP~~ [~~NP~~]]]]]

It is worth noting that not only DP, but also projections located even higher in the nominal phrase can license ellipsis. This lends support to Bošković's (2014) proposal that phases are contextually defined: the edge of the phase is constituted by the highest functional projection present. Thus, in a structure where a QP is merged above DP, Q will be the phase head. An example of ellipsis with a stranded QP element (the strong quantifier *alle* 'all') is provided in example (6):[4]

(6) Norwegian
   a. Det er noen ekstra skruer i skuff-en, men ikke ta *alle*
   there are some spare screws in drawer-DEF but not take all
   ~~*de ekstra skru-ene i skuff-en*~~
   ~~the spare screw-PL.DEF in drawer-DEF~~
   'There are some spare screws in the drawer, but don't take all of them.'
   b. alle ~~de ekstra skruene i skuffen~~ [QP [~~DP~~ [~~αP~~ [~~*n*P~~ [~~NumP~~ [~~NP~~]]]]]]

2.2 Ellipsis in the lower phase

While the data presented above seem to indicate a phase headed by the topmost projection in the nominal domain, Norwegian also allows ellipsis exclusively targeting material in the lower part of the nominal. The perhaps clearest evidence of this is ellipsis following adjectives, as illustrated in (7):

[4]It is also possible to strand a strong quantifier and a demonstrative: *Alle disse ~~bøkene~~ er solgt*, lit. 'all these ~~books~~ are sold'. Many such cases can be straightforwardly analysed as ellipsis in the lower phase, which is discussed in the next section. An issue that invites further research, both empirically and theoretically, concerns ellipsis of a noun modified by an adjective in such contexts (an elided adjective would be higher than *n*P). I leave that aside here.

(7) Norwegian

a. Vi har vanligvis t-skjorter i alle farger, men *de svarte*
we have usually t-shirts in all colours but the black
~~*t-skjort-ene*~~ er utsolgt akkurat nå.
~~t-shirt-PL.DEF~~ are sold.out just now
'We normally have t-shirts in all colours, but the black ones are sold out right now.'

b. Jeg har funnet de fleste nøkl-ene vi mistet, men *alle de fire*
I have found the most key-PL.DEF we lost but all the four
små ~~*nøkl-ene*~~ er fortsatt borte.
small ~~key-PL.DEF~~ are still missing
'I have found most of the keys that we lost, but all of the four small ones are still missing.'

Recall that adjectives are located in αP, a projection below DP and CardP. On the assumption that ellipsis can only affect phases and complements of phase heads, the examples in (7) cannot be licensed by the topmost functional projection. In example (7a), the highest element present is a pre-adjectival definite determiner, and the phase head would be D. The elided material, a noun with a definite suffix, is located in *n*P, which is a complement of α, i.e. a non-phase head. In (7b), the highest element present is a strong quantifier, and the phase head would be Q. Again, the elided material is located in *n*P, a complement of α, and in addition to αP, both CardP and DP intervene between the ellipsis site and the highest phase head. To account for the data, I propose, consistently with Julien (2005) (who reaches this conclusion on different grounds), that *n*P is a phase and that the examples in (7) are phasal ellipsis of *n*P.[5] The analysis is illustrated in (8):

(8) a. de svarte ~~t-skjortene~~
[$_{DP}$ [$_{\alpha P}$ [$_{\sout{nP}...}$]]]

b. alle de fire små ~~nøklene~~
[$_{QP}$ [$_{DP}$ [$_{CardP}$ [$_{\alpha P}$ [$_{\sout{nP}...}$]]]]]

Having looked at some ellipsis data, we now turn to speaker-perspective meanings.

[5]Simpson (2017), citing Ruda (2016), makes a similar proposal for Polish and Hungarian.

3 Speaker-perspective meanings

There is now a significant body of work developing formal syntactic accounts of phenomena related to speech acts, indexicality and speaker perspective, going back to Ross's (1970) (e.g. Speas & Tenny 2003; Giorgi 2010; Hill 2014; Sigurðsson 2014; Wiltschko & Heim 2016). While many works focus exclusively on the left periphery of CP, Sigurðsson (2014: 179) connects speaker perspective (and indexicality more generally) to phases and argues that edge linkers, a type of feature that enables narrow syntax to link to context and that includes speaker and hearer features, must be present in *any phase* (although some phases may not have a full set). This proposal, which I adopt here, is consistent with the idea that phases have a parallel structure (Poletto 2006). The edge linkers most relevant for the present discussion are the following:

(9) a. Λ_A, representing the logophoric agent (speaker).
b. Λ_P, representing the logophoric patient (hearer).

If there is evidence that speaker-perspective meanings can arise from syntactic operations both in the higher and the lower part of the nominal domain, it could be taken to suggest that there are two nominal phases.

3.1 Speaker-perspective meanings in the higher phase

In the higher nominal domain, a clear example of speaker-perspective meanings is provided by so-called psychologically proximal demonstratives (PDDs), most elaborately described by Johannessen (2008) (see also further references cited there).[6] The PDD itself has the same phonological form as a 3rd person personal pronoun, but when it combines with a (human) noun, it conveys a particular meaning: it signals psychological distance. This sets it apart from regular demonstratives. Often, the PDD is used when the speaker does not know the person under discussion personally, or when they want to signal a negative attitude towards that person (cf. examples 10a,b).[7] The reference point may also be with the hearer: the speaker uses the PDD to introduce someone that they are familiar with themselves, but that the *hearer* might not know personally (cf. 10c).

[6]Other relevant speaker-perspective phenomena are possibly the emotive adjectival construction (EAC) (Halmøy 2016: 294–297) and certain uses of *sånn* 'such' (Johannessen 2012).

[7]All examples in (10) are from Johannessen (2008); notation and translations slightly adapted.

(10) Norwegian

a. jeg og Magne vi sykla jo og *han Mikkel* da
I and Magne we cycled then and he Mikkel then
'Me and Magne and that guy Mikkel, we rode our bikes' (NoTa, M, 36, Johannessen 2008: 164)

b. *hun dam-a* hun blei jo helt nerd da
she woman-DEF she became yes totally nerd then
'That woman, she became a complete nerd, you know.' (NoTa, M, 18, Johannessen 2008: 166)

c. du vet *han kjørelærer-en* jeg har?
you know he driving.teacher-DEF I have
'You know that driving instructor I have?' (NoTa, F, 18, Johannessen 2008: 164)

Johannessen (2008: 178) shows that the PDD in Norwegian cannot co-occur with the pre-adjectival definite determiner in double definiteness constructions (example 2b); the most obvious interpretation of this is that the PDD is a D element.[8] Since no higher projections are merged in the examples in (10), DP is a phase and will contain speaker and hearer features (Λ_A and Λ_P).

I propose that the encoding of psychological distance in relation to the speaker or hearer is achieved in a way similar to that of deictic gender control (Sigurðsson 2014: 185–186). An example of deictic gender control is given in (11), where the Icelandic 1st person pronoun triggers agreement in gender (fem. or masc., depending on the speaker's gender), although the pronoun itself does not exhibit any overt gender distinctions.

(11) Icelandic (Sigurðsson 2014: 185)
Ég gerði þetta sjálfur / sjálf / *sjálft
I did this self.M self.F self.N
'I did this myself.'

Deictic gender control, according to Sigurðsson, involves gendering of the speaker/hearer features. In an example such as (11), the speaker feature at the C-edge will have the value $\Lambda_{A/M}$ if the speaker is male and $\Lambda_{A/F}$ if she is female; the value is passed down to the pronoun *ég* 'I' via Agreement with the gendered speaker feature and triggers gender agreement in *sjálfur/sjálf* 'myself'. In a similar fashion, I propose that the PDDs in (10a) and (10b) get their psychologically distal

[8]Norwegian differs from Swedish and Danish in this respect; in Swedish and Danish the PDD seems to be merged higher (Johannessen 2008: 175–176), probably in DemP.

meaning via a speaker feature at the D-edge with the specification $\Lambda_{\text{A/PSYCH-DIST}}$. The PDD in (10c) differs in that the hearer, not the speaker, is the reference point; in this case, the syntactic source of the psychologically distal meaning would be the hearer feature, with the specification $\Lambda_{\text{P/PSYCH-DIST}}$.

3.2 Speaker-perspective meanings in the lower phase?

The next question is whether there is any evidence for speaker-perspective meanings arising in the *lower* nominal domain. I would like to draw attention to a particular possessive construction that might instantiate this. The construction involves a proper or common noun and a postposed 1st person possessive pronoun, and it contrasts with the PDD in that it does not convey psychological distance; on the contrary, it yields a very affectionate reading and is only appropriate in intimate contexts.[9] The construction seems to be primarily used in vocatives, and to my knowledge, it has not been discussed much in the previous literature, although it is very briefly touched upon by Julien (2016).[10,11]

Because the construction conveys the opposite of psychological distance, namely psychological proximity, I refer to it as the *psychologically proximal possessive (PPP) construction*. Some authentic examples are given in (12):[12]

(12) a. Natt'a, *Anne min*. Jeg får vel kalle deg det?
night-night Anne my I get well call you that
'Night-night, my dearest Anne. I suppose I can call you that?' (The novel *Størst av alt*, Lillian Wirak Skow, 2010)

b. *Søte Håkon vår* du fyller 8 år den 18. juni, hipp hurra for deg!
sweet Håkon our you fill 8 years the 18 June, hip hooray for you
'Our sweet Håkon, you turn 8 on 18th June, hip hooray for you!' (Birthday greeting in local newspaper, 2013)[13]

[9] This description is based on my intuitions as a native speaker of Norwegian.

[10] Julien (2016: 90) writes: "The use of first person possessive pronouns in vocatives would be an interesting topic in itself, especially since it often appears to add a flavour of endearment to the utterance, but I will leave this topic aside here."

[11] The construction bears some resemblance to the emotive adjectival construction (EAC) (Halmøy 2016: 294ff), which consists of an adjective and a noun with a definite suffix. However, there are important differences. While the EAC is characterised by the presence of an adjective, the construction to be discussed here does not necessarily contain other modifiers than the possessive. The EAC occurs independently of possessive pronouns. Moreover, the EAC does not necessarily convey affection; it can also express negative feelings.

[12] Some speakers report that they do not use the construction with proper names, but they generally seem to be familiar with it.

[13] https://www.an.no/vis/personalia/greetings/3561747 (accessed 22/11/2017).

c. Jeg vil for alltid bære med meg minne om deg *Kari min* i
I will for always carry with me memory about you Kari my in
mitt hjerte
my heart
'I will carry with me the memory of you in my heart for ever, my dearest Kari'. (Memorial webpage, 2017)[14]

d. [...] du vil aldri bli glemt, *Godgutt-en min*
you will never be forgotten good.boy-DEF my
'You will never be forgotten, my good boy' (Kennel webpage, 2015)[15]

e. [...] Elsker deg masse *venn-en min* :-)
love you lots friend-DEF my
'I love you a lot, sweetie!' (Text message)[16]

The examples in (12a–c) illustrate the PPP construction with proper names. (12a) is taken from a novel, more precisely from a scene in which a new couple are saying good night to each other. Note that the person who addresses his girlfriend as *Anne min* (lit. 'Anne my') explicitly asks for permission to do so; this highlights the intimate style of the construction. Example (12b) is from a birthday greeting to a young boy from his parents; (12c) is taken from a memorial webpage. The examples in (12d,e) illustrate the PPP construction with common nouns; (12d) is a greeting addressed to a dog on a kennel web page; (12e) is from a text message exchange between spouses. Note that when the noun in a PPP construction is modified by an adjective, like in (12b), there is no pre-adjectival definite determiner (i.e. no double definiteness); this is a characteristic of the PPP construction (and vocatives in general).[17]

Now, it could be argued that the psychologically proximal meaning of the PPP construction is a pragmatic (i.e. non-syntactic) phenomenon that automatically

[14]https://wang.vareminnesider.no/ (accessed 22/11/2017; full URL omitted because of the sensitive nature of this example).

[15]http://kennelulwazi.com/våre%20hunder/gandhi/index.html (accessed 22/11/2017).

[16]http://www.p4.no/underholdning/p4-lytternes-beste-kjerlighetsmeldinger/artikkel/336327 (accessed 22/11/2017).

[17]Occurrences of what looks like the PPP construction can be found in non-vocative contexts too: *[...] ta godt vare på Håkon vår* 'take good care of our dearest Håkon' (http://www.torgeirogkjendisene.no/10/48/2/bangkok-og-cha-am-thailand-19-29-september/, accessed 28/11/2017). However, in this paper, I limit my attention to vocatives. Postposed possessive pronouns are regularly used in Norwegian, and in non-vocative contexts a post-nominal 1st person possessive does not necessarily yield an affectionate reading; a statement like *Jeg skal besøke broren min* 'I am going to visit my brother' comes across as neutral.

follows when certain nouns (including proper nouns) are combined with a 1st person possessive pronoun. However, although possessives are regularly postposed, Norwegian also allows preposed possessive pronouns, and, in these contexts, the degree of affection and intimacy associated with the PPP construction does not arise. Imagine a situation in which a highly respected senior member of staff in a company is about to retire and a more junior member of staff is giving a speech. The speaker could be expected to say something along the lines of (13a), with a preposed possessive pronoun. The minimally different example in (13b), on the other hand, with a postposed possessive, would come across as inappropriate; the PPP construction conveys too much intimacy in the given context.[18]

(13) Norwegian

a. *Vår* kjære *Anne*, vi ønsker deg alt godt i år-ene som
our dear Anne we wish you all good in year-PL.DEF that
kommer.
come
'Our dear Anne, we wish you all the best in the years to come.'

b. # Kjære *Anne vår*, vi ønsker deg alt godt i år-ene som
dear Anne our we wish you all good in year-PL.DEF that
kommer.
come
intended meaning: 'Our dear Anne, we wish you all the best in the years to come.'

With regard to the examples with common nouns in (12d,e), one might perhaps wonder if the proximal, affectionate reading is simply due to the lexical semantics of the cited nouns; the nouns used in the PPP construction often have a "pet-name-like" feel even in other contexts. Note, however, that nouns that are neutral with respect to such inherent properties can also be used, and the proximal reading still arises, as illustrated in (14):

(14) Norwegian
Gratulerer masse med dagen lille *brannmann-en vår*!
congratulations much with day.DEF little fire.man-DEF our
'Happy birthday, our little fire man!' (Birthday greeting in local newspaper)[19]

[18] Again, this description is based on my native-speaker intuitions; I have consulted other native speakers who agree.

[19] http://www.f-b.no/vis/personalia/greetings/7330499 (accessed 22/11/2017).

Also, note that nouns whose lexical semantics are at odds with notions such as intimacy and affection seem inappropriate in the PPP construction. Cf. the contrast between (15a) and (15b):[20]

(15) Norwegian

a. Kom hit, kjærest-en min!
come here girlfriend/boyfriend-DEF my
'Come here, my love!'

b. # Gå bort, fiend-en min!
go away, enemy-DEF my
intended meaning:'Go away, my enemy!'

The data presented in (13–15) seem to suggest that the speaker-perspective meaning of the PPP construction follows from its syntax, not from pragmatics or lexical semantics. I propose the following analysis of the PPP construction.

*n*P is a phase and thus contains edge linkers. In the PPP construction, the Λ_A feature of *n*P is equipped with a proximal counterpart of the PSYCH-DIST specification responsible for the PDD construction (see above); I call this $\Lambda_{A/\text{PSYCH-PROX}}$. Now, just as in regular possessive constructions, postposing of the possessive pronoun follows from movement of the noun from its NP-internal position past the possessive, which is first-merged in Spec-NP (Julien 2005: 143), and up to the edge of *n*P. The difference is that in the PPP construction, the possessive pronoun Agrees with $\Lambda_{A/\text{PSYCH-PROX}}$; this yields the psychologically proximal reading. A sketch of the relevant pieces of structure is given in (16) (for convenience I mark movement with traces and the Agreement relation between the possessive and the edge linker with an arrow):[21]

(16) Anne min

$[_{nP}\ [_{n}\ \Lambda_{A/\text{PSYCH-PROX}}\ \text{Anne}_i\ ]\ [_{\text{NumP}}\ [_{\text{Num}}\ t_i\ ]\ [_{\text{NP}}\ \text{min}_{A/\text{PSYCH-PROX}}\ [_{\text{N}}\ t_i\ ]]]]$

Admittedly, it is a challenge to show unequivocally that a syntactic operation in *n*P is responsible for the speaker-perspective meaning in the PPP construction; it does not have overt, phase-internal morphological or syntactic effects (unlike the PDD in the DP phase, which has a special form). A full investigation into this issue must be left for future research; in particular, it is important to

[20] Example (15b) would sound stylistically marked even with a prenominal possessive pronoun, but not as inappropriate as it does with a postnominal possessive, according to my judgement.

[21] I follow Julien (2005) in analysing the movement of the noun as head movement.

consider possible interactions with the higher phase, for which the concept of speaker/hearer-perspective is currently more established.[22] However, I would like to point out some possible indications that the PPP construction indeed gets its speaker-perspective meaning from an edge linker in *n*P.

First, as shown in example (12b), repeated below in (17), the PPP construction is compatible with a prenominal adjective:

(17) Norwegian
Søte Håkon vår du fyller 8 år den 18. juni, hipp hurra for deg!
sweet Håkon our you fill 8 years the 18 June, hip hooray for you
'Our sweet Håkon, you turn 8 on 18th June, hip hooray for you!'
(Birthday greeting in local newspaper)

Since adjectives are merged in Spec-αP (cf. example 1), this suggests that the noun does not leave *n*P, and that the postnominal possessive pronoun stays in an even lower position, in Spec-NP. This does not in itself exclude the possibility of interaction with edge linkers in the higher phase, but it is certainly compatible with *n*P as the locus of the $\Lambda_{\text{A/PSYCH-PROX}}$ feature. Second, in terms of its meaning, the PPP construction bears resemblance to diminutives; cross-linguistically it is common for diminutives to mark affection (see Jurafsky 1996 and references there). Diminutive formation is often thought to take place in a low position in the nominal; Wiltschko (2006) proposes, on independent grounds, that diminutives (e.g. in German) are light nouns in n, comparable to *n* in the framework adopted here. To me it seems plausible that the PPP construction and diminutives have structural similarities, so that arguments for diminutive formation in *n*P are also relevant for the PPP construction. I hypothesise that a speaker-perspective *n*-edge-linker is involved in diminutives marking affection, and that the PPP construction arises via syntactic operations involving the same feature. The similarity between the PPP construction and diminutives finds some support in orthography: the PPP construction can occasionally be found with a hyphen linking the noun and the possessive pronoun, as shown in (18):[23]

[22] In vocatives, the higher phase is probably not DP (Longobardi 1994); the lack of a D-layer in Norwegian vocatives is evidenced by the lack of a pre-adjectival definite determiner with modified nouns (cf. examples 12b and 14). One could perhaps argue that vocatives are small (reduced) nominals, a parallel to small clauses (Pereltsvaig 2006), consisting of the lower phase only. However, recent research argues for a Voc projection that encodes the vocative function (e.g. Hill 2007; 2014; Espinal 2013; Stavrou 2014; Julien 2014; 2016). VocP would be a phase if phases are contextually defined.

[23] I have only seen this orthographic pattern in PPP constructions involving proper names.

(18) Norwegian
Gratulerer med dagen, kjære søte fine nydelige *Marianne-min*
congratulations with day.DEF dear sweet lovely beautiful Marianne-my
'Happy birthday, my dear, sweet, lovely, beautiful Marianne' (Birthday greeting on Facebook, 2017)

The hyphen suggests a tight connection between the noun and the possessive; it could mean that the possessive pronoun in the PPP construction is a diminutive suffix (see also Lødrup 2011 and Svenonius 2017).

Many Norwegian speakers can use the suffixes *-mor* 'mother' and *-far* 'father' to form what can be described as affectionate diminutive forms of proper names. Interestingly, some of the speakers that I have informally consulted report a reluctance to use the diminutive forms in the PPP construction (I share this intuition); cf. (19):

(19) a. PPP construction
Anne min
Anne my
b. Diminutive
Annemor
Anne.DIM
c. Diminutive used in PPP construction
?? Annemor min
Anne.DIM my

There are also speakers who accept (19c); clearly, further investigations into the inter-speaker variation and its underlying reasons are needed. However, a possible interpretation of the dubious status of (19c) could be that it is not possible for both the diminutive suffix *-mor* and the possessive pronoun of the PPP to enter into a relationship with the $\Lambda_{\text{A/PSYCH-PROX}}$ feature at the *n*-edge at the same time.

4 Conclusion

In this squib, I have discussed the idea that Norwegian nominal phrases, like clauses, can consist of both a high and a low phase. I have shown that Norwegian allows ellipsis both in the higher and lower nominal domain; according to Bošković (2014), ellipsis is an indication of phasehood. Moreover, inspired by Sigurðsson (2014), I have argued that speaker-perspective meanings arise via syntactic operations in the higher nominal domain (psychologically distal demonstratives, Johannessen 2008), and, somewhat more tentatively, also in the lower

part of the nominal (*n*P) (in the psychologically proximal possessive construction). Assuming that speaker-perspective meanings are related to edge-linkers at phase edges (Sigurðsson 2014), this also corroborates a biphasal structure.

Acknowledgements

I would like to thank Ian Roberts, whom I had the pleasure of working with on the ReCoS project at the University of Cambridge, and who remains a great inspiration. This paper came about from ideas that were discussed during the 6th CamCoS conference. For comments and suggestions, I thank the editors of this volume, two anonymous reviewers, Janne Bondi Johannessen, Jan Terje Faarlund and Per Erik Solberg. Any remaining errors are my own.

Abbreviations

DEF	definite	N	neuter
DIM	diminutive	PDD	psychologically distal demonstrative
EAC	emotive adjectival construction	PL	plural
F	feminine	PPP	psychologically proximal possessive
M	masculine		

References

Abney, Steven. 1987. *The English noun phrase in its sentential aspect*. MIT. (Doctoral dissertation).

Bošković, Željko. 2014. Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45(1). 27–89. DOI: 10.1162/LING_a_00148.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.

Cornilescu, Alexandra & Alexandru Nicolae. 2011. Nominal peripheries and phase structure in the Romanian *DP*. *Revue roumaine de linguistique* 56(1). 35–68.

Espinal, M. Teresa. 2013. On the structure of vocatives. In Barbara Sonnenhauser & Patrizia Noel Aziz Hanna (eds.), *Vocative! Addressing between system and performance*, 109–132. Berlin: De Gruyter.

Giorgi, Alessandra. 2010. *About the speaker: Towards a syntax of indexicality*. Oxford: Oxford University Press.

Halmøy, Madeleine. 2016. *The Norwegian nominal system: A neo-Saussurean perspective*. Berlin: De Gruyter.

Hill, Virginia. 2007. Vocatives and the pragmatics–syntax interface. *Lingua* 117(12). 2077–2105. DOI: 10.1016/j.lingua.2007.01.002.

Hill, Virginia. 2014. *Vocatives: How syntax meets with pragmatics*. Leiden: Brill.

Johannessen, Janne Bondi. 2008. The pronominal psychological demonstrative in Scandinavian: Its syntax, semantics and pragmatics. *Nordic Journal of Linguistics* 31(2). 161–192.

Johannessen, Janne Bondi. 2012. Han professoren og sånn festskrift – om modale determinativer. In Hans-Olav Enger (ed.), *Grammatikk, bruk og norm: Festskrift til Svein Lie på 70-årsdagen 15. april 2012*, 151–172. Oslo: Novus.

Julien, Marit. 2005. *Nominal phrases from a Scandinavian perspective*. Amsterdam: John Benjamins. DOI: 10.1075/la.87.

Julien, Marit. 2014. Vokativar i norsk. *Norsk lingvistisk tidsskrift* 32. 130–165.

Julien, Marit. 2016. Possessive predicational vocatives in Scandinavian. *Journal of Comparative Germanic Linguistics* 19(2). 75–108. DOI: 10.1007/s10828-016-9081-x.

Jurafsky, Daniel. 1996. Universal tendencies in the semantics of the diminutive. *Language* 72(3). 533–578. DOI: 10.2307/416278.

Lødrup, Helge. 2011. Norwegian possessive pronouns: Phrases, words or suffixes? In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG11 conference*, 339–359. Stanford, CA: CSLI Publications.

Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry* 25(4). 609–665.

Pereltsvaig, Asya. 2006. Small nominals. *Natural Language and Linguistic Theory* 24(2). 433–500. DOI: 10.1007/s11049-005-3820-z.

Poletto, Cecilia. 2006. Parallel phases: A study on the high and low periphery of Old Italian. In Mara Frascarelli (ed.), *Phases of interpretation*, 261–294. Berlin: Mouton de Gruyter.

Roberts, Ian. 2017. The final-over-final condition in DP: Universal 20 and the nature of demonstratives. In Michelle Sheehan, Theresa Biberauer, Ian Roberts & Anders Holmberg (eds.), *The final-over-final condition*, 151–186. Cambridge, MA: MIT Press.

Ross, John Robert. 1970. On declarative sentences. In Roderick Jacobs & Peter Rosenbaum (eds.), *Readings in English transformational grammar*, 222–277. Waltham, Mass: Ginn & Co.

Ruda, Marta. 2016. NP ellipsis (effects) in Polish and Hungarian: FFs on Fs, Agree, and chain reduction. *The Linguistic Review* 33(4). 649–677.

Sigurðsson, Halldór Ármann. 2014. Context-linked grammar. *Language Sciences* 46. 175–188.

Simpson, Andrew. 2017. *Parallels in the structure of phases in clausal and nominal domains*. Paper presented at CamCoS 6, Cambridge, 6 May.

Simpson, Andrew & Saurov Syed. 2016. Blocking effects of higher numerals in Bangla: A phase-based analysis. *Linguistic Inquiry* 47(4). 754–763. DOI: 10.1162/LING_a_00231.

Speas, Peggy & Carol L. Tenny. 2003. Configurational properties of point of view roles. In Anna Maria Di Sciullo (ed.), *Asymmetry in grammar*, vol. 1: Syntax and semantics, 315–344. Amsterdam: John Benjamins. DOI: 10.1075/la.57.15spe.

Stavrou, Melita. 2014. About the vocative. In Lilia Schürcks, Anastasia Giannakidou & Urtzi Etxeberria (eds.), *The nominal structure in Slavic and beyond*, 299–342. Berlin: De Gruyter.

Svenonius, Peter. 2017. Declension class and the Norwegian definite suffix. In Vera Gribanova & Stephanie S. Shih (eds.), *The morphosyntax–morphology connection*, 325–359. Oxford: Oxford University Press.

Syed, Saurov & Andrew Simpson. 2017. On the DP/NP status of nominal projections in Bangla: Consequences for the theory of phases. *Glossa: a journal of general linguistics* 2(1), 68. DOI: 10.5334/gjgl.186.

Szabolcsi, Anna. 1994. The noun phrase. In Ferenc Kiefer & Katalin É. Kiss (eds.), *The syntactic structure of Hungarian* (Syntax and Semantics 27), 179–274. New York: Academic Press.

Wiltschko, Martina. 2006. Why should diminutives count? In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinheinz & Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, 669–679. Berlin: De Gruyter.

Wiltschko, Martina & Johannes Heim. 2016. The syntax of confirmationals: A neo-performative analysis. In Gunther Kaltenböck, Evelien Keizer & Arne Lohmann (eds.), *Outside the clause: Form and function of extra-clausal constituents*, 300–340. Amsterdam: John Benjamins. DOI: 10.1075/slcs.178.11wil.

Chapter 21

Rethinking microvariation in Romance demonstrative systems

Adam Ledgeway
University of Cambridge

This article explores the formal and functional organization of Romance demonstrative systems, providing a detailed empirical overview of the vast microvariation attested in standard and non-standard Romance varieties. Despite highlighting a considerable number of distinct demonstrative systems based on different superficial person contrasts, it is argued that the underlying number of systems can effectively be reduced to a much smaller number of systems based on a finite number of options. In particular, it is argued that the feature geometric analysis of person developed by Harley & Ritter (2002) makes some specific predictions about the range and types of person combinations, and hence by implication also the types and natural classes of demonstrative systems, that are cross-linguistically available. Adopting these assumptions, it is argued that these differing person feature specifications can be profitably modelled in terms of a set of hierarchically-organized interrelated parametric options in accordance with much recent work developed within the ReCoS group.

1 Introduction and general remarks

Traditional descriptions of Romance demonstrative systems highlight a major distinction between binary (cf. 1a below) and ternary (cf. 1b below) person-based systems (cf. Meyer-Lübke 1895: 645–647; Meyer-Lübke 1900: 95–99; Lausberg 1976: 135–140; Lyons 1999: 109–111; Stavinschi 2009: 37–46; Alkire & Rosen 2010: 301f):

Adam Ledgeway. 2020. Rethinking microvariation in Romance demonstrative systems. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 451–490. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280667

(1) a. Romanian (personal knowledge)
acest / acel copil
this that child
'This / That child'

b. Asturian (Academia de la Llingua Asturiana 2001)
esti / esi / aquel neñu
this that.2 that.3 child
'This / That (near you) / That child'

However, a more detailed examination of microvariation in this area reveals a more complex and varied picture (Ledgeway 2004; 2015; Ledgeway & Smith 2016), including both binary and ternary systems in the southern and northern Romània, respectively, and a variety of analytic formations. In what follows I shall review (cf. §§2–5) the various functional and formal organizations of a number of Romance demonstrative systems which, to varying degrees, correspond to different diachronic and diatopic groupings. Despite the identification of some quite considerable microvariation in the formal and functional structure of different Romance demonstrative systems, I shall show how the vast microvariation revealed by this overview of the Romance evidence can be effectively interpreted and reduced to a finite number of options. Following ideas proposed by Roberts & Holmberg (2010) and Roberts (2012), and further developed by the *Rethinking comparative syntax* (ReCoS) research group led by Ian Roberts,[1] I shall explore (§6.2) how a scalar interpretation of microvariation modelled in terms of parametric hierarchies can make immediate sense of the Romance data and, at the same time, make some strong predictions about the possible combinations and the markedness relations of different person features and, ultimately, how these formally map onto different demonstrative systems.

2 Binary systems

2.1 Type B1 systems

Many predominantly northern Romance varieties display a person-based binary demonstrative system (Table 21.1), in which referents which fall within the spatial, temporal or psychological domain of the speaker (the deictic centre) are

[1]For information about the ReCoS project, including recent publications, see http://recos-dtal.mml.cam.ac.uk/.

marked by a reflex of (ECCE/ECCU/*akke/*akkʊ̆-)ISTUM '(behold!) this' > (AQU)ESTO and those associated with the non-discourse participants are picked out by a reflex of (ECCE/ECCU/*akke/*akkʊ̆-)ILLUM > '(behold!) that' > (AQU)ELLO.[2]

Table 21.1: B1 systems

	Speaker[a]	Non-discourse partic.[b]
Occitan	*aqueste*	*aquel/aquéu*
Gascon (Testerin)	*aquis*	*aquits*
Ladin	*chësc*	*chël*
Northern Italian dialects	*(cu)st*	*cul*
Italian	*questo*	*quello*
Vegliot	*kost*	*kol*
Romanian	*acesta*	*acela*
Southern Daco-Romance/Moldovan	*aista*	*ăla*
Megleno-Romance	*tsista*	*tsela*

[a](ECCE/ECCU/*akke/*akkʊ̆-) ISTUM
[b](ECCE/ECCU/*akke/*akkʊ̆-) ILLUM

In these varieties the role of the addressee is not formally encoded, inasmuch as referents associated with the addressee can a priori be marked either by AQUESTO (cf. 2a) or AQUELLO (cf. 2b) in accordance with whether they are subjectively perceived to fall within the deictic centre or not (Irsara 2009: 71–77).

(2) Veronese

a. Tira via ste man!
pull.IMP.2SG away these hands
'Take these hands (of yours) away!'

b. No vardarme co quei oci
not look.INF=me with those eyes
'Don't look at me with those eyes (of yours)!'

[2]For extensive bibliography of the relevant varieties, see Ledgeway & Smith (2016: 879). When individual language forms are not of immediate interest, reflexes of (ECCE/ECCU/*akke/*akkʊ̆-)ISTE, (ECCU-)TI(BI)-ISTE, (ECCE/ECCU/*akke/*akkʊ̆-)IPSE and (ECCE/ECCU/*akke/*akkʊ̆-)ILLE are indicated with the following broadly neutral Romance forms in small caps (AQU)ESTO, (CO)TESTO, (AQU)ESSO, and (AQU)ELLO.

These broad developments can be understood in terms of the analysis proposed in Vincent (1999) who, inspired by the conception of the deictic space (cf. Figure 21.1) proposed by Benveniste (1946), argues that with the loss of the Classical Latin speaker-oriented demonstrative HIC 'this' – in large part due to the erosive effects of phonetic change – the territory HIC covered immediately fell within the domain of the addressee-oriented term ISTE.

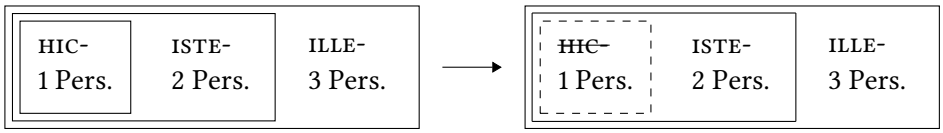


Figure 21.1: Effects of loss of HIC

This explains why in Romance ISTE comes to mark the role of the speaker, giving rise to B1 systems. However, this development necessarily presupposes that, before reflexes of ISTE grammaticalized as markers of first-person deixis, there was an earlier stage in which such reflexes marked the shared deictic spheres of both discourse participants, a stage directly attested in Old French where *(i)cist/(i)cil* mark, respectively, "proximity (to both the speaker and the addressee) […] and distance (in relation to those not present, the third person)" (CNRTL 2012: s.v. ce2; cf. also Nyrop 1925a: 293f), and which survives today in many Raeto-Romance varieties such as Surselvan and Vallader (Sornicola 2011: §2.2.1.1). We can therefore further distinguish between type $B1_A$ (Old French, Raeto-Romance) and type $B1_B$ (the rest) systems.

Formally, Italo-Romance type B1 systems typically mark a distinction between pronominal and adnominal uses of the speaker-oriented term, deploying predominantly or obligatorily ECCU-reinforced forms in pronominal uses and non-reinforced forms in adnominal functions (Rohlfs 1968: 206; Irsara 2009: 13f): Lombard *chest* vs *st.* Outside Italo-Romance, by contrast, the simple and reinforced forms appear to be in free variation (Sornicola 2011: §2.2.1.1), as in the case of Old French (cf. 3; Nyrop 1925b: 416), Old Occitan (*est* vs *(ai)cest/aquest*; Grandgent 1909: 109), and modern Romanian (*acesta/ăsta* vs *acel/ăla*), albeit subject to register variation with concomitant positional differences in the latter case where the distribution of simple vs reinforced forms is subject to considerable diachronic, diatopic, and diamesic variation (Sandfeld & Olgen 2019: 157, 161f; Caragiu Marioţeanu 1989: 418; Manea 2012: 503–505).

(3) Old French (*Strasbourg oaths*)
d' ist di / cist meon fradre
from this day this my brother
'From this day on' vs. 'This brother of mine'

Also frequent in type B1 systems (cf. Arnaud & Morin 1920: 282f; Vanelli 1997: 112; Marcato & Ursini 1998: 84, 182; Salvat 1998: 65; Bernstein 1997; Irsara 2009: 34–48, 107f; Cordin 2016) are analytic formations with the spatio-personal adverbs 'here' (*qua*, *(ei)ça(i)*, *aicí chì*, *sì*) and 'there' (*(ei)là(i)*, *alà, lì, le*) which, although originally emphatic in nature, are today generally unmarked and often preferred. In most varieties the adverb follows the demonstrative pronoun (cf. 4a,b) or the NP in a discontinuous structure (cf. 4c).

(4) a. Vegliot (Bartoli 1906)
kost káu̯k fero un músč
this here is a moss
'This one is a moss.'

b. Valéian, southeastern Occitan (Arnaud & Morin 1920)
aquéstou d eiçài / aqueous d' eilài
this.one of here that.one of there
'This one' vs. 'That one'

c. Genoese (Forner 1997)
quella scinfonìa lì
that symphony there
'That symphony'

In Emilia-Romagna (cf. 5a), the locative is frequently preceded by the relative/complementizer *che/ca* 'that', a relic of an erstwhile copular structure "… that [is] here/there" (cf. Rohlfs 1968: 206; Foresti 1988: 581), a structure also found in some Tuscan varieties (Rohlfs 1968: 203). Notable is the positional freedom of the locative in Reggiano and Ferrarese where it is also frequently preposed (cf. 5b). Some Occitan (especially Provençal) varieties use such adverbs to introduce subtle distinctions which are not canonically marked by the type B1 system (Koschwitz 1894: 88f; Ronjat 1913: 33; Salvat 1998: 65); thus alongside the *aquest(e)*/*aquéu* opposition, one can further distinguish within the conversational dyad between the speaker *aquéu-d'aqui* (lit. 'that.one-from here') and the addressee *aquéu-d'eila* ('that.one-of there').

(5) a. Emilia-Romagna (Foresti 1988)
ʃta dona ka kwe, kla dona ka le
this woman that here that woman that there
'This woman, that woman'

b. Ferrarese (Foresti 1988)
ʃti oman ki / ki ʃti oman
these men here here these men
'These men'

2.2 Type $B1_C$ systems

Northern Italian dialects also present another binary demonstrative system, henceforth type $B1_C$, the deictic organization of which is identical to that of type $B1_B$ in that it involves a simple [±1person] opposition,[3] but which formally differs quite markedly from type $B1_B$ systems. In the latter systems the demonstrative was shown to be very frequently reinforced by a spatio-personal adverb, a usage which seems to have become so entrenched over time in type $B1_C$ varieties that all deictic force has been transferred to the adverb, reducing the demonstrative to a mere marker of definiteness. This is evidenced by the fact that we find a mismatch between the original person value of the former demonstrative and that of the accompanying locative (Berruto 1974: 21; Azaretti 1982: 171; Parry 1997: 241; Vanelli 1997: 112f; Irsara 2009: 107–110), leading to the generalization either of (AQU)ESTO (cf. 6a) or AQUELLO (cf. 6b).

(6) a. Ligurian (Azaretti 1982)
stu ki invece de stu là
this here instead of this there
'This one instead of that one'

b. Friulian (Vanelli 1997)
kel libri ka / la
that book here there
'This/That book'

Interesting in this respect are some Francoprovençal dialects, such as in the Val Terbi (Jura) where the adverbs *-si* 'here' and *-li* 'there' are (optionally) employed with a suppletive paradigm (Kjellman 1928; Butz 1981: 85) that marries

[3]Here and throughout the empirical presentation, I occasionally use for informal descriptive purposes unbundled person features such as [±1], [±2] and [±3], although I shall argue in §6.2 that from a formal perspective such characterizations are ultimately flawed.

together reflexes of ISTE 'this' in the singular (*stu(-si/-li)*) with reflexes of ECCE-ILLE 'that' in the plural (*sé(-si/-li)*). Some varieties show a transitional behaviour with respect to the diachronic shift from type $B1_B$ to $B1_C$. For instance, the demonstrative system of modern Milanese is essentially of type $B1_B$ (Ledgeway 2015: 79), but also shows a progressive neutralization of adnominal *quel* 'that' which may be used with *chì* 'here' to reference the deictic sphere of the speaker (Irsara 2009: 108f).

Historically, French also belongs here inasmuch as, following the loss of the earlier *cist/cil* opposition with the refunctionalization of the latter term as the pronominal variant, the relevant binary distinction was initially maintained in conjunction with the ambiguous adnominal *ce* 'this/that' through its combination with the postnominal locatives *-(i)ci* 'here' and *-là* 'there' (Brunot 1899: 325; Nyrop 1925b: 424f; Nyrop 1925a: 292f; Price 1971: 123, 126), which became obligatory with the unmodified pronominal forms *celui-ci/-là* 'this/that one'. In the modern language, however, *-là* has encroached upon much of the territory of *-ci* (cf. 7a; Price 1971: 127; Smith 1995: §2), such that the modern French one-term system has neutralized distance distinctions (cf. 7b; Da Milano 2007: §3.4; Rowlett 2007: 67f). Where necessary, remoteness can be marked through adverbs such as *là-bas* 'over there' (cf. 7c; Brault 2004), though not actually integrated into the deictic system in that *là-bas* does not contrast with, say, *ce plat-là*, nor does it form an immediate constituent with *plat* in (7c) but, rather, modifies *ce plat* (for thorough discussion, see Smith 1995: n.5).

(7) Modern French (Smith 1995)

a. Je suis là
I am there
'I am here.'

b. ce plat-là
this dish-there
"This/That dish"

c. ce plat là-bas
this dish over.there
"That dish over there"

3 Ternary systems

3.1 Type T1 systems

In Figure 21.1 we saw how, following Vincent (1999), with the loss of HIC the deictic sphere of the speaker naturally fell within the domain of the original addressee-oriented term ISTE. Implicit in this analysis is the further implication that, initially at least, ISTE did not come to mark solely the role of the speaker as eventually happened in type $B1_{B/C}$ systems, but by inheriting the deictic territory of HIC, it saw an expansion in its original range of reference beyond the addressee to now also include the speaker (Ledgeway 2004: 91–96), producing a parallel expansion of the deictic centre, originally anchored exclusively to the speaker, to now also include the addressee (cf. type $B1_A$). The result in many Ibero-Romance and central-southern Italo-Romance varieties is an inclusive first-person term ((A)QU)ESTO (Ledgeway 2004: 78–91), as preserved in Old Neapolitan *(chi)sto* (Ledgeway 2009: 200–205) which readily marks inalienable referents pertaining uniquely to the addressee (cf. 8a), though second-person deixis could be marked separately where required (e.g. ambiguity, contrast) by innovative (ECCU)IPSU > (QU)ESSO forms, witness the contrasting deictic spheres of the speaker and addressee marked respectively by Old Neapolitan *sto* and *sso* in (8b).

(8) Old Neapolitan (Ledgeway 2009)

a. Se tu vuoi fare a muodo de 'sta capo pazza
if you want do.INF to way of this head mad
'If you want to act according to this mad mind (of yours).'

b. iettame cinco ventose a 'ste lavra co 'ssa bella
throw.IMP.2SG=me five kisses to these lips with this beautiful
vocca!
mouth
'place five kisses on these lips (of mine) with that beautiful mouth (of yours)!'

Jungbluth (2003; to appear) identifies an identical distribution for the first two terms *este* and *ese* of the European Spanish ternary system where,[4] contrary to traditional studies which treat the system as simply person-oriented (*Diccionario de la lengua española* 1970: 109, 581, 585; Eguren 1999: 940; Eguren 2012: 557) or

[4] Cf. also Gutiérrez-Rexach (2002; 2005), Langacker (1990: 52), Gómez Sánchez & Jungbluth (2015: 245–247).

distance-oriented (Hottenroth 1982; Diessel 1999: 39), she highlights how in default face-to-face encounters the deictic spheres of both discourse participants (the *inside* space) are indiscriminately marked by *este* (cf. 9), with referents situated outside the conversational dyad (the outside space) marked by the third term *aquel*.

(9) European Spanish (Jungbluth 2003)
¡AH! Pues este reloj es BUENO
ah then this watch is good
'Ah! Well that watch [that you're wearing] is shipshape!'

That the deictic domain marked by ISTE must have come to include both the speaker and addressee in late Latin/early Romance is reflected formally in the development of the Tuscan and Umbrian addressee-oriented forms *codesto/cotesto* and *tisto*. Significantly, both these second-person forms are forged from a form of ISTE, reinforced in turn by an explicit second-person marker, namely (ECCU)TI(BI) '(behold) for you'. If in early Romance ISTE only marked speaker-oriented deixis, its presence in the term used to mark the addressee in Tuscan and Umbrian would remain inexplicable. Instead, ISTE in Tuscany and Umbria, as in many Romance dialects (Ledgeway 2004), must have generalized as a demonstrative marking the deictic domains of both discourse participants. However, in certain cases (e.g., ambiguity, contrast) speakers would have felt it necessary to clearly distinguish between the deictic domains of the addressee and speaker, a distinction which could have been marked by simply adding a second-person marker such as (ECCU)TI(BI) to ISTE. This mechanism in time then would have become conventionalized, giving rise to the modern lexicalized forms *codesto/cotesto* and *tisto*.

As illustrated in detail in Ledgeway (2004), in type T1 systems the fundamental deictic contrast therefore involves a binary opposition between AQUESTO [−3person] and AQUELLO [+3person], inasmuch as the unmarked addressee-oriented demonstrative is AQUESTO, the competing AQUESSO/(CO)TESTO forms constituting marked variants restricted to contexts where particular attention has to be drawn to the addressee. This explains why the textual distribution of the latter forms is systematically very low in all statistical studies to date: 4.8% for 15th-c. Neapolitan (Vincent 1999), 6.4% for 13th–18th-c. Neapolitan (Ledgeway 2004: 89), and 4.3% for 19th-c. Sicilian (Ledgeway 2004: 92). Indeed, it has not gone unnoticed in descriptions of southern Italian dialects and Tuscan-Italian (Ledgeway 2004: 68–70), Peninsular Spanish (Eguren 1999: fn. 31; Eguren 2012: 558f; Gutiérrez-Rexach 2002; 2005) and European Portuguese (Teyssier 1980; Salvi

2011: 325) how in many apparently ternary systems the use of the addressee-oriented term proves somewhat restricted, ultimately pointing to the essential binary organization of the systems. Indeed, Jungbluth (to appear: §3.1) and Gómez Sánchez & Jungbluth (2015: 245f) observe how in face-to-face encounters in European Spanish addressee-oriented deixis is only exceptionally marked by *ese*, rather than the more usual *este*, thereby subdividing the *inside* space of the conversational dyad, when: (i) the speaker focuses on referents in contact with the addressee's body; (ii) strong emotions are aroused in relation to divisive disputes or refusals; and (iii) quarrels about possessions are at stake.

As already noted, type T1 demonstrative systems are principally found in Ibero-Romance, large areas of southern Italy, and more limitedly in some Occitan varieties. Representative of the former group is European Portuguese where, in contrast to traditional person-based treatments (Cunha & Cintra 1984); Tláskal 1994: 166; Topa Valentim 2015), Jungbluth (2000: 93–95; 2003: 31; to appear: §3.2.3.2) characterizes the demonstrative system in terms of a fundamental binary opposition on a par with that analysed above for European Spanish which contrasts the *inside* space of the conversational dyad (*este*) with the *outside* space of non-discourse participants (*aquele*), with *esse* reserved for marked addressee-oriented uses (cf. Carvalho 1976: 247–251). A similar picture arises for Asturian which, although standardly described as displaying a person-based system (García de Diego 1946: 166; Frías Conde 1999: 8; Academia de la Llingua Asturiana 2001: 103), employs the first term *esti* to mark referents that fall within the deictic spheres of both the speaker and the hearer (Academia de la Llingua Asturiana 2001: 105). Similar observations apply to Galician *(aqu)iste* (/*(aqu)este*) / *(aqu)ise* /*(aqu)ese* / *aquil* (/*aquel*) (García de Diego 1946: 94), Leonese *este/ese/aquel* (Zamora Vicente 1967: 176) and Aragonese *este/eše(/iše)/aquel* (García de Diego 1946: 260).

Almost without exception type T1 systems in southern Italy, at least in the modern dialects, formally mark the pronominal/adnominal paradigmatic opposition through the use of ECCU-reinforced and non-reinforced forms of (AQU)ESTO and (AQU)ESSO (Ledgeway 2004: 71–74), e.g. Anzese *kwéstə/stú*, *kwéssə/ssú*. Within Ibero-Romance the distribution of simple and reinforced forms in the first two terms (*(aqu)este*, *(aqu)e(s)se*) is generally subject to diachronic and diatopic variation (cf. use of *aqueste/aquesse* alongside of *este/e(s)se* in Old Portuguese and Spanish; Kjellman 1928: 5; Teyssier 1980: 39; Penny 2000: 211; Sornicola 2011: §2.2.1.1), with reinforced forms in the first two terms today surviving only in rural dialects.

Spatio-personal adverbial reinforcement is much less frequent in type T1 systems, generally assuming, in contrast to B1 systems, an emphatic interpretation and more frequently found with the pronominal demonstratives: Sicilian *chistu*

cà, *chissu dd(u)ocu*, *chiddu ddà* (Pitré & Wentrup 1995: 72). In Ibero-Romance, alongside the canonical, unmarked prenominal position the demonstrative may also occur in postnominal position in the modern languages in conjunction with a prenominal definite article (Butt & Benjamin 1994: 84; Brugè 1996; Brugè 2002; Eguren 2012: 559–561; Ledgeway 2012: 113f), witness the Asturian alternations in (10a; Academia de la Llingua Asturiana 2001: 104f). Unlike in Romanian where postnominal demonstratives are immediately postnominal (cf. 10b), in Ibero-Romance postnominal demonstratives can either precede or follow postnominal direct modifiers (cf. 10c). A further difference is that whereas in Romanian the postnominal position is very frequent in neutral registers where it may also license contrastive focus, in Ibero-Romance the postnominal position is marked, typically associated with topical interpretations and pejorative readings, hence its incompatibility with contrastive focus (cf. 10d; Roca 2009).

(10) a. Asturian (Academia de la Llingua Asturiana 2001)
esti homi / l' homi esti
this man the man this
'This man'

b. Romanian (personal knowledge)
cartea aceasta veche (*aceasta)
book.the this old this
'This old book'

c. Spanish (personal knowledge)
el libro (este) viejo (este)
the book this old this
'This old book'

d. Spanish (personal knowledge)
este libro / ??el libro este, no aquel
this book the book this not that.one
'This book, not that one'

3.2 Type T2 systems

Alongside type T1 systems we also find, especially throughout most of central Italy (Vignuzzi 1988: 616; Vignuzzi 1997: 315; Loporcaro 2009: 129) and in Abruzzo and Molise (Marinucci 1988: 647; Stavinschi 2009: 161f), a genuinely ternary demonstrative system (viz. type T2), in which reference to the deictic sphere of the addressee is no longer canonically marked by (AQU)ESTO as in type T1 systems,

but has now come to be systematically marked by (AQU)ESSO. Representative examples among the many central dialects reported in this respect include Maceratese (*kwiſtu/kissu/kwillu*; Regnicoli 1995: 232), the southern Umbrian dialect of Cascia (*vistu* (*kuistu*)/*vissu* (*kuissu*)/*villu* (*kuillu*); Moretti 1987: 123), and the central Laziale dialect of Sant'Oreste (*kweſtu/kwessu/kwellu*; Cimarra 1998: 74). For Abruzzo and Molise, Finamore (1893: 22) reports contrasts such as those in (11a) below for Abruzzese (cf. also Verratti 1968: 47), and Vincelli (1995: 75) notes for the Molisan dialect of Casacalenda that in the ternary opposition (11b) each of the three demonstratives refers exclusively to the spatio-personal domains of the speaker, addressee, and the non-discourse participants, respectively.

(11) a. Abruzzese (Finamore 1893)
šta case / ssa mane / cla case
this house that hand that house
'This house' vs. 'That hand (of yours)' vs. 'That house'

b. Molisan (Vincelli 1995)
cuisc_t' uóve / cuiss' albere / cuill'u maleditte
this egg this tree that damned.one
'This egg' vs. 'That tree' vs. 'That damned man'

Outside central Italy and Abruzzo and Molise, type T2 systems are distributed somewhat less densely across Basilicata (Lüdtke 1979: 29), northern Puglia (Valente & Mancarella 1975: 27, 60), central-southern Calabria (Ledgeway 2004: 92 n.41, 107) and Sicily (Leone 1995: 29, 41). Outside Italo-Romance, T2 systems are even less frequent, but are reported for: (i) Old Catalan (e.g. *(aqu)est, (aqu)eix*, *aquell*, and still occasionally found in the modern literary language) and some conservative (eastern and southern) Catalan varieties (Badia i Margarit 1995: 500f; Duarte i Montserrat & Alsina i Keith 1986: 81; Veny 1991: 256; Wheeler et al. 1999: 107; Moll 2006: 179; Nogué-Serrano 2015: 208f); and (ii) some Sardinian dialects (Blasco Ferrer 1988: 839; Jones 1993: 34, 203; Corda 1994: 44; Da Milano 2007: §3.6; Putzu 2015: 48).

Formally, most Italo-Romance type T2 demonstrative systems display a paradigmatic distinction, though less frequently in the distal term, between adnominal and pronominal demonstratives through the use of simple and ECCU-reinforced forms, respectively. In some varieties the distinction is systematic, for example western Abruzzese/Molisan *štu/ssu/quillu libbre* 'this/that/that book' vs *quiste*/*quisse*/*quille* 'this/that/that one' (Finamore 1893: 22; Marinucci 1988: 647), while in others the reinforced forms can also be used in adnominal functions, for

example Teramano *(cu)štu/(que)ssú/(que)llu* vs *cuštə/quessə/quellə* 'this/that/that (one)' (Savini 1881: 62; Mantenuto 2016).

Outside Italo-Romance, however, the distribution of simple and reinforced forms is not correlated with the adnominal/pronominal opposition, but tends to involve diachronic and diatopic variation (Sornicola 2011: §§2.1.1–4). For instance, in the history of Catalan simple (*est, eix*) and reinforced (*aquest, aqueix*) forms alternated up until the Middle Ages (Badia i Margarit 1991: 141; Duarte i Montserrat & Alsina i Keith 1986: 79f; Moll 2006: 179), but are today distributed according to areal tendencies, with the simple forms preferred in north-western dialects and Valencian.

Typologically noteworthy within Romance is the emphatic pattern of demonstrative doubling found in Abruzzese (Savini 1881: 62; Finamore 1893: 22; Rohlfs 1968: 209; Verratti 1968: 48f) where the NP is sandwiched between a non-reinforced demonstrative to its left and a corresponding reinforced form to its right:

(12) Eastern Abruzzese (Verratti 1968)

a. štu=cavalla quéšte
this=horse this
'This horse'

b. ssu=cane quésse
that=dog that
'That dog (near you)'

c. chelu=vóve quélle
that=ox that
'That ox'

3.2.1 Type $T2_A$ systems

Within type T2 systems, we must also recognize at least two formal subtypes, henceforth types $T2_A$ and $T2_B$, in which the deictic space continues to display a strict ternary organization, but the markers of each of the three deictic divisions belong to a distinct system of formal exponence.

Type $T2_A$ demonstrative systems are reported to occur widely in Piedmont and Liguria. For example, Parry (1997: 241) notes that most Piedmontese dialects present as many as three demonstratives continuing reflexes of (ECCU-)ISTE, IPSE and ECCU-ILLE. Fundamentally, the system of most dialects operates in terms of a simple type $B1_B$ opposition (cf. §2.1), namely *cust/stu* 'this' vs *cul* 'that'. However, this basic binary system can be expanded into a strict ternary system through its

combination with one of the three spatio-personal adverbs *sì* 'here', *lì* 'there' (addressee-oriented), and *là* 'there' (cf. Lombardi Vallauri 1995: 219): *cust sì* 'this' [+1person], *cul lì* 'that' [+2person], *cul là* 'that' [−1/−2person]. As for the third term *(ë)s(ë)* (< IPSE; cf. Ascoli 1901), Parry describes it as spatially unmarked, coming close in some respects to the functions of a definite article (cf. Lombardi Vallauri 1995: 214). Indeed, the weakened deictic force of *(ë)s(ë)* is reflected by its frequent use in conjunction with the three spatio-personal adverbs above to produce an alternative ternary adnominal demonstrative system, viz. *(ë)s(ë) sì/lì/là* (cf. discussion of type $B1_C$ systems in §2.2).

This latter formal development is widely found in dialects on the Piedmontese-Ligurian border (Forner 1997: 251; Irsara 2009: 98f). For instance, Parry (1991; 2005: 150–153) reports for Cairese the presence of a single demonstrative, namely IPSE > *es*, with reflexes of ISTE today limited to a handful of lexicalized temporal expressions (e.g. *sc-tamatin* 'this morning') and reflexes of ECCU-ILLE employed solely as adjectival/pronominal cataphors (e.g. *chi u l'è cul óm ch'u vénn?* 'who's the/that man who is coming?'). Just like *(ë)s(ë)* above, Cairese *es* is spatially unmarked, freely referring to the deictic space of any of the three grammatical persons (cf. 13a–c; see also discussion of modern French *ce* in §2.2).

(13) Cairese (Parry 1991; 2005)

a. sa sc-pala a= 'm= fa mò
this shoulder SBJ.CL.3= me= does bad
'I've got this painful shoulder.'

b. do=me sa bursa
give.IMP.2SG=me this bag
'Give me that bag (of yours)!'

c. cum i='s=ciamu sci brichi?
how them=self=call these mountains
'What's the name of those mountains?'

In its pronominal uses, and also very frequently in its adnominal functions, however, *es* is combined with one of the three spatio-personal adverbs *chì* 'here', *lì* 'there' (addressee-oriented), and *là* 'there' yielding once again an analytic ternary system: *es chì/lì/là* 'this one/that one (addressee-oriented)/that one'.

Identical T_{2A} systems are found in many (neighbouring) Occitan dialects (Collègi d'Occitania 2010: 21) which, alongside a simple type B_{1B} opposition *aqueste* 'this' [+1person] vs *aquel* [−1person], may optionally operate a ternary system through the undifferentiated use of *aquel* in conjunction with *d'aicí* 'here', *d'aquí* 'there' (addressee-oriented), and *d'alai* 'there'.

3.2.2 Type $T2_B$ systems

The second formal variant of the type T2 system is found in various parts of Salento, Gascony and south-western Romania (Oltenia) and involves a remarkable functional reanalysis of the dual formal outcomes of the reflex of AQUELLO (Mancarella 1998: 159f; Sornicola 2011: §2.2.1.1). In the Salentino dialects affected, the original long lateral of ECCU-ILLE is subject to various changes, including both a more conservative plosive stage [-ll-] > [-dd-] / > [-ḍḍ-] (e.g. *kwiddu/kwiddə, kuddu/kuddə, kwíḍḍu*) and a more advanced rhotic stage [-ll-] (> [-dd-] > [-ḍḍ-]) > [-r] (e.g. *kwiru/kwirə, kuru/kurə*). Although originally the plosive and rhotic outcomes in reflexes of ECCU-ILLE were presumably variant realizations of the long lateral (cf. dialect of Andrano described by Mancarella 1998: 157), in the relevant dialects the two outcomes have today specialized as distinct formal markers, with the plosive and rhotic outcomes coming to mark the deictic spheres of the addressee and non-discourse participants, respectively.

A not too dissimilar development characterizes many Gascon dialects where, alongside reflexes of *akkʊ̆-ISTE > *aquest(e)* 'this', reinforced reflexes of ILLE combine both with ECCU (> *akkʊ̆) and ECCE (> *akke) to produce velar and palatal outcomes, respectively aligned with the second and third persons (Rohlfs 1970: 188; Sornicola 2011: §2.2.1.1), namely (M/F) *aquéste/aquésto* vs *aquét(ch)/aquéro* vs *acét(ch)/acéro* (cf. 14a). Gascon too frequently employs spatio-personal adverbs in conjunction with the pronominal series (cf. 14b; Daugé 2000: 34). Exceptionally, in Aranés the roles of the palatal and velar variants are reversed, with the former (*acetch*) referencing the addressee and the latter (*aquet*) the non-discourse participants (Rohlfs 1970: 188, n. 323).

(14) a. Armagnac (Rohlfs 1970)
aquést' / aquét / acét òmi
this that that man
'This/That (by you)/that man'

b. Aire-sur-l'Adour, Landes (Daugé 2000)
aquesteací, aqueth aquí, aceth aciu
this here that there that over.there
'This one, that one (by you), that one over there'

Finally, some Oltenian varieties of Daco-Romanian contrast *ăsta*, *ala*, *ăla* (Ionaşcu 1960). Once again, although it is a ternary system which continues Latin terms, namely ISTE > *ăsta* 'this' and two reflexes of ILLE > *ala* 'this/that (addressee-oriented)' and *ăla* 'that (over there)', it does not continue the Latin ternary system, and may in fact, according to Ionaşcu, be a calque on Slavonic.

Among type $T2_B$ dialects we can formally distinguish between type $T2_{B1}$ and type $T2_{B2}$ systems which contrast AQUESTO and AQUESSO, respectively, with the dual outcomes of AQUELLO: (i) type $T2_{B1}$, e.g. province of Lecce *kwíštu* vs *kwíddu* vs *kiru* (Miggiano, Surano, Presicce, Montesano); Gascon dialects, e.g. Béarnais *aqueste/aquesta* vs *aqueth/aquera* vs *aceth/acera* (Rohlfs 1970: 188); and Oltenian dialects, e.g. *ăsta*, *ala*, *ăla*; (ii) type $T2_{B2}$, e.g. province of Brindisi *kussə* vs *kuddə* vs *kurə* (Ostuni, Villa Castelli) and province of Taranto (Ginosa, Martina Franca, Laterza, Palagianello). Both $T2_{B1}$ and $T2_{B2}$ variants of this system would appear then to represent developments from earlier $B2_A$ and $B2_B$ systems (§§4.1–4.2) in which formal marking of the addressee role has been reintroduced into the system through the exaptive reanalysis of erstwhile free phonetic variants of the distal term. This development can apparently be observed in progress in the northern Salentino dialect of Mottola for which Mancarella (1998: 157, 160) reports a four-way system, namely *kustə* vs *kussə* vs *kuddə* vs *kurə*, characterizing the distribution of *kustə* as sporadic. Consequently, speaker-oriented deixis in this dialect now shows advanced on-going competition between AQUESTO and AQUESSO to the advantage of the latter, the predominant outcome in this area (Mancarella 1998: 157), such that the specialization of AQUESSO in this role left a potential gap in the system. In response to this development, the plosive variant (*kuddə*) of the distal term has been pressed into service and deployed to mark addressee-oriented deixis, perhaps still alongside residual uses of *kussə*.

4 Type B2 systems

4.1 Type $B2_A$ systems

I noted in §3 how in a number of central-southern Italian type T1 systems AQUESSO is not integrated into the core demonstrative system, but is largely restricted to the periphery of speakers' grammars as a marked term. In particular, reference to the deictic domain of the addressee is in most cases already marked by AQUESTO in its inclusive functions, so that the role of AQUESSO proves in any case largely redundant. In view of its marginal status, it is not therefore surprising to observe that AQUESSO may frequently fall entirely from usage leaving a new binary system, type $B2_A$, in which reference to the shared deictic domain of both discourse participants in the conversational dyad continues to be marked by the inclusive term AQUESTO, with AQUELLO marking all referents falling outside this domain. This is the situation reported for some varieties of modern Sardinian (Blasco Ferrer 1988: 839), Judaeo-Spanish, and modern Catalan (cf. Badia i Margarit 1951: 281; Badia i Margarit 1995: 501; Duarte i Montserrat & Alsina

i Keith 1986: 81; Hualde 1992: 120f; Wheeler et al. 1999: 106; Da Milano 2007: §3.3; Nogué-Serrano 2015: 208f) where, following the loss of *cussu/ese/aqueix*, the deictic sphere of both discourse participants is now marked by *custu/este/aquest*, contrasting with *cuddu/akel/aquell* which marks referents that fall outside the conversational dyad (cf. 15a,b).

(15) Catalan (Wheeler et al. 1999)

a. Aquest abric que porto / portes
this overcoat that wear.1SG wear.2SG
'This overcoat that I am/you are wearing.'

b. Aquell abric que porta
that overcoat that wear.3SG
'That overcoat which s/he's wearing.'

An identical system is documented and analysed in detail in Ledgeway (2004: 96–104) for modern Neapolitan (cf. also Ledgeway 2009: 195–212) and, more briefly, for some other southern dialects where there obtains a binary opposition *chisto* [−3person] vs *chillo* [+3person]. Thus despite their formal similarity with the Italian dyad *questo* vs *quello*, the modern Neapolitan pair entail a quite different reading, since the Italian opposition makes reference only to the speaker, drawing a contrast between *questo* [+1person] and *quello* [−1person] (Maiden 1995: 125; Vanelli 1995: 324; Maiden & Robustelli 2000: 82f).

Revealing in respect to the diachronic development sketched above are some dialects from the province of Reggio Calabria which typically display a type T2 system, but which in more recent times are reported (Loporcaro 2009: 129) to have all but lost the original addressee-oriented term *ssu*, namely *stu*/$^{(\dagger)}$*ssu*/*ḍḍu mulu* 'this/this/that mule', playing out changes which have long been completed in other varieties. Analogously, in the dialect of Anzi the original addressee-oriented term *kwéssə* is today nothing more than an occasional relic of a former type T1 system with the deictic domain of the addressee all but systematically marked, together with that of the speaker, by the inclusive term *kwéstə* (Ruggieri & Batinti 1992: 50), exemplifying the final stages of a transitional phase from a type T1 to a type $B2_A$ system. In addition to these varieties, type $B2_A$ systems are reported to occur in: (i) most of northern Lazio (Stavinschi 2009: 140); (ii) large areas of Campania (Parascandola 1976: 74; Castagna 1982: 79, 81f); (iii) most dialects south of Taranto-Brindisi (Mancarella 1975: 16, 36; Mancarella 1998: 159; Loporcaro 2009: 129f); (iv) small parts of Calabria (Tassone 2000: 33); and (v) much of Sicily (Varvaro 1988: 722; Ledgeway 2004: 92).

Quite exceptional among the northern Italian dialects, which as we have seen in §§2.1–2.2 predominantly operate a binary [±1person] opposition in which reference to the addressee is neutralized and freely marked by either of the two available terms, is the Romagnol dialect. According to Masotti (1999: 64f), here *stè/quèst* 'this' and *chè/quèl* 'that' are organized in terms of a type $B2_A$ system with the latter indicating "distance from both the speaker and the addressee":

(16) Romagnol (Masotti 1999)

a. [−3pers.]
quest l'=è mi zej; i vòstar dirèt j' è quist
this SBJ.CL.3SG=is my uncle the your rights SBJ.CL.3PL is these
'This is my uncle; your rights are these.'

b. [+3pers.]
quell l' è mi nòn
that SBJ.CL.3SG is my grandfather
'That is my grandfather.'

As with the other southern Italian dialects, pronominal forms in type $B2_A$ systems are typically reinforced by ECCU, whereas in their adnominal functions the demonstratives typically favour unsupported ESTO and, especially in the extreme south (e.g. central-southern Salento, Sicilian), ELLO (Parascandola 1976: 74; Mancarella 1998: 156, 158f; Abbate 1995: 69). In some Salentino varieties where the reinforced forms are also employed with adnominal functions, the paradigmatic distinction between the pronominal/adnominal series continues to be marked by the realization of the post-verbal labial as a glide or in nuclear position (Mancarella 1998: 158):

(17) Cellinese (Mancarella 1998)

a. kwíḍḍu tisse
that.one said
'That one said.'

b. kuḍḍu paíse
that village
'That village'

Locative reinforced forms are also occasionally encountered in type $B2_A$ systems but are typically employed with, though not restricted to, the pronominal demonstratives: Viterbo *quésto qqui(ne)* lit. 'this one here' (Petroselli 2009: 484f),

Neapolitan *chisti ccà* 'these here', *chilli llà* 'those there' (Iandolo 1994: 168; Iandolo 2001: 208, 212). On a par with Emilian-Romagnol varieties characterized by type $B1_B$ systems, Romagnol also displays a reduced copular structure (Masotti 1999: 65): *stucaquè* < *stu ch'è acquè* 'this one that is here', *clucalè* < *clu ch'è lè* 'that one that is there'.

Observe, finally, how the availability of the discontinuous periphrasis AQUESTO (NP) + 'there (near you)' allows type $B2_A$ systems to single out reference to the addressee on those rare occasions when particular emphasis is required and simple AQUESTO is not suitable (Parascandola 1976: 74; Vann 1995: 258; Ledgeway 2004: 102f; Ledgeway 2009: 211; Jungbluth to appear: §5). In particular, despite having entirely lost AQUESSO, the organization of the type $B2_A$ demonstrative system functionally replicates the T1 system through the ternary opposition instantiated by the use of spatio-personal adverbs, e.g., southern Italo-Romance ECCU-HAC (> *(a)ccà*) 'here' [+1/±2person], *'llɔko (> *ll(u)oco, ddh(r)(u)ocu*) 'there' [−1/+2person], and ILLAC (> *llà*, *ddh(r)à*) 'there' [−1/−2person]. For example, in Messinese *chistu (…) ccà* lit. 'this (…) here' constitutes an inclusive expression marking referents "close to both the speaker and the addressee", while *chistu (…) ddhocu* lit. 'this (…) there (near you)' only picks out referents "far from the speaker but close to the addressee", and *chillu ddhà* lit. 'that (over) there' marks referents 'distant from both the speaker and addressee' (Quartarone 1998: 30). Effectively, then, type $B2_A$ dialects like Messinese operate a binary distinction between discourse and non-discourse participants (viz. *chistu (ccà)* vs. *chillu (ddhà)*), with *chistu ddhocu* representing a marked expression of addressee-oriented deixis (cf. also Stavinschi 2009: 76f). It is significant to note that the addressee-oriented spatio-personal adverb *lloco* (and local variants) is only compatible with AQUESTO, and not AQUELLO, an observation entirely in line with my claim that AQUESTO alone may (inclusively) mark the deictic sphere of the addressee.

4.2 Type $B2_B$ systems

In type T_1 systems such as Old Neapolitan there is considerable overlap in the use of the first two terms as a result of their inclusive values,[5] which we have just

[5]As for the inclusive value of AQUESSO, one could assume that it acquired this value by analogy with AQUESTO, with which it enjoyed, as we have seen, a certain degree of distributional overlap. But in any case the inclusive value of AQUESSO was probably already present in the deictic ECCU-IPSU > AQUESSO from the beginning, in that the presentative ECCU (and variants: ECCE, *akke, *akkŭ), besides calling attention to the addressee, also serves to identify a referent in relation to the speaker, as noted by Anderson & Keenan (1985: 279); for further detailed discussion, see Ledgeway (2004: 78–87).

seen in the case of modern Neapolitan and other varieties to have led to the generalization of AQUESTO at the expense of the marked and more restricted member of the system AQUESSO (⇒ type $B2_A$ system). Equally, however, the overlap in the use of AQUESTO and AQUESSO, which guarantees their frequent near equivalence, might just as easily have given rise to an increased use of AQUESSO at the expense of AQUESTO, a state of affairs which could ultimately, though not necessarily, lead to the total loss of AQUESTO. This in fact must be what happened in a large number of southern dialects, including many northern Calabrian (Rohlfs 1977: 167; Ledgeway 2004: 104–107) and most Pugliese dialects (Rohlfs 1968: 207; Valente & Mancarella 1975: 27; Loporcaro 1988: 248; Loporcaro 1997: 344; Loporcaro 2009: 129f; Ledgeway 2004: 107f), which now present a type $B2_B$ system opposing AQUESSO [−3person] vs AQUELLO [+3person], witness (18) below:

(18) Cosentino (personal knowledge)
Ssu vrazzu mi= / ti=fa male / Chiru vrazzu cci=fa
this arm me= you.SG=makes bad that arm him=makes
male
bad
'This arm (of mine/of yours) hurts.' vs. 'That arm (of his) hurts.'

Other Italo-Romance varieties reported to display a type $B2_B$ system include: (i) dialects around Spoleto where *tistu/testo* is reported to include reference to the speaker (Moretti 1987: 98; Stavinschi 2009: 171); (ii) the central Laziale dialect of Palombara (Stavinschi 2009: 140); and (iii) several dialects of northern Salento (Mancarella 1998: 157, 159).

Outside Italo-Romance, type $B2_B$ systems are found in south-eastern Catalan dialects in and around Tarragona (Badia i Margarit 1991: 141; Badia i Margarit 1995: 501), some Latin-American varieties of Spanish (Kany 1945: 170; Zamora Vicente 1967: 434; Stavinschi 2009: 42, 44), and Brazilian Portuguese (Câmara 1971; Teyssier 1976: 114f; Jungbluth 2000; Jungbluth to appear: §5; Jungbluth & Vallentin 2015: 317–319). Although the basic Brazilian Portuguese system is of type $B2_B$ in which *esse* marks the shared deictic sphere of both discourse participants, the so-called *inside* space of the conversational dyad, Jungbluth (2000) has shown that, when necessary, the deictic spheres of the speaker and addressee can still be formally marked off through the use of the postnominal speaker- and addressee-oriented spatio-personal adverbs *aqui* and *aí*, respectively (cf. Carvalho 1976: 27–51; Jungbluth & Vallentin 2015: 317), effectively restoring a type T1 system *esse (aqui)* vs. *esse aí* vs *aquele (lá)*.

It is also possible to identify transitional type $B2_B$ varieties including, for instance, the northern Pugliese variety described by Imperio (1990: 201) which,

although canonically contrasting *cussə* 'this/that' (speaker-/addressee-oriented) with *cuddə* 'that', is reported as still displaying occasional residual uses of *custə* 'this'. Also revealing in this respect is the description of the northern Salentino dialect of Crispiano in Mancarella (1998: 155) where, alongside the standard formal opposition *kussə* [−3person] vs *kuddə* [+3person], *kuštə* is also reported to occur sporadically in place of *kussə* as part of the final stage in the transition from a type T1/2 to a type $B2_B$ system. A similar picture is reported for several northern-western and eastern Catalan dialects (cf. Duarte i Montserrat & Alsina i Keith 1986: 81; Veny 1991: 250) where, following the loss of the original type T1 system, non-discourse participant deixis is invariably marked by *aquell*, but the shared deictic domain of both discourse participants is variously marked, without any distinction of meaning, either by *aquest* (type $B2_A$) or *aquei(x)* (type $B2_B$).

Significantly, the loss of AQUESTO from the demonstrative system of type $B2_B$ varieties faithfully reproduces what must have happened in late Latin following the loss of HIC hypothesized above in §2.1. In this respect, these varieties serve as important models in verifying the reconstruction of the developments in the demonstrative system proposed for late Latin. Above I claimed that with the loss of HIC, the deictic territory it covered and therefore the deictic centre, were inherited by ISTE, whose domain of deictic reference was extended to include the role of the speaker in addition to that of the addressee. This development is accurately reflected in type $B2_B$ dialects where AQUESSO, having replaced AQUESTO, now functions as the term marking referents in the deictic domains of both discourse participants, whereas AQUELLO, in contrast to its reflexes in type $B1_B$ systems (cf. Italian *quello*), picks out referents that fall outside the deictic domain of both discourse participants. Thus, although differing formally from one another with respect to the choice of term employed to mark both discourse participants (AQUESSO vs AQUESTO), functionally type $B2_B$ demonstrative systems are identical to type $B2_A$ systems.

4.3 Type $B2_C$ systems

A number of southern Italian dialects present an interesting development of the type B2 demonstrative system which marries together formal developments of type $B2_A$ and $B2_B$ systems. For instance, several northern Salentino varieties operate a binary opposition in which the distal [+3person] term is standardly represented by AQUELLO, but the deictic space associated with the discourse participants is marked in part by AQUESTO and in part by AQUESSO (Mancarella 1998:

157). For instance, in Castellaneta the pronominal form associated with the discourse participants is AQUESSO (viz. *kussə*), occasionally also found in adnominal functions (e.g., *kussə vagnonə* 'this/that boy'), whereas the usual adnominal form is represented by non-reinforced ESTO (e.g. *štu libbrə* 'this book'). A similar (partially) suppletive paradigmatic distinction is also reported for Massafra and Ginosa, e.g. *kussə (figghiə)* 'this one (son)' vs *štu fratə tuə* 'this brother of yours', as well as for the Pugliese dialect of Mola (Cox Mildare 2001: 62f) where, alongside the core adnominal/pronominal opposition *kɔss* 'this' vs *kəd* 'that', we also find a restricted use of ESTO (viz. *stu*) in adnominal functions alone.

More robust suppletive paradigmatic oppositions of this kind are found in Calabria. For example, Ledgeway (2004: 107) observes that, alongside the traditional Cosentino type $B2_B$ system (*(chi)ssu* vs *chiru*), younger speakers, under the influence of regional Italian, have innovated a compromise suppletive system which for the first term makes recourse to ESTO in adnominal functions (*stu cane* 'this dog'), but which draws on the conservative AQUESSO forms for pronominal uses (*chissu* 'this one'), yielding a mixed system *stu/chissu* vs *chiru*.

4.4 Type B3 systems

Finally I consider one additional binary system, henceforth B3. This system proves relatively rare in Romance and is limited to a number of Latin-American Spanish varieties, e.g. Chile, Venezuela, Ecuador and Cuba (Zamora Vicente 1967: 434; de Bruyne & Pountain 1995: 171). Already we have seen in §4.1 how, from an original T1 system in which AQUESSO was not integrated into the core system, a number of Romance varieties have developed a $B2_B$ type demonstrative system in which the latter term has now fallen from usage such that reference to the deictic sphere of both discourse participants is now marked compositionally by AQUESTO. In the relevant Latin-American Spanish varieties a similar development from an original T1 system has occurred, but with the difference that reference to the deictic sphere of the addressee, previously marked by *ese*, has not been usurped by the erstwhile speaker-oriented term *este* but, rather, by the original non-discourse participant term *aquel*. The result then is a novel binary system in which AQUESTO (viz. *este*) is limited to marking referents that fall exclusively within the deictic sphere of the speaker, whereas AQUELLO functions as an inclusive category marking both addressee and non-discourse participants. Consequently, in these Latin-American varieties *este* is marked [+1person] excluding reference to the addressee, whereas *aquel* is marked [−1person] thereby including reference also to the deictic sphere of the addressee.

5 Type U(nary) systems

I noted above the existence of what are effectively one-term demonstrative systems, typified by French (§2.2), where the single form *ce* (F *cette*, PL *ces*) functions as a demonstrative without specification of place or person; it can be combined with a postnominal locative, but can also occur independently, without a locative element. Cairese (§3.2.1) behaves similarly, as do the other Piedmontese, Ligurian, Francoprovençal and langue d'Oïl varieties reviewed in §2.2. The fact that in these varieties there is only a single demonstrative, which is often not combined with a postnominal locative, implies that the systems in question are best analysed as underlyingly U(nary), with the addition of the locative element yielding derived B(inary) or T(ernary) systems.

6 Rethinking demonstratives

6.1 Summary of findings

In Table 21.2 (page 474) I summarize the various formal and functional characteristics of the thirteen demonstrative systems reviewed above.

6.2 Romance demonstrative systems: A parametric hierarchy approach

Since the conception in early government and binding theory of Universal Grammar in terms of a small set of abstract parameterized options, much work over recent decades has radically departed from this view with a focus on predominantly surface-oriented variation (Borer 1984). This has led to the proliferation of a remarkable number of local, low-level parameters interpreted as the (PF-)lexicalization of specific formal feature values of individual functional heads in accordance with the so-called Borer–Chomsky conjecture (Baker 2008a: 353). While this approach may prove descriptively adequate in that it predicts what precisely may vary (cf. Kayne 2000; 2005a,b; Manzini & Savoia 2005), it suffers considerably from explanatory inadequacy. Among other things, it necessarily assumes such microparameters to be highly local and independent of one another. This assumption seriously increments the acquisitional task of the child who has to set each value in isolation of the next on the basis of the primary linguistic data alone, and at the same time exponentially multiplies the number of parametric systems and, in turn, the number of possible grammars predicted by UG (cf. Kayne 2005b: 11–15; Roberts 2014).

Table 21.2: Overview of Romance demonstrative systems

	Discourse participants		Non-discourse participants	Geographic distribution
Type	Speaker	Addressee		
$B1_A$	(AQU)ESTO		AQUELLO	OFr., RaeR. (Srs., Vld.)
$B1_B$	(AQU)ESTO		AQUELLO	NWGaR., Occ., Gsc., Lad., NIDs, It., Dal., DRo.
$B1_C$	(AQU)ESTO / AQUELLO QUA/QUI		(AQU)ESTO / AQUELLO LÀ/LÌ	Lig., Pie., Frl., (Mil.), Frp., ((O)Fr.)
T1	(AQU)ESTO	((AQU)ESSO / (CO)TESTO)	(AQU)ELLO	CIDS, Tsc., Umb., SIDS, EuSp., EuPt., Ast., Occ.,
T2	(AQU)ESTO	(AQU)ESSO	(AQU)ELLO	CIDs, SIDS (esp. Abr., Mol.), OCat., E/SCat., Vlc., Gsc., Srd.
$T2_A$	ESSO QUI	ESSO LÌ	ESSO LÀ	Lig., Pie., Occ.
$T2_{B1}$	(AQU)ESTO	AQUELLO (*-dd-*,*-qu-*)	AQUELLO (*-r-*, *-c-*)	Sal.; Gsc.; Oltenian Ro.
$T2_{B2}$	(AQU)ESSO	AQUELLO (*-dd-*)	AQUELLO (*-r-*)	NSal.
$B2_A$	(AQU)ESTO		(AQU)ELLO	Srd., JuSp., Cat., Rmg., NLaz., Cmp., Sal., Sic.
$B2_B$	(AQU)ESSO		(AQU)ELLO	NCal., NPug., SUmb., CLaz., NSal., SECat., LAmSp., BrPt., Alg.
$B2_C$	AQUESSO/ESTO		(AQU)ELLO	Cal., Pug., NSal.
B3	ESTO	AQUELLO		LAmSp.
U	←——————————CE—————————→			Fr., Pie-Lig.

One way to avoid the proliferation of grammatical systems that such a microparametric approach predicts is to assume a theory that combines some notion of macroparameters alongside microparameters (Baker 1996; 2008a,b). Following ideas first proposed by Kayne (2005b: 10) and further developed by Roberts & Holmberg (2010) and Roberts (2012), considerable progress in this direction has recently been made by the ReCoS research group; their central idea is that macroparameters should be construed as the surface effect of aggregates of microparameters acting in unison, ultimately as some sort of composite single parameter (cf. Biberauer & Roberts 2017). On this view, macroparametric effects obtain whenever all individual functional heads behave in concert, namely are set identically for the same feature value, whereas microparametric variation arises when different subsets of functional heads present distinct featural specifications.

Conceived in this way, parametric variation can be interpreted in a scalar fashion and modelled in terms of parametric hierarchies. Macroparameters, the simplest and least marked options that uniformly apply to all functional heads, are placed at the very top of the hierarchy, but, as we move downwards, variation becomes progressively less "macro" and, at the same time, more restricted with choices becoming progressively limited to smaller and smaller proper subsets of features, namely, no F(p) > all F(p) > some F(p), for F a feature and p some grammatical behaviour. More specifically, functional heads increasingly display a disparate behaviour in relation to particular feature values which may, for example, characterize: (1) a naturally definable class of functional heads (e.g. [+N], [+finite]), a case of mesoparametric variation; (2) a small, lexically definable subclass of functional heads (e.g. pronominals, auxiliaries), a case of microparametric variation proper; and (3) one or more individual lexical items, a case of nanoparametric variation.

These assumptions then open the way for us to reinterpret the forms and functions of Romance demonstrative systems in terms of a set of hierarchically-organized interrelated parametric options based on differing person feature specifications. In particular, I adopt here the feature geometric analysis of person and number developed by Harley & Ritter (2002), represented schematically in Figure 21.2, which makes specific predictions about the range and types of person combinations, and hence by implication also the types and natural classes of demonstrative systems, that are cross-linguistically available.

For my purposes I focus here on person, namely the PARTICIPANT node and its possible dependents, from which we can derive the four person specifications in Figure 21.3 where projection of the PART(icipant) node indicates the presence of

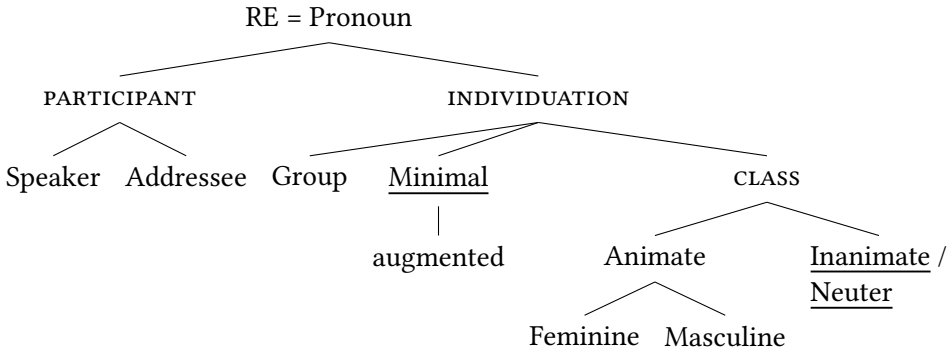


Figure 21.2: Feature geometric analysis of person and number (Harley & Ritter 2002)

person (first and second persons), whereas its absence indicates the lack of person which, following the seminal intuition in Benveniste (1956), corresponds to the so-called third person, the non-person (cf. Harley & Ritter 2002: 488). When projected, in the unmarked case the underspecified value (indicated by underlying) is Sp(eaker) expressing the default first person value as indicated in (a). On the other hand, second person forms are represented by projection of the dependent Ad(dressee) node without the Sp node, as illustrated in (b). When, however, the node for the default Sp value is explicitly filled in without specification of the Ad node (cf. c), we then derive a contrastive first person reading, albeit a marked exclusive interpretation. Finally, the most marked option obtains whenever the PART node is maximally specified as in (d), projecting both Sp and Ad nodes to license an inclusive first person interpretation uniting the deictic spheres connected to the speaker and addressee features.

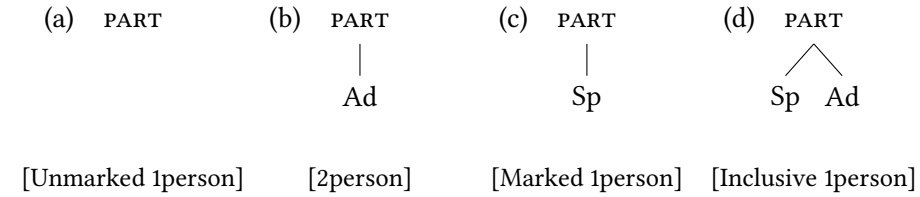


Figure 21.3: Possible person specifications

With these fundamental person specifications in place, I now turn to consider the formal representation of Romance demonstrative systems sketched in the parameter hierarchy in Figure 21.4.

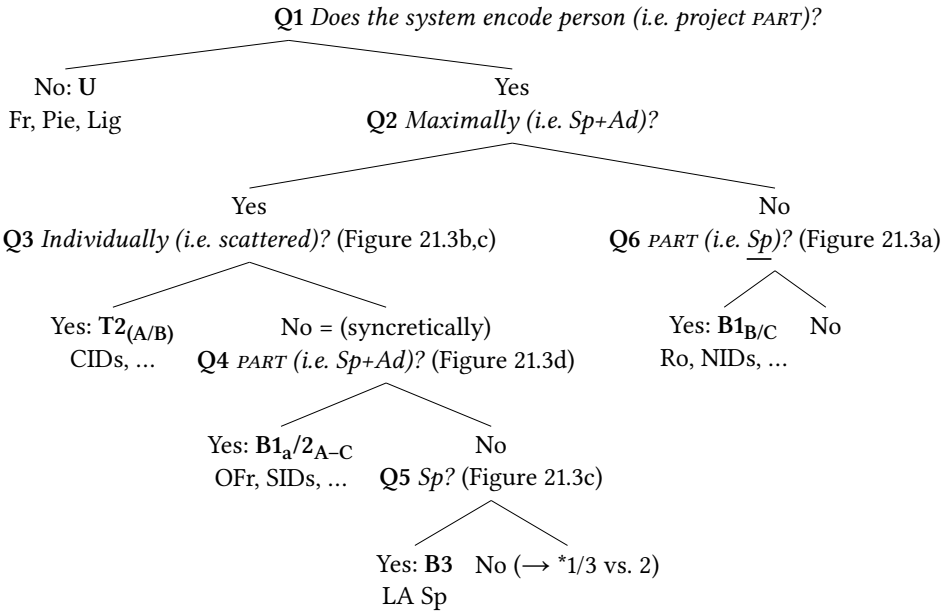


Figure 21.4: Parametric hierarchy for Romance demonstrative systems

In line with our markedness expectations (no F(p) > all F(p) > some F(p)), the first question in Figure 21.4 simply asks whether a given demonstrative system encodes person, albeit projects the PART node. The least marked option is represented by varieties such as modern French and many Piedmontese and Ligurian varieties (cf. §5) whose demonstrative systems I have characterized as unary, in that they fail to encode any person distinctions (cf. languages lacking pronouns such as Japanese; Harley & Ritter 2002: 512). However, as we have seen, most Romance varieties do in fact encode person, such that the next question (viz. Q2) in Figure 21.4 asks whether person is maximally encoded such that all possible person features (viz. Sp and Ad) are grammaticalized within the system. If the answer to this question is positive, then this immediately triggers the follow-up question whether the maximal representation of person features within the system is realized in a scattered fashion (Q3). In the case of a positive answer to this question, we correctly identify $T2_{(A/B)}$ systems (cf. §§3.2.1–3.2.2) including, among others, many central Italian dialects which reserve a distinct term for each of the three person specifications variously projecting fully specified PART nodes (cf. options b,c in Figure 21.3) or no PART node at all in the case of the so-called third person. If, however, the answer to Q3 is negative, this necessarily

implies that the maximal representation of person features must be realized syncretically, giving rise to inclusive forms which are typologically rarer (Harley & Ritter 2002: 496) and hence more marked, as reflected by their concomitant placement towards the end of the hierarchy in Figure 21.4.

Here there arise two possibilities. The first and least marked, as formalized in Q4, is to ask whether the syncretic realization of maximal person features involves projection of the PART node, giving rise to the Sp and Ad inclusive forms (cf. option d in Figure 21.3) found in $B1_A/2_{(A\text{-}C)}$ systems which operate a [±discourse participant] opposition through the formal binary distinction between AQUESTO (or AQUESSO) and AQUELLO. The second and more marked option is formalized through Q5 which asks whether maximal representation of person features when realized syncretically involves a different type of split which privileges the Sp as an exclusive first person category. This marked option perfectly describes B3 systems which we have seen are quite rare in Romance, only occurring in a limited number of Latin-American Spanish varieties where an exclusively speaker-oriented form *este* contrasts with *aquel* which syncretically marks referents that fall within the deictic sphere of the addressee and non-discourse participants. As predicted by its position towards the bottom of hierarchy in Figure 21.4, this latter possibility admittedly represents a marked option from a cross-linguistic perspective and is even argued by Harley & Ritter to be unattested in their sample of 110 languages. In particular, they maintain:

> "[w]hat we predict NOT to exist are languages that use the same pronoun (or in a language with cases, the same set of pronouns) for both 1st and 3rd or both 2nd and 3rd persons. In fact, none of the languages we looked at has such a pronoun or set of pronouns in its inventory." (Harley & Ritter 2002: 513)

Admittedly, the highly marked option of a single demonstrative term that syncretically marks first and third persons in opposition to a term uniquely restricted to referencing the second person is not attested in my Romance sample, witness the position of this unattested option at the very bottom of the hierarchy in Figure 21.4 which no doubt represents a no choice parameter. However, we have seen that the less marked option of a formal opposition between a marked Sp category and all other persons is not only attested in Romance, but, is also predicted by Harley & Ritter's system which readily allows for a marked first person category (cf. option c in Figure 21.3) that formally excludes reference to the Ad.

Finally, I turn to Q6, a possibility that arises whenever person is not encoded maximally in a given language (cf. Q1). In particular, if person is not encoded

maximally, then in accordance with Harley & Ritter's claims about markedness and person features I ask whether at the very least encoding of person features includes the projection of the PART node, represented in the unmarked case by the underspecified value of Sp instantiating the default first person value (cf. option a in Figure 21.3). In reality, this question involves a no choice parameter, inasmuch as a negative response, which would produce a hypothetical system that only references the deictic sphere of the Ad, is not an option since deictic systems must at the very least make reference to the Sp, the deictic centre to which all deictic relations are anchored. Consequently, the positive answer to Q6 allows us to identify $B1_{B/C}$ demonstrative systems such as Romanian and northern Italian dialects (§§2.1–2.2), where projection of PART yielding the underspecified Sp value does not necessarily exclude the Ad, which we have seen may be encoded by either of the two terms of the system, but correctly places by default the Sp at the centre of the opposition.

7 Concluding remarks

To conclude, I briefly look at a number of other significant implications of the parametric representation in Figure 21.4. First, despite my identification of 13 formal systems in Table 21.2, the hierarchy in Figure 21.4 reduces this superficial variation in demonstrative systems to just five featural parametric options. This is clearly a welcome result since it underlines how cross-linguistic variation should not necessarily be taken at face value as instantiating distinct parametric choices, but can often be reduced to a finite set of natural classes and options.

Second, although I have identified a number of binary formal systems, this does not a priori presuppose a binary featural opposition. Rather, we have seen that, despite operating on the surface in terms of a binary formal opposition, $B1_A/2_{A\text{-}C}$ demonstrative systems nonetheless involve a syncretic ternary featural opposition in that they refer to three person values.

Third, the representation in Figure 21.4 reveals how a formal analysis in terms of unbundled feature specifications such as [±1], [±2], and [±3] proves entirely inadequate at all relevant levels (cf. footnote 4). For example, if we were to characterize $B1_{B/C}$ systems in terms of a simple [±1] feature, then it would incorrectly predict that the first term of the system exclusively marks reference to the speaker, with reference to the addressee marked solely through the second term of the system together with the so-called third person. By contrast, we have observed how in these systems reference to the addressee may ambiguously fall between both terms of the system, a fact which is immediately captured by our

analysis in terms of Sp which, while not formally excluding reference to the Ad, nonetheless places the speaker at the centre of the opposition. In a similar fashion, a simple [±1] feature would equally make incorrect predictions about B3 systems: if in such Latin-American Spanish varieties we were to characterize the superficial binary opposition in terms of a [+1] (= *este*) vs [−1] (= *aquel*) contrast, then we would fail to capture the fact that only the second term also explicitly includes reference to the deictic sphere of the addressee, since under this simple representation reference to the addressee could a priori also be marked by the first term, contrary to fact.

Analogous arguments carry over to $B1_A$ and B2 systems where we might a priori be tempted to analyse the relevant contrasts in terms of a simple [±3] opposition. In principle, it would be possible to analyse the first and second terms of such binary systems in terms of the feature specifications [−3] and [+3], respectively, while still maintaining the correct empirical generalization that the first term of the opposition is an inclusive category marking reference to both discourse participants. However, to do so would force us to lose the significant generalization (cf. Harley & Ritter 2002: 504f) that the relevant inclusive forms are built on the saliency of the Sp (AQUESTO = $B2_A$) or the Ad (AQUESSO = $B2_B$). Equally unsatisfactory would be any attempt to analyse $B1_{B/C}$ systems by way of a simple [±3] opposition, since this would incorrectly entail that in such systems reference to the addressee can only be marked through the first term of the system, but never by the second term of the system (viz. AQUELLO).

Finally, another important consequence of the hierarchical representation in Figure 21.4 is the conclusion that the T1 systems observed above in §3 do not constitute under the analysis developed here independent person systems, but, rather, represent a transitional phase in the passage from an original T2 system to a $B2_A$ system.

Abbreviations

1	first person	PF	phonetic form
2	second person	PL	plural
3	third person	SBJ	subject
CL	clitic	SG	singular
IMP	imperative	UG	Universal Grammar
INF	infinitive		

Acknowledgements

Over many years Ian Roberts has been an important influence on my research, especially, but not only, in relation to his groundbreaking work within Romance and theoretical linguistics. It is fitting therefore that the present article, which I dedicate to my good friend and colleague, should also attempt to show how a number of key theoretical ideas developed in large part by Ian himself can provide original insights into a traditional topic in comparative Romance linguistics.

References

Abbate, Lucia. 1995. *Contributo allo studio del dialetto di Frazzanò*. Messina: Armando Siciliano Editore.

Academia de la Llingua Asturiana. 2001. *Gramática de la llingua asturiana*. Oviedo: Academia Llingua Asturiana.

Alkire, Ti & Carol Rosen. 2010. *Romance languages: A historical introduction*. Cambridge: Cambridge University Press.

Anderson, Stephen R. & Edward L. Keenan. 1985. Deixis. In Timothy Shopen (ed.), *Language typology and syntactic description*, vol. III, 259–308. Cambridge: Cambridge University Press.

Arnaud, François & Gabriel Morin. 1920. *Le langage de la Vallée de Barcelonnette*. Paris: Champion.

Ascoli, Graziadio. 1901. Intorno ai continuatori neolatini del latino "ipsu". *Archivio glottologico italiano* 15. 303–316.

Azaretti, Emilio. 1982. *L'evoluzione dei dialetti liguri*. Sanremo: Casabianca.

Badia i Margarit, Antoni M. 1951. *Gramática histórica catalana*. Barcelona: Noguer.

Badia i Margarit, Antoni M. 1991. Evolución lingüística interna i. gramática. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. V, 2. Okzitanisch, Katalanisch, 127–152. Tübingen: Niemeyer.

Badia i Margarit, Antoni M. 1995. *Gramàtica de la llengua catalana: Descriptiva, normativa, diatòpica, diastràtica*. Barcelona: Biblioteca universitària Edicions Proa.

Baker, Mark C. 1996. *The polysynthesis parameter*. Oxford University Press: Oxford.

Baker, Mark C. 2008a. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of syntactic variation*, 351–373. Amsterdam: John Benjamins. DOI: 10.1075/la.132.16bak.

Baker, Mark C. 2008b. *The syntax of agreement and concord*. Cambridge: Cambridge University Press.

Bartoli, Mario. 1906. *Das Dalmatische: Altromanische Sprachreste von Veglia bis Ragusa und ihre Stellung in der Apennino-balkanische Romània. I, II*. Vienna: Hölder.

Benveniste, Émile. 1946. Structure des relations de personne dans le verbe. *Bulletin de la Société de Linguistique de Paris* 43. 1–12.

Benveniste, Émile. 1956. La nature de pronoms: Essays on the occasion of his sixtieth birthday, 11 October 1956. In Morris Halle (ed.), *For Roman Jakobson*, 34–37. The Hague: Mouton.

Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102(2–3). 87–13. DOI: 10.1016/S0024-3841(96)00046-0.

Berruto, Gaetano. 1974. *Piemonte e Valle d'Aosta* (Centro di studio per la dialettologia italiana 1). Pisa: Pacini.

Biberauer, Theresa & Ian Roberts. 2017. Parameter setting. In Adam Ledgeway & Ian Roberts (eds.), *The Cambridge handbook of historical syntax*, 134–161. Cambridge: Cambridge University Press.

Blasco Ferrer, Eduardo. 1988. Sardisch, interne Sprachgeschichte i. gramamtik. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. Band IV Italienisch, Korsisch, Sardisch, 836–853. Tübingen: Niemeyer.

Borer, Hagit. 1984. *Parametric syntax: Case studies in Semitic and Romance languages*. Dordrecht: Foris.

Brault, Grégoire. 2004. Pour une typologie des emplois de *là-bas*. *Linguisticae Investigationes* 27(1). 25–45.

Brugè, Laura. 1996. Demonstratives in Spanish: A comparative approach. *University of Venice Working Papers in Linguistics* 16. 1–53.

Brugè, Laura. 2002. The position of demonstratives in the extended nominal projection. In Guglielmo Cinque (ed.), *Functional structure in DP and IP* (The cartography of syntactic structures 1), 15–53. Oxford: Oxford University Press.

Brunot, Ferdinand. 1899. *Précis de grammaire historique française*. Quatrième édition. Paris: Masson et Cie.

Butt, John & Carmen Benjamin. 1994. *A new reference grammar of modern Spanish*. 2nd ed. London: Edward Arnold.

Butz, Beat. 1981. *Morphosyntax der Mundart von Vesmes (Val Terbi)*. Bern: Francke.

Câmara, Jnr, Joaquim Mattoso. 1971. Uma evolução em marcha: A relação entre êste e êsse. In Eugenio Coseriu & Wolf-Dieter Stempel (eds.), *Sprache und Geschichte: Festschrift für Harri Meier*, 327–331. München: Fink.

Caragiu Marioţeanu, Matilda. 1989. Rumänisch. Areallinguistik I. Dakorumänisch / Les aires linguistiques I. Dacoroumain. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. III, Die einzelnen romanischen Sprachen und Sprachgebiete von der Renaissance bis zur Gegenwart: Rumänisch, Dalmatisch / Istroromanisch, Friaulisch, Ladinisch, Bündnerromanisch, 405–423. Tübingen: Niemeyer.

Carvalho, José G. Herculano de. 1976. Systems of deictics in Portuguese. In Jürgen Schmidt-Radefeldt (ed.), *Readings in Portuguese linguistics*, 245–266. Amsterdam: John Benjamins.

Castagna, Giovanni. 1982. *Guida grammaticale del dialetto foriano letterario*. Forio d'Ischia: Epomeo.

Cimarra, Luigi. 1998. *Il dialetto di Sant'Oreste*. Sant'Oreste: Apeiron.

CNRTL. 2012. *Centre National de Ressources Textuelles et Lexicales*. http://www.cnrtl.fr.

Collègi d'Occitania. 2010. *Cahier de grammaire: Cours de 1ère année*. Toulouse: Collègi d'Occitania.

Corda, Francesco. 1994. *Grammatica moderna del sardo logudorese con una proposta ortografica, elementi di metrica e un glossario*. Cagliari: Edizioni della Torre.

Cordin, Patrizia. 2016. *On lexical marking Romance dialects: Person and space deixis*. 11th Cambridge Italian Dialect Syntax–Morphology Meeting, University of Vienna, 4-6 July 2016.

Cox Mildare, Terry Brian. 2001. *Aspetti della fonologia e della morfologia del molese, un dialetto pugliese dell'Italia sud-orientale*. Bari: Levante.

Cunha, Celso Ferreira da & Luís F. Lindley Cintra Cintra. 1984. *Nova gramática do português contemporâneo*. Lisboa: Edições J. Sá da Costa.

Da Milano, Federica. 2007. Demonstratives in parallel texts: A case study. *Sprachtypologie und Universalienforschung* 60(2). 135–147.

Daugé, Césaire. 2000. *Grammaire élémentaire de la langue Gasconne (dialecte d'Aire)*. Monein: Princi Réguer.

de Bruyne, Jacques & Christopher J. Pountain. 1995. *A comprehensive Spanish grammar*. Malden, MA: Blackwell Publishers.

Diccionario de la lengua española. 1970. Madrid: Real Academia Española.

Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins. DOI: 10.1075/tsl.42.

Duarte i Montserrat, Carles & Alex Alsina i Keith. 1986. *Gramàtica històrica del català*. Vol. 7, 11 (Manuals curial). Barcelona: Curial.

Eguren, Luis. 1999. Pronombres y adverbios demostrativos: Las relaciones deícticas. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la lengua española*, vol. I, 929–973. Madrid: Espasa.

Eguren, Luis. 2012. Binding: Deixis, anaphors, pronominals. In José Ignacio Hualde, Antxon Olarrea & Erin O'Rourke (eds.), *The handbook of hispanic linguistics*, 557–577. Oxford: Wiley-Blackwell.

Finamore, Gennaro. 1893. *Vocabolario dell'uso abruzzese*. 2. ed. Città di Castello: Tip. dello stabilimento S. Lapi.

Foresti, Fabio. 1988. Aree linguistiche v. emilia e Romagna. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen linguistik*, vol. Band IV. Italienisch, Korsisch, Sardisch, 569–593. Tübingen: Niemeyer.

Forner, Werner. 1997. Liguria. In Martin Maiden & M. Mair Parry (eds.), *The dialects of Italy*, 245–252. London: Routledge.

Frías Conde, Xavier. 1999. *An approach to the Asturian language*. Romania Minor, www.romaniaminor.net.

García de Diego, Vicente. 1946. *Manual de dialectología española*. Madrid: Cultura hispánica.

Gómez Sánchez, Elena & Konstanze Jungbluth. 2015. European Spanish. In Konstanze Jungbluth & Federica Da Milano (eds.), *Manual of deixis in Romance languages*, 240–257. Berlin: De Gruyter. DOI: 10.1515/9783110317732-013.

Grandgent, Charles. 1909. *An outline of the phonology and morphology of old Provençal*. 2nd. Boston: Heath.

Gutiérrez-Rexach, Javier. 2002. Demonstratives in context: Trends in Spanish semantics and pragmatics. In Javier Gutiérrez-Rexach (ed.), *From words to discourse*, 195–238. Amsterdam: Elsevier.

Gutiérrez-Rexach, Javier. 2005. Setting up perspectives for demonstratives in discourse. *Traitement Automatique des Langues* 46. 141–162.

Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3). 482–526. DOI: 10.1353/lan.2002.0158.

Hottenroth, Priska-Monika. 1982. The system of local deixis in Spanish: Crosslinguistic studies on deixis and demonstration. In Jürgen Weissenborn & Wolfgang Klein (eds.), *Here and there*, 133–153. Amsterdam: John Benjamins. DOI: 10.1075/pb.iii.2-3.07hot.

Hualde, José. 1992. *Catalan*. London: Routledge.

Iandolo, Antonio. 2001. *Parlare e scrivere in dialetto napoletano*. Naples: Tempolungo Edizioni.

Iandolo, Carlo. 1994. *'A lengua 'e pulecenella: Grammatica napoletana*. Castellammare di Stabia: Franco Di Mauro.

Imperio, Stefano Leonardo. 1990. *Alle origini del dialetto pugliese*. Fasano di Puglia (Italia): Schena.
Ionaşcu, Al. 1960. O paralelă gramaticală slavo-romînă (sistemul cuvintelor deictice în graiurile olteneşti). *Romanoslavica* 4. 73–86.
Irsara, Martina. 2009. *Demonstratives and adverbs of place in early and modern texts of northern Italy*. Padua: Monografie LabRoman 3/II-2009.
Jones, Michael Allan. 1993. *Sardinian syntax*. London: Routledge.
Jungbluth, Konstanze. 2000. Os pronomes demonstrativos do português brasileiro na fala e na escrita. *Caderno de linguagem e sociedade* 7. 83–105.
Jungbluth, Konstanze. 2003. Deictics in the conversational dyad: Findings in Spanish and some cross-linguistic outlines. In Friedrich Lenz (ed.), *Deictic conceptualisation of space, time and person*, 13–40. Amsterdam: John Benjamins.
Jungbluth, Konstanze. To appear. Positions – constellations – practices: Referring to the space in front of the hearer. In Federica Da Milano & Paolo Zublena (eds.), *Space and language: On deixis*. Amsterdam: John Benjamins.
Jungbluth, Konstanze & Rita Vallentin. 2015. Brazilian Portuguese. In Konstanze Jungbluth & Federica Da Milano (eds.), *Manual of deixis in Romance languages*, 315–331. Berlin: De Gruyter.
Kany, Charles E. 1945. *American-Spanish syntax*. Chicago: University of Chicago Press.
Kayne, Richard S. 2000. *Parameters and universals*. Oxford University Press.
Kayne, Richard S. 2005a. On the left edge in UG: A reply to McCloskey. In *Movement and silence*, 50–56. Oxford: Oxford University Press.
Kayne, Richard S. 2005b. Some notes on comparative syntax, with special reference to English and French. In Guglielmo Cinque & Richard S. Kayne (eds.), *Handbook of comparative syntax*, 3–69. Oxford: Oxford University Press.
Kjellman, Hilding. 1928. *Étude sur les termes démonstratifs en provençal*. Göteborg: Elanders boktryckeri aktiebolag.
Koschwitz, Eduard. 1894. *Grammaire historique de la langue des Félibres*. Greifswald: J. Abel.
Langacker, Ronald W. 1990. *Concept, image, and symbol: The cognitive basis of grammar*. Berlin: De Gruyter. DOI: 10.1515/9783110857733.
Lausberg, Heinrich. 1976. *Linguistica romanza*. Vol. II, Morfologia. Milan: Feltrinelli.
Ledgeway, Adam. 2004. Lo sviluppo dei dimostrativi nei dialetti centromeridionali. *Lingua e stile* 39(1). 65–112.
Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*. Vol. Bd. 350 (Beihefte zur Zeitschrift für romanische Philologie). Tübingen: Niemeyer.

Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press.
Ledgeway, Adam. 2015. Varieties in Italy. In Konstanze Jungbluth & Federica Da Milano (eds.), *Manual of deixis in Romance languages*, 146–162. Berlin: De Gruyter.
Ledgeway, Adam & John Charles Smith. 2016. Deixis. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 879–896. Oxford: Oxford University Press.
Leone, Alfonso. 1995. *Profilo di sintassi siciliana*. Vol. 3. Palermo: Centro studi filologici e linguistici siciliani.
Lombardi Vallauri, Edoardo. 1995. Il sistema dei pronomi dimostrativi dal latino al piemontese (varietà torinese), una catena di trazione morfologica: Atti del XXVVII congresso della Società di Linguistica Italiana, Lecce, 28-30 ottobre 1993. In Maria Teresa Romanello & Immacolata Tempesta (eds.), *Dialetti e lingue nazionali*, 209–225. Rome: Bulzoni.
Loporcaro, Michele. 1988. *Grammatica storica del dialetto di Altamura*. Pisa: Giardini editori e stampatori.
Loporcaro, Michele. 1997. Puglia and Salento. In Martin Maiden & M. Mair Parry (eds.), *The dialects of Italy*, 341–351. London: Routledge.
Loporcaro, Michele. 2009. *Profilo linguistico dei dialetti italiani* (Manuali Laterza 275). Rome: Laterza.
Lüdtke, Helmut. 1979. *Lucania* (Consiglio nazionale delle ricerche, Centro di studio per la dialettologia italiana 17). Pisa: Pacini.
Lyons, Christopher. 1999. *Definiteness*. Cambridge: Cambridge University Press.
Maiden, Martin. 1995. *A linguistic history of Italian*. London: Longman.
Maiden, Martin & Cecilia Robustelli. 2000. *A reference grammar of modern Italian*. Chicago, Ill: McGraw-Hill.
Mancarella, Giovan Battista. 1975. *Salento*. Pisa: Pacini.
Mancarella, Giovan Battista. 1998. *Salento: Monografia regionale della Carta dei dialetti italiani*. Lecce: Edizioni del Grifo.
Manea, Dana. 2012. Pronumele: Morfosintaxa limbii literare în secolele al xix-lea şi al xx-lea. In Gheorge Chivu, Gabriela Pană Dindelegan, Adina Dragomirescu, Isabela Nedelcu & Irina Nicula (eds.), *Studii de istorie a limbii române*, 479–524. Bucharest: Editura Academiei Române.
Mantenuto, Isara. 2016. *Demonstratives in Teramano: A decompositional approach*. Poster presented at Going Romance 30, Frankfurt, Germany.
Manzini, M. Rita & Leonardo M. Savoia. 2005. *I dialetti italiani e romanci: Morfosintassi generativa*. 3 vols. Alessandria: Edizioni dell'Orso.

Marcato, Gianna & Flavia Ursini. 1998. *Dialetti veneti: Grammatica e storia*. Padua: Unipress.

Marinucci, Marcello. 1988. Aree linguistiche VIII: Abruzzo e Molise. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. IV. Italienisch, Korsisch, Sardisch, 643–652. Tübingen: Niemeyer.

Masotti, Adelmo. 1999. *Grammatica romagnola*. Ravenna: Edizioni del Girasole.

Meyer-Lübke, Wilhelm. 1895. *Grammaire des langues romanes*. Vol. Tome deuxième: Morphologie (traduction française par auguste Doutrepont et Georges Doutrepont). Paris: Welter.

Meyer-Lübke, Wilhelm. 1900. *Grammaire des langues romanes*. Vol. Tome troisième: Syntaxe (traduction française par auguste Doutrepont et Georges Doutrepont). Paris: Welter.

Moll, Francesc de Borja. 2006. *Gramàtica històrica catalana*. Valencia: Publicacions de Universitat de València.

Moretti, Giovanni. 1987. *Umbria*. Pisa: Pacini.

Nogué-Serrano, Neus. 2015. Catalan. In Konstanze Jungbluth & Federica Da Milano (eds.), *Manual of deixis in Romance languages*, 206–239. Berlin: De Gruyter.

Nyrop, Kristoffer. 1925a. *Grammaire historique de la langue française*. Vol. 5. Syntaxe. Noms et pronoms. Copenhagen: Gyldendal.

Nyrop, Kristoffer. 1925b. *Grammaire historique de la langue française*. Vol. 2. Morphologie. Copenhagen: Gyldendal.

Parascandola, Vittorio. 1976. *Vèfio: Folk-glossario del dialetto procidano*. Napoli: A. Berisio.

Parry, M. Mair. 1991. Le système démonstratif du cairese. In Dieter Kremer (ed.), *Actes du XVIIIe congrès international de linguistique et de philologie romanes*, 625–631. Tübingen: Niemeyer.

Parry, M. Mair. 1997. Piedmont. In Martin Maiden & M. Mair Parry (eds.), *The dialects of Italy*, 237–244. London: Routledge.

Parry, M. Mair. 2005. *Sociolinguistica e grammatica del dialetto di Cairo Montenotte: Parluma 'd còiri* (Zetesis 2). Savona: Società savonese di storia patria.

Penny, Ralph J. 2000. *Variation and change in Spanish*. Cambridge, UK: Cambridge University Press. http://www.loc.gov/catdir/samples/cam031/00023265.html.

Petroselli, Francesco. 2009. *Il lessico dialettale viterbese nelle testimonianze di Emilio Maggini*. Viterbo: Banca di Viterbo.

Pitré, Giuseppe & Christian Wentrup. 1995. *Grammatica siciliana del dialetto e delle parlate*. Palermo: Flaccovio.

Price, Glanville. 1971. *The French language: Present and past*. London: Edward Arnold.

Putzu, Ignazio. 2015. Sardinian. In Konstanze Jungbluth & Federica Da Milano (eds.), *Manual of deixis in Romance languages*, 45–58. Berlin: De Gruyter. DOI: 10.1515/9783110317732-005.

Quartarone, Saverio. 1998. *'A parrata missinisa: Il dialetto siciliano in Messina*. Messina: Antonino Sfameni.

Regnicoli, Agostino. 1995. Questione di organizzazione linguistica dello spazio a partire da fenomeni del dialetto maceratese. In Raffaella Bombi (ed.), *Lingue speciali e interferenza. atti del Convegno seminariale, Udine 16-17 maggio 1994*, 227–247. Rome: Il Calamo.

Roberts, Ian. 2012. Macroparameters and minimalism: A programme for comparative research. In Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo & Juanito Avelar (eds.), *Parameter theory and linguistic change*, 320–335. Oxford: Oxford University Press.

Roberts, Ian. 2014. Subject clitics and macroparameters. In Paola Benincà, Adam Ledgeway & Nigel Vincent (eds.), *Diachrony and dialects: Grammatical change in the dialects of Italy*, 177–201. Oxford: Oxford University Press.

Roberts, Ian & Anders Holmberg. 2010. Introduction: Parameters in minimalist theory. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 1–57. Cambridge: Cambridge University Press.

Roca, Ignacio. 2009. Todas las vascas son vascos, y muchos vascos también vascas: Género y sexo en el castellano. *Boletín de la Real Academia Española* 89. 77–117.

Rohlfs, Gerhard. 1968. *Grammatica storica della lingua italiana e dei suoi dialetti*. Vol. II. Morfologia. Turin: Einaudi.

Rohlfs, Gerhard. 1970. *Le gascon: Études de philologie pyrénéenne*. Second edition (Beihefte zur Zeitschrift für romanische Philologie, 85. Heft). Tübingen: Niemeyer.

Rohlfs, Gerhard. 1977. *Nuovo dizionario dialettale della Calabria (con repertorio italo-calabro)*. Nuova ed. interamente rielaborata, ampliata ed aggiornata. Ravenna: Longo.

Ronjat, Jules. 1913. *Essai de syntaxe des parlers provençaux modernes*. Macon: Protat.

Rowlett, Paul. 2007. *The syntax of French*. Cambridge, UK: Cambridge University Press. http://www.loc.gov/catdir/enhancements/fy0729/2007297864-b.html.

Ruggieri, Donato & Antonio Batinti. 1992. *Lingua e dialetto ad Anzi (Potenza): Note sulla situazione linguistica*. Vol. 3. Potenza: Il Salice.

Salvat, Joseph. 1998. *Gramatica occitana: Quinta edicion revista per Ernèst Nègre*. Toulouse: Collège d'Occitanie.

Salvi, Giampaolo. 2011. Morphosyntactic persistence. In Martin Maiden, John Charles Smith & Adam Ledgeway (eds.), *The Cambridge history of the Romance languages*, vol. 1: Structures, 318–381. Cambridge: Cambridge University Press.

Sandfeld, Kristian & Hedvig Olgen. 2019. *Syntaxe roumaine*. Copenhague: Munksgaard.

Savini, Giuseppe. 1881. *La grammatica ed il lessico del dialetto teramano: Due saggi*. Bologna: Forni.

Smith, John Charles. 1995. L'évolution sémantique et pragmatique des adverbes déictiques ici, là et là-bas. *Langue française* 107. 43–57.

Sornicola, Rosanna. 2011. Per la storia dei dimostrativi romanzi: I tipi neutri [tso], [so], [ço], [to] e la diacronia dei dimostrativi latini. *Zeitschrift für romnische Philologie* 127. 1–80.

Stavinschi, Alexandra. 2009. *Sullo sviluppo del sistema dimostrativo italo-romanzo: Una nuova prospettiva*. Padua: Monografie LabRoman 3/I-2009 (www.maldura.unipd.it/ddlcs/laboratorio/home.html).

Tassone, Bruno. 2000. *Grammatica e dizionario del dialetto crotonese*. Soveria Mannelli (Catanzaro): Calabria letteraria.

Teyssier, Paul. 1976. *Manuel de langue portugaise: Portugal-Brésil*. Paris: Klincksieck.

Teyssier, Paul. 1980. *Histoire de la langue portugaise*. Vol. 1864 (Que sais-je?). Paris: Presses universitaires de France.

Tláskal, Jaromir. 1994. Flexion. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. Band VI, 2. Galegisch, Portugiesisch, 160–172. Tübingen: Niemeyer.

Topa Valentim, Helena. 2015. European Portuguese. In Konstanze Jungbluth & Federica Da Milano (eds.), *Manual of deixis in Romance languages*, 297–314. Berlin: De Gruyter.

Valente, Vincenzo & Giovan Battista Mancarella. 1975. *Puglia*. Vol. 15-16 (Consiglio nazionale delle ricerche, Centro di studio per la dialettologia italiana 5). Pisa: Pacini.

Vanelli, Laura. 1995. La deissi. In Lorenzo Renzi, Giampaolo Salvi & Anna Cardinaletti (eds.), *Grande grammatica italiana di consultazione*, 261–375. Bologna: Il Mulino.

Vanelli, Laura. 1997. Personal pronouns and demonstratives. In Martin Maiden & M. Mair Parry (eds.), *The dialects of Italy*, 106–113. London: Routledge.

Vann, Robert E. 1995. Constructing Catalanism: Motion verbs, demonstratives, and locatives in the Spanish of Barcelona. *Catalan Review* IX(2). 253–274.

Varvaro, Alberto. 1988. Aree linguistiche XII. Sicilia. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. IV. Italienisch, Korsisch, Sardisch, 716–731. Tübingen: Niemeyer.

Veny, Joan. 1991. Català: Áreas lingüísticas. In Günter Holtus, Michael Metzeltin Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. V, 2. Okzitanisch, Katalanisch, 243–261. Tübingen: Niemeyer.

Verratti, Vittore. 1968. *Fonologia e morfologia del volgare abruzzese*. Lanciano: Cooperativa editorale tipografica.

Vignuzzi, Ugo. 1988. Aree linguistiche VII: Marche, Umbria, Lazio. In Günter Holtus, Michael Metzeltin & Christian Schmitt (eds.), *Lexikon der romanistischen Linguistik*, vol. IV. Italienisch, Korsisch, Sardisch, 606–642. Tübingen: Niemeyer.

Vignuzzi, Ugo. 1997. Lazio, Umbria and the Marche. In Martin Maiden & M. Mair Parry (eds.), *The dialects of Italy*, 311–320. London: Routledge.

Vincelli, Antonio. 1995. *Grammatica descrittiva del dialetto di Casacalenda*. 1. ed. Vol. 2. Campobasso: Edizioni Enne.

Vincent, Nigel. 1999. The synchrony and diachrony of deixis in Old Neapolitan (with special reference to the Cronache e ricordi of Loise De Rosa). Ms., University of Manchester.

Wheeler, Max, Alan Yates & Nicolau Dols. 1999. *Catalan: A comprehensive grammar*. London: Routledge.

Zamora Vicente, Alonso. 1967. *Dialectología española*. Second edition (Biblioteca románica hispánica 3: Manuales). Madrid: Editorial Gredos.

Chapter 22

Preliminary notes on the Merge position of deictic, anaphoric, distal and proximal demonstratives

Guglielmo Cinque
Ca' Foscari University, Venice

In many languages the same demonstrative forms can be used either deictically (to point to some entity present in the speech act situation) or anaphorically (to refer back to some entity already mentioned in the previous discourse). In other languages deictic and anaphoric demonstratives are expressed by different forms, and in a subset of the latter group of languages the deictic and anaphoric demonstratives can co-occur, in a certain order. The two thus appear to be merged in different positions of the nominal extended projection, with deictic demonstratives arguably merged higher than anaphoric demonstratives, as is more clearly evident in certain languages. I submit that this is true of all languages even if most do not provide any overt indication of a different Merge position. Some languages also appear to provide evidence that distal and proximal demonstratives are merged in distinct positions of the nominal extended projection.

1 Introduction

Demonstratives, whether used deictically or anaphorically,[1] are usually taken to be merged in the same position of the extended nominal projection. While most languages do not provide evidence to the contrary, there are some that do show a

[1]Anaphoric demonstratives, together with "cataphoric" and "recognitional" demonstratives (the latter used for entities known from shared knowledge, Diessel 1999), are often termed "endophoric", and are opposed to "exophoric" (deictic) demonstratives, though anaphoric demonstratives may also show distal/proximal/etc. deictic distinctions. For simplicity I will keep here to the traditional terms "anaphoric" and "deictic".

Guglielmo Cinque. 2020. Preliminary notes on the Merge position of deictic, anaphoric, distal and proximal demonstratives. In András Bárány, Theresa Biberauer, Jamie Douglas & Sten Vikner (eds.), *Syntactic architecture and its consequences II: Between syntax and morphology*, 491–503. Berlin: Language Science Press. DOI: 10.5281/zenodo.4280669

distinct Merge position for their deictic and anaphoric demonstratives (pointing to a higher Merge position for the deictic ones). Rather than taking this to be a parametric difference among languages, I submit that all languages merge their deictic and anaphoric demonstratives in two distinct positions. This will simply not be visible in those languages where the two cannot co-occur and/or where nothing raises between the position occupied by anaphoric demonstratives and that occupied by deictic ones.

2 Languages where deictic and anaphoric demonstratives are formally distinct and can co-occur

I consider first those languages where the two types of demonstratives are represented by different forms[2] and overtly display their distinct Merge position by occurring together.

One such language is Ngiti, a Central Sudanic Nilo-Saharan language. Demonstrative, numeral and adjectival nominal modifiers precede the head noun (Kutsch Lojenga 1994: §9) and deictic demonstratives are formally distinct from anaphoric ones (cf. Kutsch Lojenga 1994: §§9.5.1–9.5.2). See (1a,b).[3]

(1) Ngiti (Kutsch Lojenga 1994: 373, 375)

a. yà dza
Dem$_{deictic}$ house
'this house'

b. ndɨ dza
Dem$_{anaphoric}$ house
'that house (mentioned before)'

As apparent from (2), the two types of demonstratives can co-occur, with the deictic demonstratives preceding the anaphoric ones:[4]

[2]Diessel (1999: §5.5) states that anaphoric demonstratives are morphologically more complex than deictic demonstratives, citing a number of languages where the former are formed by adding a morpheme to the latter. Dixon (2003: 76f) however, documents the opposite case, where the deictic demonstrative is formed by adding a morpheme to the anaphoric one. For the internal complexity of demonstratives, composed of a determiner and an adjectival deictic adjective, see Leu (2007; 2015: §2.5) (pace Kleiber 1986).

[3]The question arises whether the "anaphoric" demonstrative of Ngiti and that of the other languages mentioned below are distinct from determiners. In Loniu at least (see footnote 5 below) the post-nominal anaphoric and deictic demonstratives are distinct from the determiners, which are pre-nominal. In the other languages, which lack determiners, this is harder to tell, though the relevant grammatical descriptions seem not to assimilate the anaphoric demonstratives to determiners. I thank Richard Kayne for raising this general question. Possibly some of the anaphoric demonstratives discussed below correspond to the "neutral" demonstratives of Kayne (2014: §11).

[4]If nominal modifiers can move only as part of a constituent containing the N (Cinque 2005), the possibility that the deictic demonstrative of (2) is merged below the anaphoric one and is raised above it is not viable.

(2) Ngiti (Kutsch Lojenga 1994: 376)
yà ndɨ dza
$\text{Dem}_{\text{deictic}}$ $\text{Dem}_{\text{anaphoric}}$ house
'this house (mentioned before)'

As pre-nominal modifiers (as opposed to post-nominal ones) reflect the order of Merge, with elements to the left higher than those to the right (Kayne 1994, Cinque 2009; 2017), this language provides direct evidence that deictic demonstratives are merged higher than anaphoric demonstratives.

Another language showing the distinct Merge position of deictic and anaphoric demonstratives, with the former arguably higher than the latter, is the Papuan (Yam) language Komnzo.

In addition to deictic demonstratives, Komnzo has one demonstrative, *ane*, which

> has no spatial reference, but it is used for anaphoric reference. It marks a referent which has been established in the preceding context. […] It may combine with the proximal and the medial demonstrative identifier as can be seen in example [(3)] (Döhler 2016: 128f)

in the order N > anaphoric demonstrative > deictic demonstrative:

(3) Komnzo (Döhler 2016: 129)
fintäth *ane* *z=iyé* … yem=anme dagon.
PROPN $\text{Dem}_{\text{anaphoric}}$ PROX=3SG.M:NPST.be cassowary=POSS.NSG food
'This fintath (Semecarpus sp.) here is the cassowaries' food.'

The relative order of the two is with the anaphoric demonstrative closer to the noun than the deictic demonstrative, as was the case in Ngiti. The linear order, however, is the reverse, arguably due to the successive raising of the NP, with pied piping of the *whose picture*-type, first above the lower anaphoric and then above the higher deictic demonstrative dragging along the lower anaphoric one, with the result of reversing the order entirely (cf. Cinque 2005; 2017).

Identical to the Komnzo situation is that of the Alor Pantar (Papuan) language Kaere, where the anaphoric demonstrative *erang* can combine with the deictic demonstratives *ga* 'this' or *gu* 'that' (Klamer 2014: §4) (see 4), and that of the Oceanic language Loniu (Hamel 1994: §4.3.7), where the anaphoric demonstrative

nropo can co-occur with the deictic demonstrative *itiyen* 'that (relatively distant from speaker)' (see 5), in both cases with the order N $\text{Dem}_{\text{anaphoric}}$ $\text{Dem}_{\text{deictic}}$:[5]

(4) Kaere (Klamer 2014: 120)
kunang masik utug *erang* *gu*
child male three $\text{DEM}_{\text{anaphoric}}$ that
'those three boys (mentioned earlier)'

(5) Loniu (Hamel 1994: 99)
... hetow law a iy *nropo* *itiyen* ...
3PL.CL REL POSS 3SG $\text{DEM}_{\text{anaphoric}}$ $\text{DEM}_{\text{deictic}}$
'...to those aforementioned relatives of his ...'

The Austronesian, Malayo-Polynesian, languages Gayo (Eades 2005) and Nias (Brown 2005) and the Niger-Congo languages Samba Leko (Fabre 2004) and Kitalinga (Paluku 1998) instead show post-nominally the same order shown pre-nominally by Ngiti: NP > deictic demonstrative > anaphoric demonstrative:[6]

[5] "[T]he two together are equivalent to English 'aforementioned'" (Hamel 1994: 99). In addition to the anaphoric and deictic demonstratives in post-nominal position, Loniu appears to also have determiners, in pre-nominal position. "The order of constituents in the noun phrase is, generally, as shown in the formula in [(i)] below" (Hamel 1994: 89).

(i) (Det) Noun (Possessor NP) (Associated NP) (Descriptive Adjunct) (Quantifier) (Prepositional Phrase) (Relative Clause) (Demonstrative)

"The personal pronouns which function as determiner are the same as those used as nominals for subject, object, and so on. Although they may co-occur with inanimate nouns, the majority of NPs in the data which contain personal pronoun determiners are animate. [...] These personal pronoun determiners, however, seem to be present only in NPs which are definite" (Hamel 1994: 90). See the example in (ii):

(ii) Loniu
iy pihin iy huti kawa
3SG woman 3SG take basket
'The woman takes the basket'

[6] In (6b), the anaphoric demonstrative *nomema* contains *mema* 'earlier'. Adjectives and numerals follow the two demonstratives in that order (Brown 2001: 412). Another language with an anaphoric demonstrative meaning 'earlier/before' is Madurese:

(i) Madurese (Davies 2010: 192)f)
Reng lake' gella' entar ka Sorbaja
person male before go to Surabaja
'That man (we were talking about just now) went to Surabaja'

(6) a. Gayo (Eades 2005: 225))
Serule-*ni-ne*
Serule-this-earlier
'this Serule' [Serule-this-MENTIONED earlier] (the aforementioned Serule)

b. Nias (Brown 2005: 579))
Ba si'ulu wa e nama-da *andre*
CNJ noble DPTCL DPTCL father:MUT-1PL.INCL.POSS $\text{DEM}_{\text{Deictic}}$
nomema'e!?
$\text{DEM}_{\text{anaphoric}}$
'And you mean that ancestor you've been talking about was a noble!?'

c. Samba Leko (Fabre 2004: 173)
bā?–ā *yê* *dō*
iron $\text{DEM}_{\text{deictic}}$ $\text{DEM}_{\text{anaphoric}}$
'that iron we talked about' [our translation]

d. Kitalinga (Paluku 1998: 203)
omumelo ʏú-*nì-lá*
throat ?-$\text{DEM}_{\text{deictic}}$-$\text{DEM}_{\text{anaphoric}}$
'this aforementioned throat', orig. French 'gorge celui-ci – en question'

My interpretation of the orders in (6) is that they are derived by raising the NP (or constituents containing the NP) above the two demonstratives in one fell swoop (without pied piping) (cf. Cinque 2005; 2017).[7]

3 Languages where deictic and anaphoric demonstratives are formally distinct, occupy different positions, but cannot co-occur

In the Trans-New Guinea Alor-Pantar language Abui (Kratochvíl 2007: §3.5.2; 2011) "[t]he deictic demonstratives precede the head noun while the anaphoric demonstratives follow it" (Kratochvíl 2007: 156). See the overall structure of Abui determiner phrases in (7) (Kratochvíl 2007: 156), and the illustrative examples of the order of the two types of demonstratives in (8):

[7]For evidence that constituents appearing to the right of N/V/etc. cannot be taken to be merged there, but come to be there as a function of the N(P)/V(P)/etc. moving above them, see Cinque (2009).

(7) (DEM$_{deictic}$) (NPOSS PROPOSS-) N (NMOD) (ADJ/V) (QUANT) (DEM$_{anaphoric}$)

(8) a. Abui (Kratochvíl 2007: 111)
oro fala
DEM$_{deictic}$ house
'that house over there (far from us)'

b. Abui (Kratochvíl 2007: 114)
fala *to*
house DEM$_{anaphoric}$
'the house (you just talked about)'

If deictic demonstratives are merged higher than anaphoric demonstratives, the Abui DP internal order Dem$_{deictic}$ N A Num Dem$_{anaphoric}$ can be analysed as involving successive raisings of the NP, with pied piping of the *whose picture*-type above the lower anaphoric demonstrative but not above the higher deictic demonstrative.[8]

In the Dogon language Jamsay, where the deictic demonstrative follows the noun (cf. 9a)[9] and the anaphoric one precedes it (cf. 9b),[10] within the overall order ⟨Dem$_{anaphoric}$⟩ N A Num ⟨Dem$_{deictic}$⟩, the derivation must be different, involving raising of the constituent [Dem$_{anaphoric}$ N A Num] (itself obtained via raising of the NP around A and Num) above the higher deictic demonstrative (cf. Cinque 2005; 2017).

(9) a. Jamsay (Heath 2008: 161)
èjù *núŋò*
field.L DEM$_{deictic}$
'this/that field'

b. Jamsay (Heath 2008: 164)
kò kùmàndâw kù^{n} bé
DEM$_{anaphoric}$ Major DEF PL
'those (aforementioned) Majors'

[8] The situation in Topoke (Bantu, C53) is only slightly different, as "the anaphoric demonstrative always follows the noun, whereas other demonstratives can either precede or follow" (Van de Velde 2005: §2.4). This suggests that anaphoric demonstratives are obligatorily crossed over by the NP, while deictic demonstratives are crossed over by the NP only optionally. Only slightly different is the case of Rama (Chibchan; Craig Grinevald 1988: §6.6), where the deictic demonstrative is only pre-nominal while the anaphoric one "meaning 'previously mentioned' […] is found either pre- or post-nominally" (p. 15).

[9] "*núŋò* is deictic, and may be accompanied by pointing or a similar gesture" (Heath 2008: 162).

[10] "Unlike deictic [noun + *núŋò*], the phrase [*kò* + noun] is discourse anaphoric …" (Heath 2008: 164).

4 Languages where deictic and anaphoric demonstratives are formally identical, occupy different positions, but cannot co-occur

The same pattern is instantiated by a number of other languages, modulo the formal identity of the deictic and the anaphoric demonstratives.

Migdalski (2001: 142) notes that "demonstratives may either precede or follow a noun in Polish. The latter option is stylistically marked and is used only when the noun followed by a demonstrative has been previously mentioned, […] as in [(10)]":[11]

(10) Polish

a. *Ta* ksiazka
'*this* book'

b. Ksiazka *ta* (acceptable if the book has been mentioned previously)
book this

Here too it is possible to analyse the pattern in $\text{Dem}_{\text{deictic}}$ NP $\text{Dem}_{\text{anaphoric}}$ as involving raising of the NP (with possible pied piping) above the lower anaphoric demonstrative but not above the higher deictic one.[12]

The opposite pattern $\text{Dem}_{\text{anaphoric}}$ NP $\text{Dem}_{\text{deictic}}$ is instantiated by Thimbukushu (Bantu language of Namibia; Fisch 1998), where "[u]sually demonstratives

[11]The Polish situation recalls the semantic difference between pre- and post-nominal demonstratives in Spanish and Modern Greek (modulo the obligatory presence of a determiner in pre-nominal position when the demonstrative is post-nominal). As observed by Bernstein (1997) and Taboada (2007) for Spanish and Panagiotidis (2000) for Modern Greek, a post-nominal demonstrative is only interpreted anaphorically (unless a demonstrative reinforcer is added), while a pre-nominal one can be interpreted deictically. But see Brugè (2002: 50, n. 27) and Brugè (2000: §2.5.3, p. 167, n. 51) for discussion of a number of complexities and of differences among the Spanish distal and proximal demonstratives.

[12]In Italian, where no evidence exists of a different Merge position of deictic and anaphoric demonstratives, there is still a difference between the two in the possibility for the former but not for the latter, in its neuter usage (presumably with a silent head noun THING; cf. Kayne & Pollock 2009), to take a locative "reinforcer". See (i):

(i) a. Questo (*qui) non lo so
This (here) not it I.know
'This I don't know'

b. Quello (*lì) me lo sono chiesto anch'io
That (there) to.me it am asked even-I
'That I wondered myself'

[…] occur as postpositive determiners after the nouns to which they refer" (Fisch 1998: 50), see (11):

(11) Thimbukushu (Fisch 1998: 50)
[Mugenda *oyu*] na haka
guest this I like
'I like this guest'

"If the demonstrative preposes the noun, it carries the meaning of 'this aforementioned', 'this one mentioned'" (Fisch 1998: 50), see (12):[13]

(12) Thimbukushu (Fisch 1998: 50)
[*oyu* ngombe]
the.aforementioned cow
'this cow'

This pattern can be taken to involve no movement of the NP above the lower anaphoric demonstrative (or possibly movement of the NP in the *picture of whom*-mode, which has the effect of not changing the relative order of the two elements), and raising of the NP (or of larger constituents containing the NP) above the higher deictic demonstrative.

5 Languages where distal and proximal demonstratives occupy different positions

In Nawdm (Niger-Congo, Gur; Albro 1998: §2.4)

> there are two basic demonstratives […], corresponding to 'this' and 'that' in English. Their distribution within the DP is different. The demonstrative corresponding to 'this' appears at the end of the DP […], and the demonstrative corresponding to 'that' appears at the beginning of the DP.

See (13):

(13) Nawdm (Albro 1998: 6)
a. *làʔà* bà hɔ́l̀ə té tèréʔété:
that dog black CL.PL CL-two-CL
'those two black (big) dogs'

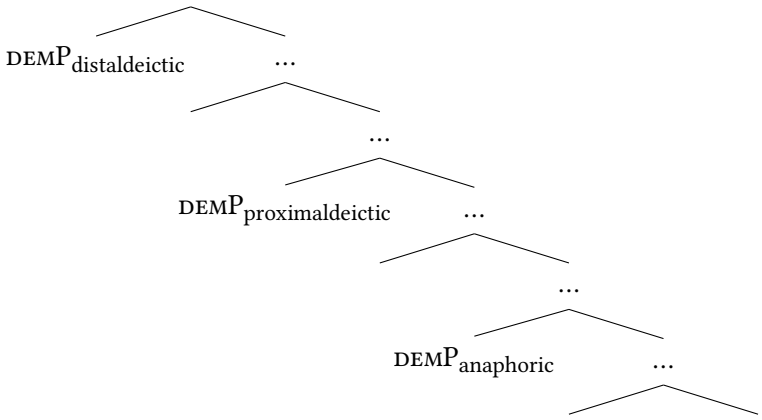
[13]Romanian appears to be similar. Post-nominal demonstratives have a deictic interpretation while pre-nominal ones, which belong to a non-colloquial style (cf. Brugè 2002: n. 32), have an anaphoric interpretation (Giusti 2005: 31; Nicolae 2013: 299f).

b. bà hɔ́l`ə té tèréʔètèn *tènté*
dog black CL.PL CL-two-CL CL-this-CL
'these two black (big) dogs'

According to Apronti (1971: 66ff), the same distribution (Dem_{that} N A Num and N A Num Dem_{this}) is found in the Kwa language Dangme.

It is thus tempting to assume that the distal and proximal deictic demonstratives occupy two distinct Merge positions, with distal demonstratives higher than proximal demonstratives, as shown in (14):

(14)



The order in Nawdm and Dangme would then involve raising of the NP with pied piping of the *whose picture*-type around A, Num and the lower proximal demonstrative, but not above the higher distal one, which then appears prenominally.

As in the case of Jamsay above, a different derivation must be involved to yield the order $Dem_{proximal}$ (Num) N (A) Dem_{distal} of Tigre (Afro-Asiatic, Semitic), where it is the proximal demonstrative that precedes the noun and the distal one that follows it (see 15):

(15) Tigre (Dryer 2013, after Raz 1983: 45)

a. *ʔəllan* ʔamʕəlāt
this.F.PL days
'these days'

b. ʔəb laʔawkād *lahay*
at time that.M
'at that time'

The NP must raise around A with pied piping of the *whose picture*-type (or with no pied piping), and then around Num and the lower proximal demonstrative with pied piping of the *picture of whom*-type, after which it raises around the higher distal demonstrative again with pied piping of the *whose picture*-type (a mixture of movements typically involved in the derivation of non-consistent languages; see Cinque 2017).

The fact that the two positions are presumably close to each other may give the impression in those languages where no material raises between them that they are one and the same position.

Abbreviations

1	first person
3	third person
CL	clitic
CNJ	conjunction
DEM	demonstrative
DPTCL	discourse particle
F	feminine
INCL	inclusive
L	low tone
M	masculine
MUT	mutated nominal
NPST	non-past
NSG	non-singular
PL	plural
POSS	possessive
PROPN	proper name
PROX	proximal, proximate
REL	relative
SG	singular

Acknowledgements

To Ian, to whom *nihil alienum est* in things linguistic, with fond memories and admiration. For helpful comments to a previous draft of this squib I am indebted to Laura Brugè, Richard Kayne and two anonymous reviewers.

References

Albro, Daniel M. 1998. Some investigations into the structure of DPs in Nawdm. Ms., UCLA.

Apronti, E.O. 1971. The structure of the nominal group in Dangme. *Journal of African Languages* 10(3). 65–72.

Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102(2–3). 87–13. DOI: 10.1016/S0024-3841(96)00046-0.

Brown, Lea. 2001. *A gramar of Nias Selatan*. University of Sydney. (Doctoral dissertation).

Brown, Lea. 2005. Nias. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 562–589. London: Routledge.

Brugè, Laura. 2000. *Categorie funzionali del nome nelle lingue romanze*. Milano: Cisalpino.

Brugè, Laura. 2002. The positions of demonstratives in the extended nominal projection. In Guglielmo Cinque (ed.), *Functional structure in DP and IP* (The Cartography of Syntactic Structures 1), 15–53. Oxford: Oxford University Press.

Cinque, Guglielmo. 2005. Deriving Greenberg's universal 20 and its exceptions. *Linguistic Inquiry* 36(3). 315–332. DOI: 10.1162/0024389054396917.

Cinque, Guglielmo. 2009. The fundamental left-right asymmetry of natural languages. In Sergio Scalise, Elisabetta Magni & Antonietta Bisetto (eds.), *Universals of language today*, 165–184. Dordrecht: Springer. DOI: 10.1007/978-1-4020-8825-4_9.

Cinque, Guglielmo. 2017. A microparametric approach the head-initial/head-final parameter. *Linguistic Analysis* 41(3–4). 309–366.

Craig Grinevald, Colette. 1988. A grammar of Rama. Report to the National Science Foundation. BNS 8511156, Université de Lyon.

Davies, William D. 2010. *A grammar of Madurese*. Berlin: De Gruyter.

Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins. DOI: 10.1075/tsl.42.

Dixon, R.M.W. 2003. Demonstratives: A cross-linguistic typology. *Studies in Language* 27(1). 61–112. DOI: 10.1075/sl.27.1.04dix.

Döhler, Christian. 2016. *Komnzo: A language of Southern New Guinea*. The Austrian National University. (Doctoral dissertation).

Dryer, Matthew S. 2013. Order of demonstrative and noun. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/88 (18 January, 2019).

Eades, Domenyk. 2005. *A grammar of Gayo: A language of Aceh, Sumatra*. Canberra: Pacific Linguistics.

Fabre, Gwenaëlle. 2004. *Le samba leko, langue Adamawa du Cameroun*. München: Lincom Europa.

Fisch, Maria. 1998. *Thimbukushu grammar*. Windhoek: Out of Africa Publishers.

Giusti, Giuliana. 2005. At the left periphery of the Romanian noun phrase. In Martine Coene & Liliane Tasmowski (eds.), *On space and time in language*, 23–49. Cluj Napoca: Clusium.

Hamel, Patricia J. 1994. *A grammar and lexicon of Loniu, Papua New Guinea*. Canberra: Pacific Linguistics.

Heath, Jeffrey. 2008. *A grammar of Jamsay*. Berlin: De Gruyter.

Kayne, Richard S. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kayne, Richard S. 2014. Why isn't *this* a complementizer? In Peter Svenonius (ed.), *Functional structure from top to toe* (The Cartography of Syntactic Structures 9), 188–231. Oxford: Oxford University Press.

Kayne, Richard S. & Jean-Yves Pollock. 2009. Notes on French and English demonstratives. Ms., New York University and Université de Paris Est, EA 4120.

Klamer, Marian. 2014. Kaera. In Antoinette Schapper (ed.), *The Papuan languages of Timor, Alor and Pantar: Volume 1, Sketch grammars*, 97–146. Berlin: De Gruyter. DOI: 10.1515/9781614515241.97.

Kleiber, Georges. 1986. À propos de l'analyse adjectif démonstratif = article défini + élément déictique ou: Sur l'irréductibilité des symboles indexicaux. In *Actes du VIIe congrès international de linguistique et de philologie romanes, Aix-en-Provence, 29 août–3 septembre 1983*, vol. A, 195–212. Aix: Publications de l'Université de Provence.

Kratochvíl, František. 2007. *A grammar of Abui: A Papuan language of Alor*. Universiteit Leiden. (Doctoral dissertation).

Kratochvíl, František. 2011. Discourse-structuring functions of Abui demonstratives. In Foong Ha Yap, Karen Grunow-Hårsta & Janick Wrona (eds.), *Nominalization in Asian languages: Diachronic and typological perspectives*, 757–788. Amsterdam: John Benjamins. DOI: 10.1075/tsl.96.26kra.

Kutsch Lojenga, Constance. 1994. *Ngiti: A Central-Sudanic language of Zaire*. Köln: Köppe.

Leu, Thomas. 2007. These HERE demonstratives. *University of Pennsylvania Working Papers in Linguistics* 13(1), 12. 141–154.

Leu, Thomas. 2015. *The architecture of determiners*. New York: Oxford University Press.

Migdalski, Krzysztof. 2001. A determiner phrase approach to the structure of Polish nominals. In Adam Przepiórkowski & Piotr Barański (eds.), *Generative linguistics in Poland: Syntax and morphosyntax*, 137–150. Warszawa: Instytut podstaw informatyki PAN.

Nicolae, Alexandru. 2013. Demonstratives. In Gabriela Pană Dindelegan (ed.), *The grammar of Romanian*, 294–300. Oxford: Oxford University Press.

Paluku, André Mbula. 1998. *Description grammaticale du Kitalinga (Langue bantu du Nord-Est du Zaïre)*. München: Lincom Europa.

Panagiotidis, Phoevos. 2000. Demonstrative determiners and operators: The case of Greek. *Lingua* 110(10). 717–742. DOI: 10.1016/S0024-3841(00)00014-0.

Raz, Shlomo. 1983. *Tigre grammar and texts*. Malibu: Undena.
Taboada, Inma. 2007. Prenominal and postnominal demonstratives in Spanish: A [±deictic] approach. *Anuario del Seminario de Filologí Vasca "Julio de Urquijo"* 41(2). 323–332.
Van de Velde, Mark. 2005. The order of noun and demonstrative in Bantu. Ms., University of Leuven/FWO, http://llacan.vjf.cnrs.fr/pers/vandevelde/files/pdfs/The-order-of-noun-and-demonstrative-in-Bantu.pdf.

Name index

Language index

Subject index

Syntactic architecture and its consequences II

This volume collects novel contributions to comparative generative linguistics that "rethink" existing approaches to an extensive range of phenomena, domains, and architectural questions in linguistic theory. At the heart of the contributions is the tension between descriptive and explanatory adequacy which has long animated generative linguistics and which continues to grow thanks to the increasing amount and diversity of data available to us.

The chapters address research questions in comparative morphosyntax, including the modelling of syntactic categories, relative clauses, and demonstrative systems. Many of these contributions show the influence of research by Ian Roberts and collaborators and give the reader a sense of the lively nature of current discussion of topics in morphosyntax and morphosyntactic variation.

This book is complemented by two other volumes.

ISBN 978-3-96110-288-4