HERITAGE LANGUAGES & SYNTACTIC THEORY

edited by
Roberta **D'Alessandro**
Michael T. **Putnam**
Silvia **Terenghi**

Heritage Languages and Syntactic Theory

Heritage Languages and Syntactic Theory

Edited by

ROBERTA D'ALESSANDRO
MICHAEL T. PUTNAM
SILVIA TERENGHI

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries.

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging-in-Publication Data: 2024951638

ISBN 9780198876182

DOI: 10.1093/9780191987731.001.0001

Printed and bound by
CPI Group (UK) Ltd, Croydon, CR0 4YY

Contents

PART IV THE DP

List of abbreviations

BP	Brazilian Portuguese
CANS	Corpus of American Nordic Speech
CG	Clausal Gerund
CIC	Change in Contact
CLLD	Clitic left Dislocation
DOM	Differential Object Marking
ECM	Exceptionally Case Marked
EP	European Portuguese
HL	Heritage Language
IRH	Implicational Referential Hierarchy
LEE	Left-edge Ellipsis
PAH	Position of Antecedent Hypothesis
PAS	Position of Antecedent Strategy
PD	Pennsylvania Dutch
QP	Quantifier Phrase
RQ	Research Question
SCH	Syntactic Complexity Hypothesis
UG	Universal Grammar

List of contributors

Artemis Alexiadou is Director, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Germany, Professor of General Linguistics, Humboldt Universität zu Berlin, Germany, Director, Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS), Germany, and Professor of General Linguistics, Humboldt Universität zu Berlin, Germany. https://orcid.org/ 0000-0002-6790-232X

Luigi Andriani is Junior Professor of Italian and Spanish Linguistics, Institut für Romanistik, Universität Hamburg, Germany. https://orcid.org/0000-0002-8543-155X

Roberta D'Alessandro is Professor of Linguistics/Syntax and Language Variation, Utrecht University, Utrecht, Netherlands.

Cristina Flores PhD is Professor of German Linguistics, Departamento de Estudos Germanísticos e Eslavos, Centro de Estudos Humanísticos, Universidade do Minho, Braga, Portugal. https://orcid.org/0000-0001-5629-9556

Alberto Frasson PhD is Post-Doctoral Researcher in Linguistics at the Faculty of Neophilology, University of Wrocław, Poland. https://orcid.org/0000-0001-9619-2761

Oksana Laleko PhD is Professor of English and Linguistics, State University of New York at New Paltz. https://orcid.org/0000-0001-5611-1912

Terje Lohndal PhD is Professor of English linguistics, Norwegian University of Science and Technology, Trondheim, Norway, and Adjunct Professor of Linguistics, UiT The Arctic University of Norway, Tromsø, Norway. https://orcid.org/0000-0002-8514-1499

Manuela Pinto PhD is Assistant Professor, Intercultural Communication, Utrecht University, Netherlands. https://orcid.org/0000-0003-1893-1797

Maria Polinsky PhD is Professor, Department of Linguistics, University of Maryland, College Park, MD, US. https://orcid.org/0000-0003-1460-5089

Michael T. Putnam PhD is Professor of German and Linguistics, Penn State University, University Park, PA, US. https://orcid.org/0000-0002-7758-8266

Esther Rinke is Professor of Portuguese and Spanish Linguistics, Goethe Universität Frankfurt, Frankfurt, Germany. https://orcid.org/0000-0001-9378-9079

Vasiliki Rizou is PhD Candidate in Linguistics, Department of German Studies and Linguistics, Humboldt Universität zu Berlin, Berlin, Germany. https://orcid.org/0000-0002-5804-4976

Luana Sorgini MA is PhD candidate, Utrecht Univerity, Utrecht, Netherlands.

Silvia Terenghi is Assistant Professor of Multilingualism, Institute for Language Sciences, Utrecht University, Netherlands. https://orcid.org/0000-0002-6907-8602

Yvonne van Baal PhD is Researcher, Department of Language and Literature, NTNU – Norwegian University of Science and Technology, Trondheim, Norway, and Associate Professor of Nordic Languages, Department of Cultural Studies and Languages, University of Stavanger, Norway. https://orcid.org/0000-0002-1766-2911

This volume has been financed by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement No. 681959_Microcontact).

Heritage languages and syntactic theory

An introduction

Roberta D'Alessandro, Michael T. Putnam, and Silvia Terenghi

1 Introduction

Heritage languages (HLs) are spoken in restricted environments, such as emigrant language communities or families, surrounded by a stronger and more dominant language (Polinsky 2018). Over the past two decades, HLs have taken a prominent position in linguistic research due to the significant attention they attract from scholars who are focused on exploring linguistic diversity from myriad perspectives. By virtue of their inherent diversity stemming from varied initial learning conditions, HLs pose significant challenges to traditional methods of linguistic description that rely on a uniform and monolithic conception of language. Even though numerous psycholinguistic studies have recognized certain overarching patterns in the acquisition of HLs (Montrul 2016), the integration of data from HL grammars – both with respect to language acquisition and language maintenance across the lifespan – into larger discussions (focused primarily on theory-shaping) has not yet taken place. One factor that may have prevented this is the observation of occasional (significant) variance in the grammars of some HL speakers when compared with monolingual baselines. Such comparisons run the risk of viewing bilingual outputs as being "deviant" or "incomplete" (Montrul 2008) – especially those found among HL speakers.[1] In light of this view of HL grammars, little has been done concerning the use of data and generalizations gained from the analysis of said data in theory-building efforts.

[1] Since their inception, scholars promoting generative treatments of aspects of HL grammars have struggled with the proper way to describe mental representations that appear to (significantly) differ from monolingual baselines. Montrul (2008, 2016) appeals to the term of "*incomplete acquisition*"; however, this distinction has been roundly criticized both on theoretical (Putnam and Sánchez 2013) and sociolinguistic grounds (Bayram et al. 2019). Speaking strictly from the perspective of the quality of mental representations, Domínguez, Hicks, and Slabakova (2019) advance the claim that the distinction of *incomplete acquisition* holds as an accurate depiction of these representations (in some instances), while Perez-Cortes, Putnam, and Sánchez (2019) provide a more gradient view of the quality of HL syntactic representations that does not support this designation.

Roberta D'Alessandro, Michael T. Putnam, and Silvia Terenghi, Heritage languages and syntactic theory.
In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro et al., Oxford University Press.
 DOI: 10.1093/9780191987731.003.0001

Two additional trends in generative theory-building efforts are worth noting here: First, research on bi- and multilinguals from experimental, formal, and functional approaches has aided in generating an increased interest in these communities (D'Alessandro, Natvig, and Putnam 2021; Grosjean 2008; Polinsky 2018; Rothman, González Alonso, and Puig-Mayenco 2019). Second, the inclusion of typologically distinct languages – especially those that are minority or endangered (Nevins 2022) – has also been a motivating factor to pay more attention to HL grammars. What modern HL studies contribute to the traditional methodology regarding comparative microvariation is attention to individual speakers and their language development patterns (for an overview, see Luk and Rothman 2022). Redirecting focus toward the *circumstances* of language acquisition has occasionally led to a reduced emphasis on structural patterns, giving rise to a fresh array of scientific inquiries concerning the precise influence of developmental and social elements in HL acquisition. One key question that has emerged is whether these languages can genuinely be regarded as equivalent to monolingual grammars in terms of completeness and complexity. The emerging consensus suggests that HL grammars are indeed quite complex in spite of the qualitative and quantitative differences in input that heritage speakers receive during formative years of language acquisition and throughout the course of life in opportunities to produce and comprehend their HL (Polinsky and Putnam 2024).

This volume aims to direct attention back to structural issues. Once it is ascertained that, despite their diversity, HL grammars are complete systems worthy of serious investigation (D'Alessandro et al. 2021; Polinsky and Scontras 2020), we can confidently adopt the methodological assumption that the investigation of HL syntactic structures is a valid and meaningful pursuit. Even more: precisely because their patterns are deviant or different from those of monolingual grammars, they can provide supplementary insights into syntactic theory and contribute to addressing the question of what a natural grammar can or cannot generate (D'Alessandro 2021a). The chapters collected here all stem from the same methodological assumption, namely that structural phenomena found in HLs, particularly if they are different (or "deviant") from those which are found in monolingual languages, can offer precious empirical evidence to test syntactic hypotheses and ultimately for the development of syntactic theory.

This introduction includes a note on the conceptualization of the ideal speaker-hearer (Section 2), then turns to discuss the issue of variable input and shared syntax (Section 3). Section 4 highlights shared development trends and trajectories found in typologically distant HLs. The existence of these

trends confirms the development of large-scale, systemic representational "constraints" that are best modeled in a comprehensive theoretical approach to syntactic structure. We continue with a reflection on the importance of the search for invariant principles and what data from HLs can do to help in this enterprise (Section 5). Finally, Section 6 presents an overview of the chapters, highlighting the contribution they offer to the development of syntactic theory.

2 The *ideal speaker-hearer* concept revisited

We understand a *generative grammar* to be one that is *formal* and *explicit*. A core desideratum of generative syntactic theory-building efforts has traditionally rested on the notion of *the ideal speaker-hearer*. This archetype finds its origin in the following quote by Chomsky (1965: 3):

> Linguistic theory is concerned primarily with an ideal speaker-listener, in a completely homogeneous speech-community, who knows its language perfectly and is unaffected by such grammatically irrelevant conditions as memory limitations, distractions, shifts of attention and interest, and errors (random or characteristic) in applying his knowledge of the language in actual performance.

The notion of a *grammar* of a particular language under these contexts "purports to be a description of the ideal speaker-hearer's intrinsic competence" (Chomsky 1965: 4). Within the same framework, the earliest attempts toward building a generative grammar were held to be guided by innate principles, e.g. Universal Grammar (UG), which represented invariant principles of language and cognition that were neither acquired through experience nor affected by "non-linguistic" factors. The past 70 years have borne witness to interesting proposals within the generative enterprise that have resulted in shifts in focus at both the operational and architectural levels. Some proposed adjustments have been relatively small in nature, while others have been seismic, necessitating a reconceptualization of the entire system. The Minimalist Program (Chomsky 1995) represents a particular implementation of a generative grammar (while acknowledging that possible architectural designs and associated operational procedures are permissible) which has dominated the syntactic scene in the last 30 years.

One of the notions that has been most affected by the evolution of syntactic theory is that of the *ideal speaker-hearer*. In light of recent developments and minimalist assumptions, the concept of I-language is best understood (for

the most part) as the cardinal value of the systematic relations that are established during language acquisition and maintained throughout the course of the lifespan. *The ideal speaker-hearer* is that which embodies these systematic relations. The competence (and performance) of individuals will of course display a wide array of variance; however, notwithstanding impairment or pathological disorders, languages will show coherent and systematic structural realizations, and underlying general invariable principles, since invariant structure-building procedures will remain constant (see Lohndal 2013 for a related discussion).

The notion of *parameters* and their development over the course of the last 40 years are another excellent case in point of these developmental trends. Once thought to be innate principles that assist language acquirers in sifting through ambient input, parameters are nowadays no longer regarded as such (Eguren, Fernández-Soriano, and Mendikoetxea 2016). Rather, they are viewed as the result of more fundamental processes; i.e. as "third factor" elements that are second-order objects derived from more general principles and axioms (Biberauer 2019; Boeckx 2014; Chomsky 2005). Continued research on the cognitive underpinnings of the design of the faculty of language supports the perspective of a radically reduced UG, with some prominent scholars, such as Lightfoot (2020), calling for an "open UG" that is guided and determined by the parsing of input. Variants of minimalist grammars are now paired with statistical learning models (Lidz and Gagliardi 2015; Yang 2016, 2018), further demonstrating the attractiveness of the appeal to an "open UG" that allows parsing strategies to be led by a finite number of structure-building principles. Arriving at this eventual conclusion was, in fact, the desired outcome in an architecture that views linguistic expressions as "the optimal realization conditions of the interface conditions, where 'optimality' is determined by the economy conditions of UG" (Chomsky 1995: 171). The refinement of minimalist inquiries into the nature of grammar continues through the present, which is a clear sign of a growing and vibrant field of inquiry. Probing into the accomplishments and advances of the Minimalist Program (Leivada 2020), while simultaneously acknowledging architectural, operational, and terminological elements that require revision or elimination (Leivada and Murphy 2021), has become a healthy common practice that shapes the landscape of generative linguistics today.

Concepts such as *the ideal speaker-hearer*, parameters, and UG have also been widely discredited in research on bi- and multi-competence individuals and populations (commonly referred to as "*bilinguals*"). Agreeing with Luk and Bialystok (2013), we too support the view that the distinction of

bilingual is not a categorial variable; i.e., it is virtually impossible to find populations of bilinguals who have had exactly the same experiences, especially with respect to qualitative and quantitative input exposure during the language acquisition process. The reliance on parameters has also largely been abandoned in generative approaches to L2 acquisition and heritage languages in favor of feature-based models (Putnam 2019, 2020a; Putnam, Perez-Cortes, and Sánchez 2019; Slabakova 2009; Slabakova et al. 2020). In this volume, we focus our collective attention on syntactic properties and trends observed in heritage varieties of various languages. Collectively, the chapters in this volume adopt this modern minimalist perspective in their individual studies, which holds that the invariant structure-building procedures remain constant in these grammars, as stated above (Lohndal 2021; Polinsky 2018). There are, however, noted differences, sometimes referred to as *divergent developments*, which result in the development of HLs, especially in connection with Factors 2 (Primary Linguistic Data) and 3 (principles of data analysis and computational efficiency) (Chomsky 2005). One of the principal debates the chapters in this volume investigate is whether or not a finite set of developmental trends in HL syntax is discernible from the studies that now exist (Polinsky and Scontras 2020). The answer provided by the chapters contained in this volume is positive: despite the considerable diversity in the many factors relevant to language acquisition, it is possible to discern robust developmental patterns and common trends that are shared among all heritage language grammars.

3 Shared syntax and variable input

Critics of generative models shaped by some version of *the ideal speaker-hearer* often point to variability in input – both from a qualitative and quantitative perspective – as a significant challenge. This state of affairs is particularly poignant in our treatment of HL speakers, who, as bilinguals, constitute an extremely heterogeneous group (Luk and Bialystok 2013). HL speakers vary significantly from one another in this respect, as well as opportunities later in life to continue to use their HL in myriad settings and domains. A highly relevant and justified question is whether or not this variable input may result in mental representations that diverge from a monolingual baseline. This question rests at the center of discussions and debates focusing on whether or not aspects of grammatical competence can be *incompletely* acquired (Domínguez et al. 2019; Montrul 2008, 2016). Although this controversial proposal, and the large amount of research it has generated, is certainly worthy of continued

debate and inquiry, a detailed treatment of this topic will not be carried out here. For our immediate purposes, however, we wish to point out that the extant body of research on HL syntax suggests that these core-structural building principles remain largely intact in these grammar systems. The domains in which *incomplete acquisition* or *attrition* may affect HL grammars are generally assumed to be restricted to interface conditions.[2]

We affirmed in the previous section that minimalist generative models are compatible with variationist and statistic approaches to language acquisition (and change) (Lidz and Gagliardi 2015; Yang 2016, 2018), and adopt the position that a finite number of operational mechanisms, e.g. Merge, Agree, etc., restrict the number of possible grammars learners need to consider (Gould 2017; Leivada 2015; Moro 2008). A separate, yet extremely relevant matter that deserves our attention at this point is whether or not we assume that these core structure-building principles construct two separate syntactic systems, or one combined, or "shared," syntax. The consensus in current psycholinguistic and neurolinguistic studies is that syntactic structures occupy a shared cognitive space (Branigan and Pickering 2017; Kroll and Gollan 2014; Putnam, Carlson, and Reitter 2018). Some of the strongest evidence adduced in support of this concept of "shared syntax" is found in priming effects and reaction times in connection with related structures. In sum, even though a language may not be "active" at the moment, latent effects of its structure impact parsing strategies for the grammar currently in use. This proposal is not inconsequential, because it raises a number of serious questions for the study of bilingual syntax more generally, and HL syntax more specifically. The first question concerns how the mind switches between two language systems. In a model of "shared syntax," we should anticipate the principal area of variance to reside in the realization of syntactico-semantic features post-syntactically (at PF). To phrase this as a question, how might this be particularly challenging for HL speakers? One possibility is that languages, such as HLs, that are infrequently activated, or may have truncated representations, must ward off the intrusion of elements and parsing preferences from "stronger" grammars, resulting in the need for bilinguals to develop "adaptive control" to mediate between these two integrated systems (Green and Abutalebi 2013). Grosjean (2008: 63) points out that "future research will have to investigate the underlying mechanisms that

[2] *Incomplete acquisition* is usually meant as the process whereby HL speakers do not attain a complete grammar, equivalent to those of monolingual speakers, due to quantitatively and qualitatively different exposure; *attrition* is the phenomenon whereby the grammar of a speaker erodes due to long exposure to a different dominant language.

make a stronger language 'seep through'" even in instances when the stronger language has been "deactivated."

The development of any notion of "adaptive control" involves two essential prerequisites: (i) a better understanding of the (third factor-) cognitive factors that can adjudicate between (ii) linguistic representations. This brings us to our second question, namely, how are bilinguals able to navigate through competing, or *dueling* grammars (Myers-Scotton 2002)? A detailed account of how the representations in both grammars exist simultaneously in a bilingual's mind is essential here. Moreover, we need a set of universal features in order to ensure that the elements in one source grammar are compatible and equivalent with the other. In the absence of a universalist approach to features and syntactic structures, this is simply not possible (Reiss in press). Putnam et al. (2018) highlight the necessity of features and representations in order to determine an approximate measure of the typological proximity (and, contrariwise, distance) that exists between two (or more) languages. This point highlights the importance of including situations of *microcontact*, i.e. where the two grammars under investigation are "more closely related" than others, alongside studies involving two (or more) "more distant" grammars (Andriani et al. 2022b; D'Alessandro 2021a).

HL grammars are aptly classified as hybrid grammars (Aboh 2015); however, in all fairness, the "strong" grammar in these dyads can be influenced to various degrees, too. Scontras and Putnam (2020) call for increased attention to the dyads in which HL grammars exist, and this point reinforces the importance of including situations involving *microcontact* into this more comprehensive treatment of HL syntax, although these comparisons are only possible under the assumptions that the atoms that constitute the make-up of one grammar equate to those in the other. This presupposition is often taken for granted in formal syntax that focuses predominantly on monolinguals (i.e., *ideal speaker-hearers*); however, once we incorporate bilinguals into our research as legitimate grammars, the need for universals comes to the forefront.

4 Developmental trends in HL syntax

From a holistic perspective, the extant research on HL syntax suggests a number of developmental trends observed across HL grammars of varying proficiencies and dyads (Lohndal 2021; Polinsky 2018; Polinsky and Scontras 2020). As one might expect, invariant structure-building principles such as

Merge remain unaffected; however, other aspects of syntactic representations have proven to be more vulnerable to changes due to the status of these languages as HLs.

Morphology has proven to be a domain of significant alteration in the course of HL acquisition and development. Putnam, Schwarz, and Hoffman (2021) survey and summarize a number of common characteristics of HL morphological systems (see also Polinsky 2018: Chapter 5 for a related treatment of these issues). These include (but are not limited to) phenomena such as the overregularization and overmarking of unmarked forms (often resulting in paradigmatic leveling) and the increase in analyticity, i.e. the push toward one-to-one form-meaning correspondences. A specific implementation of this preference for analytic forms has been advanced by Laleko and Polinsky (2017), which they call the "Silent Problem." According to this proposal, bilinguals, including HLs, show a strong dispreference for "empty elements" in the syntax and their subsequent null realization at an external interface.

In line with a modular interpretation of the cognitive underpinnings of the language faculty, these morphological trends observed in HLs are perhaps best understood as an "interface phenomenon" that exists between syntax and morphology (however one wishes to construe what "morphology" means here). Morphological and morphosyntactic changes can certainly have a substantial impact on syntactic representations and operations that affect them; therefore, it stands to reason that a majority of research on HL "syntax" to date has focused on the syntax-morphology interface. Interfaces between syntax and other modules of grammar have been the focus of a substantial amount of research in HL syntax and bi/multilingualism more generally (see Sorace 2011 for a discussion of the central importance of the concept of "interface" in bilingualism research). The interface between syntax and discourse has been a particularly fruitful domain of study and inquiry. For example, Hoot and Leal (2023) investigate how information focus may affect the word order of heritage Spanish.

Developments in HL syntax are not exclusively restricted to interface phenomena; in fact, there are some properties that suggest a restriction in the size of computational domains utilized in HL grammars. Two phenomena that have been the focus of study in HL syntax in this regard concern (i) long-distance anaphoric binding and (ii) *wh*-movement involving intermediate gaps. Previous research on heritage Korean (Kim, Montrul, and Yoon 2009, 2010), Icelandic (Putnam and Arnbjörnsdóttir 2015), and Turkish (Gürel 2004) evince the reduction in binding domains that are associated with weak pronominal reflexives in HL varieties. In the baseline variant of

these aforementioned HL grammars, each language has multiple ways of establishing antecedent-referent dependencies; however, in the heritage variants, the dependency that represents the shortest distance generalizes as the only possibility. In similar fashion, *wh*-movement involving intermediate gaps has also been shown to be problematic for HL speakers with intermediate to low proficiency. Hopp, Putnam, and Vosburg (2019) demonstrate a restriction in the licensing of these complex filler-gap dependencies in Mennonite Low German (aka Plautdietsch), spoken in southwest Kansas. Based on data obtained from two production tasks, Hopp et al. (2019) observed that Plautdietsch speakers with intermediate to low proficiency (based on speaker self-ratings) would produce an additional *wh*-item in place of the intermediate gap position. These findings are important for two reasons: First, they confirm that some version of derivational complexity is at play in shaping the computational domains in HL syntax. Second, and related to the first point, any notion of derivational complexity intrinsically ties to some conceptualization of locality and how it applies to abstract representations.

5 The need for invariant operations

We agree with the observation regarding heritage language grammars made by Bayram, Di Pisa, Rothman, and Slabakova (2021: 545), who recognize that the field of heritage linguistics "has emerged in earnest over the past two decades as part of a wave of expansion in investigating bilingual language scenarios beyond the two previously widely studied cases: (i) simultaneous child bilingual development (2L1) and (ii) nonnative adult second language (L2) acquisition." Although they differ in fundamental ways from monolinguals and other bi/multilingual populations, HL speakers constitute a special case of individuals who as "children [were] exposed to a language from birth [and] who nevertheless appear to deviate from the expected native-like mastery in pronounced and principled ways" (Polinsky and Scontras 2020: 2). The wide variance in abilities and grammatical expressions is not without form and direction. These data, although extremely fascinating, are not the deliverable in scientific inquiry; rather, it is the lens provided by formal proposals that enables us, as linguists, to interpret these data, leading to trends and generalizations that extend beyond individual speakers and languages.

Invariant operations are a fundamental prerequisite for interpreting (highly) variable data such as those found in heritage language grammars.

D'Alessandro et al. (2021) offer a detailed summary of how theoretical proposals are essential to gaining a deeper understanding of developmental trends found in heritage language grammars. One of the goals of this volume is to clearly demonstrate that the treatment of aspects of HL syntax is mutually beneficial for scholars interested in HL grammars and the theory-building efforts. Even in the face of significant representational differences when compared with more proficient speakers, we see how generative approaches to syntax contribute to a potentially deeper understanding of these phenomena. A final important piece of evidence in favor of the inclusion of speakers with limited capacity in their HL grammar comes from Sherkina-Lieber's (2015) study of the ability of *receptive bilinguals*, i.e. individuals who can comprehend a language, but only produce it with pronounced difficulty, to understand functional morphology. Sherkina-Lieber (2015) investigated whether *receptive bilinguals* could decode tense, aspect, and agreement in heritage Labrador Inuttitut. The results of her study confirm that these speakers could interpret overarching categories, although they struggled with more specific intra-categorical distinctions. Assuming that the morphological reflexes of tense, aspect, and agreement are vehicles of syntactico-semantic features, these findings affirm the permanence, i.e. invulnerability, of certain elements of mental representations, while also making the case for the possibility of the weakening of connections between elements of syntactic structure (e.g., features) and their interpretation at the interfaces.

The syntax of HL grammars has also proven to be fertile ground for the creation of innovative structures, confirming once again the need for formal analysis to aid in better understanding these sorts of developments (Bousquette and Putnam 2020; Rinke and Flores 2014). Examples of these novel forms in HL syntax are found in case systems that show the development of Direct Object Marking in Germanic, a language that commonly does not have such forms (Yager et al. 2015); the development of proclitic forms involving separable prefix verbs in Gottscheerisch, a Bavarian German-based dialect spoken in Slovenia (Putnam and Hoffman 2021); and in the syntax of infinitives in North American Norwegian (Putnam and Søfteland 2022).

6 Summary of chapters

Against this background, the chapters collected in this book successfully showcase both the crucial role that syntactic theory plays in allowing us to better understand developments and common trends in heritage languages,

and, contrariwise, how theory-building efforts (specifically in the syntactic domain) benefit from the inclusion of heritage grammars into formal theorizing.

The first two chapters take a general approach to linguistic theory and language variation from the standpoint of heritage languages. In *Microcontact and syntactic theory*, D'Alessandro et al. delve into the underlying principles that govern syntactic change resulting from language contact. They do so by investigating Italo-Romance heritage languages by means of the microcontact methodology, which involves performing pairwise comparisons between one and the same heritage variety across different dominant languages. This chapter discusses the main generalization that can be drawn from various case studies performed with this methodology: namely, that the nature of a given syntactic phenomenon influences its reaction to contact-induced change. In particular, phenomena that involve core syntactic features (such as auxiliary selection and demonstratives) follow a predictable pattern of change, while phenomena that are discourse related (such as null subjects and differential object marking (DOM)) exhibit more complex patterns of change, as further determined by (perceived) structural similarity across the languages in a given dyad. The authors argue that, whenever the locus of variation can be neatly perceived by speakers, then 1:1 structural mapping between the two relevant languages is possible, allowing the speaker to identify complex or redundant structures and, possibly, to simplify them; this is the situation most commonly described for heritage languages, for which clear-cut divergences with respect to the baseline and homeland varieties, often discussed in terms of simplification, can be identified. However, if the speakers cannot perceive the locus of variation, as is the case in microcontact situations, then 1:1 structural mapping between the two relevant languages is not possible and the outcome of contact is more unpredictable; crucially, this does not neatly align with existing models for heritage languages. As such, D'Alessandro et al. highlight the significance of incorporating the microcontact dimension into the analysis of change in contact, to clarify the conditions that drive change.

Rinke and Flores consider more closely how variation, amplified in the input (the Primary Linguistic Data, or second factor, in Chomskyian terms) that heritage speakers receive, fundamentally shapes the idiosyncrasies observed in heritage grammars. More specifically, focusing on research on heritage European Portuguese, their chapter *Systematic and predictable variation in heritage grammars: The role of complexity, diachronic change, and linguistic ambiguity in the input* identifies three distinct dimensions of variation in the input heritage speakers receive. The authors argue that, once

those aspects of variation inherent to the input received by heritage speakers are factored in, apparently "deviant" phenomena in heritage European Portuguese are in fact fully predictable. As per the title, variation can be induced by the complexity of given linguistic phenomena, such as clitic placement; by diachronic change that is ongoing in the baseline grammar, as exemplified by the case of null objects; and by ambiguity, be it in the lexical or in the grammatical domain, as illustrated by the interpretation of overt subject pronouns and by the multifunctionality of the lexical item *que*, respectively. This chapter highlights both the need for correct descriptions of the data, and in particular of the variation inherent to them, as a prerequisite for the correct interpretation of seemingly "deviant" features and, in turn, how looking at acquisition in heritage contexts grants us enhanced theoretical understanding of the systematic nature of linguistic variation and change.

The chapters by Polinsky and Laleko shift the focus to sentence structure in heritage languages. In *Heritage language gaps*, Polinsky investigates how overt and null expressions alternate within heritage languages and takes issue with the widely accepted claim that heritage languages tend to favor overt expressions over null ones (the already mentioned "Silent Problem" of Laleko and Polinsky 2017). Upon closer inspection, Polinsky concludes that the licensing conditions for null elements are intact in heritage languages; the apparent difficulties that heritage speakers have with null elements are rather epiphenomenal to the shortening or even the elimination of the long-distance dependencies in which such null elements typically occur. This revision is based on data coming from the domains of both overt vs. null pronouns and relative clauses. In relation to the former, it is observed that, besides the lack of perceptual salience displayed by null elements, other factors that contribute to their difficulty for heritage speakers include the variation that the input attests when it comes to determining the antecedent of overt vs. null pronouns (this is a matter of preference, not a categorical rule) and general difficulties with long-distance dependencies. This latter point is explored more in detail in relation to relative clauses: relative clauses admit both anaphoric and syntactic dependencies, but Polinsky shows that heritage speakers prefer the former type, as realized by coindexation, over the latter, which would require instead A-bar movement. Also in this case, the grammar of relativization in the baseline is not categorical but allows for optionality: Polinsky speculates that this indeterminacy might ultimately propel reanalysis in heritage grammars. Importantly for the wider aims of this volume, Polinsky notes various missing data points in the domain of syntactic gaps and calls for a better integration of the

empirical observations, available and still to be gathered, relative to different heritage languages, working toward a more general theory-building effort.

Laleko, too, underscores the validity and the importance of heritage languages as yet another empirical domain for linguistic theorizing. Laleko's chapter, *Word order and prosody in the expression of information structure*, substantiates this call by exploring the formal representation of information structure in both homeland and heritage Russian and, more specifically, by probing into the interplay between word order and prosody in this domain. The experiments conducted by Laleko make a good case for the departure from conventional theories about the relation between word order and information structure, according to which givenness and newness are primarily marked by word order across Slavic languages, and not by prosodic means. On the contrary, both homeland speakers and heritage speakers are shown to display no preference for word order over prosody in marking information structure. Moreover, Laleko discusses how both the word-order and the prosodic strategies are used in an innovative fashion by heritage speakers, and how their weight in encoding information structure is different than what is attested among homeland speakers. This change is explained by appealing to the concept of plasticity (Vallduví and Engdahl 1995), which captures cross-linguistic variation in the domain of discourse relations by considering the extent to which languages allow for deviations from default word order and prosodic patterns. Based on this, Laleko ultimately likens heritage languages to other instances of microvariation across Slavic languages, thereby bringing them to the fore in the theory-building effort: in this sense, heritage languages are not just a valid extension to more conventional empirical domains, but they are also argued to be complementary to those, thanks to the insights that they offer into the basic tenets of linguistic structure and into how languages vary and change.

The verbal domain in HL syntax is investigated in the chapters by Alexiadou and Rizou and Putnam. In *Non-active Voice in heritage Greek*, Alexiadou and Rizou investigate the production of non-active Voice in heritage Greek as spoken in Germany and in the United States. While the extant literature found non-active morphology to be vulnerable in heritage populations, due to its more complex and underspecified syntax, this chapter presents novel production data that conflict with those conclusions. More specifically, Alexiadou and Rizou show that the non-active morphology is not vulnerable in certain low-proficiency speakers; rather, these speakers generalize non-active morphology in anticausative verbs and may even generate new transitive deponents, although more restrictedly. The authors argue that the former

observation suggests that non-active Voice is generalized as an intransitivity marker by heritage Greek speakers in the United States and Germany; importantly, the convergence across different contact languages suggests that transfer cannot be a valid explanation for the observed patterns. The latter observation, instead, is taken to align with an analysis of the Greek Voice system whereby structures that contain non-active Voice can be effectively harnessed to generate new meanings (Oikonomou and Alexiadou 2022). While the (over)generalization of inflectional morphology attested in the data is a widely acknowledged feature of heritage morphology, non-active Voice is connected with the formal register in Greek; as such, this investigation contributes to our understanding of the different factors that shape heritage grammars and their relative weights.

Putnam's chapter, *The shape and size of defective domains: Non-finite clauses in Pennsylvania Dutch*, sets off from an exploration of the expression of non-finiteness in Pennsylvania Dutch to make a novel case regarding the notion of Representational Economy in heritage languages (Scontras, Polinsky, and Fuchs 2018). Pennsylvania Dutch has begun to deviate from typical German(ic) syntactic features in defective domains, and in particular with respect to the use of the complementizer *fer* as an infinitival marker (and the concurrent loss of *zu*) and the development of defective clausal gerunds. Putnam traces these innovations back to the fact that Pennsylvania Dutch, consistent with Germanic syntax more generally, still avoids the projection of Spec,TP as the final landing position for subjects; this results in the clausal spine to be stretched in the two domains under investigation. This chapter advances our understanding of the mechanisms that shape the syntax of heritage languages in two innovative ways. First, Putnam argues that the mental representation of German(ic) syntax plays a pivotal role in predicting the outcomes of contact-induced change in heritage languages, setting clear boundaries for syntactic change and transfer. Second, based on these observations, Putnam recasts the concept of Representational Economy, which is commonly assumed as one of the main driving forces behind the shaping of heritage languages, in more nuanced terms: while Representational Economy is typically shown to amount to the shrinking of computational domains, the cases discussed in this chapter point rather to structural stretching driven by the same principle (see Lohndal and Putnam 2024 for a related proposal). Putnam argues in fact that syntax stretching, when functional to the avoidance of representational conflicts across the languages in the dyad, likewise results in more efficient processing.

Finally, Frasson and Lohndal and van Baal address issues related to the DP-domain. Frasson's chapter, *Parallel changes in pronominal clitic systems: A view*

from heritage Romance and Slavic, deals with the syntax of pronominal clitics across heritage Romance (Venetian) and Slavic (Bulgarian). Frasson notes a similar trend across the two varieties, albeit in different domains: the distribution of subject clitics in heritage Venetian is parallel to that of object clitics in heritage Bulgarian. In both varieties, in fact, these clitic pronouns occur in the same positions in which their tonic counterparts may also appear. This is taken to be consistent with an analysis whereby these clitics are arguments, and not highly deficient agreement markers. This behavior is traced back to heritage speakers' avoidance of morphological complexity and, possibly, with their difficulties with long-distance dependencies (in this case: agreement): clitics are agreement heads in the baseline grammars, but get reanalyzed as phrases in the heritage varieties, although they are shown to retain some phonological and morphological clitic properties. Importantly, the syntactic behavior of heritage Venetian subject clitics and heritage Bulgarian object clitics is at odds with the clitic syntax described for the respective homeland varieties; hence, heritage languages call for a revision of Cardinaletti and Starke's (1999) typology of pronominal elements and as such offer a new vantage point from which our understanding of and theorizing about clitic syntax can be furthered.

Lohndal and van Baal delve into the domain of DP-syntax to make the case for the need for nuanced predictions as regards the vulnerability and resilience of grammatical features in heritage languages. In *The DP layer in heritage Norwegian: Vulnerability and nominal architecture*, they take as their departure point Polinsky's (2018) proposal that salience correlates with resilience: in the DP domain, this amounts to predicting that the D-layer will be resilient. This prediction is tested against heritage Norwegian. The data discussed, however, do not seem to support the resilience of the D-layer: here, determiners are vulnerable, as shown by the fact that they can undergo omission in modified definite phrases or that they can be replaced by the English determiner. Lohndal and van Baal analyze these data by proposing a revised version of Polinsky's proposal: perceptual salience and structural salience should be considered disjointly for different syntactic elements. Determiners are structurally salient (high in the DP), but perceptually non-salient (phonologically reduced): in this case, the low perceptual salience supersedes the high structural salience, accounting for the vulnerability of determiners.

In conclusion, as a whole, the chapters in this volume are the source of interesting empirical investigations into new and under-researched domains as well as the catalyst for revisiting traditional theoretical assumptions surrounding HL syntax. Collectively, they point toward the importance of not only continued research into HL syntax involving lesser studied HLs and the dyads they

exist in (Scontras and Putnam 2020), but also the central role that mental representations found in generative approaches to grammar play in elucidating and clarifying these empirical findings.

Acknowledgments

We owe a tremendous thanks to Andrew Hoffman and Mike McLaughlin for proofreading a final version of this introduction. The usual disclaimers apply.

PART I

LINGUISTIC THEORY AND LANGUAGE VARIATION

1

Microcontact and syntactic theory

Roberta D'Alessandro, Luigi Andriani, Alberto Frasson, Manuela Pinto, Luana Sorgini, and Silvia Terenghi

1.1 Setting the ground

Syntactic theory aims to uncover the principles that govern language, and syntax in particular. Like every other theory, syntactic theory makes predictions about what can and cannot be found in languages. For a theory to be successful, it needs to meet at least observational, descriptive, and explanatory adequacy requirements (Chomsky 2000). However, discussions concerning these levels of adequacy often involve considering systems in an "idealized" or "static" state, assuming a monolithic entity known as *grammar*. This grammar is presumed to be unchanging and constant over time and uniform across all speakers of a given language. The flaw in this assumption, deemed incorrect and somewhat naïve, has been scrutinized in numerous articles, including a recent contribution by Rothman et al. (2022), as well as works by Chandra, D'Alessandro, and Putnam (2024). This article contributes additional evidence to highlight the inconsistency of the assumption of a static, unvarying grammar.

Operating under the premise that grammars are invariant in a synchronous state implies that we examine the system at a specific point in time. This is a straightforward approach to pursue when seeking to uncover the governing principles of a system, by examining the interconnections among its elements (which bears a resemblance to Saussure's methodology) and formulating hypotheses regarding how they were established. However, concentrating solely on a single moment in time may lead to misunderstandings and potentially obscure crucial underlying mechanisms at work in grammar. In the not-too-distant past, the diachronic viewpoint was added to the general methodology for generative grammar: observing language change can give very important insights into the laws regulating the interaction of all elements

Roberta D'Alessandro et al., *Microcontact and syntactic theory*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro et al., Oxford University Press.
DOI: 10.1093/9780191987731.003.0002

in a grammar (Battye and Roberts 1995; Pintzuk et al. 2000; Roberts and Roussou 2003; van Kemenade and Vincent 1997; and many others). In contrast, *change in contact* (CIC henceforth) has remained rather confined in generative theoretical studies: while much research has been carried out from an experimental/psycholinguistic viewpoint, the formalization of the laws governing grammars in contact is outstanding. One notable exception is the study of heritage languages (HLs henceforth), for which a recent, but quite robust, community of scholars has been coming together over the last few years. Many of the studies within this community are theoretically informed, but most of them are concerned with only a part of what Chomsky would call critical analysis: "explanatory power, simplicity, learnability, generality, evolvability" (Chomsky in Collins 2021: 5).

The issue of CIC has usually been tackled through quantitative analysis instead, which has yielded a comprehensive inventory of potential outcomes of change arising from language contact. However, there has been relatively limited exploration of the overarching principles governing such changes, and even less effort dedicated to using insights from language change to deduce the principles that govern grammar. This chapter is dedicated precisely to identifying shared foundational principles governing grammar by drawing comparisons between change in diachrony and in contact (CID and CIC, respectively; see Andriani et al. 2022a,b; D'Alessandro 2021). It does so by developing a methodology, which we call *microcontact*, well suited for conducting this comparison since it enables us to assess the outcomes of change that impact linguistic features in contact as well as to trace their development over time. As such, this methodology provides a more fine-grained analysis of CIC, moving far beyond obvious explanations in terms of transfer. Furthermore, an important aim of this study is to enlarge the pool of HLs on the basis of which generalizations can be drawn, in a field that is mostly dominated by studies on heritage Spanish in the United States, a plea which can already be found in Scontras and Putnam (2020). Including understudied minority languages among the HL under investigation not only contributes to their documentation, but allows us to expand the empirical base against which the various generalizations and proposed analyses can be checked.

In this chapter, we will not describe the methodology for data collection, but will focus on the generalizations regarding grammar. For a detailed description of the data elicitation methods, the reader is referred to Andriani et al. (2022a,b).

This chapter is organized as follows. First, we present an overview of the microcontact methodology and what is needed to compare as well as distinguish CIC from CID, also known as endogenous change (Section 1.2). After outlining our methodology, in Section 1.3, we share findings from the project *Microcontact: Language variation and change from the Italian heritage perspective*, carried out at Utrecht University between 2017 and 2022. This project focused on various syntactic phenomena, encompassing both purely grammatical aspects and those situated at the intersection of grammar and discourse. In so doing, we highlight the insights into grammatical theory that the microcontact methodology provides. In Section 1.4, we present data from several phenomena, in contact and in diachrony: demonstratives, auxiliary selection, differential object marking, subject clitics, and null subjects. Section 1.5 highlights the main theoretical contributions of each of our four case studies. Section 1.6 concludes the chapter.

1.2 Change

While change is rather easy to identify in language, the conditions that cause it are much less transparent. A challenge in addressing issues related to change rests in teasing apart change caused by contact from internally caused change. If we observe change after it has happened, and given the number of extralinguistic and accidental factors determining the output of language change, it is almost impossible to ascertain whether the change that we see was determined by an external factor or it happened "from within," for instance because of the reanalysis of some element following a phonological change in the system, but crucially with no intervention from a putative contact variety. One of the aims of the *Microcontact* project was therefore to identify a methodology to be able to determine whether the results of change were determined by contact or whether they were spontaneous. The rest of this section describes this methodology.

1.2.1 Teasing apart the causes of change

Identifying the triggers for language change is a challenging task. One of the possible paths to follow is to try to design some diagnostics, based on the preexisting knowledge of the mechanisms of CID or CIC. Unfortunately, such an attempt is bound to fail, as the object of study of CIC seems to be rather

different than that of CID, and therefore the diagnostics for CIC, in turn, are rather different (and target different issues) than those for CID. This disparity can be attributed to the different traditions in which these studies are cast: CID studies usually focus on identifying the directionality of change and on drawing generalizations about this directionality (from more complex to less complex, from more lexical to more grammatical, from phonologically heavier to phonologically reduced, and so on). One of the major research areas of CID studies is to identify the paths of grammaticalization taken by the elements undergoing change. Until very recently, when socio-historical studies have started to gain a foothold in linguistics (see, for instance, the work by George Walkden and colleagues: Hejná and Walkden 2022; Kauhanen et al. 2023; Walkden and Breitbarth 2019) the sociolinguistic causes of CID have rarely been tackled: CID was almost exclusively tackled from a grammar-internal viewpoint.

CIC studies address a completely distinct set of research questions, primarily concerning the causes of change and the paths along which change unfolds, and the matter of structural simplification, to which we will return below. First, it is obvious that CIC takes place in speakers who master two or more languages. However, some scholars focus their attention on the grammatical systems in contact while others focus on the speakers themselves and what happens to their mental representations, as well as the cognitive processes involved in bi- or multilingualism. Within the realm of examining grammatical systems in contact, scholars are divided into two main groups: (i) a minority focuses on studying grammars in contact in a manner akin to the analysis of diachronic linguistics and (ii) those who subscribe to majority approaches that view language systems in contact as an extension of social groups that are in contact.

These differences notwithstanding, all scholars agree that the causes of CIC are not exclusively grammatical: grammar-external factors have a prominent role in this domain. Among these, there are purely social factors, such as the attitude of the speaker, the extension (partial, complete), or the nature of contact (balanced, displacive). Whether we are talking of a scenario akin to Basque in Spain or similar to Spanish in the United States, the speaker's perception of the languages they speak plays a pivotal role in either facilitating or impeding linguistic change. These social factors are usually investigated by sociolinguists as well as typologists.[1] In addition to these, many other factors play a role in

[1] For a full overview of factors impacting CIC, see Aikhenvald and Dixon (2006) and Hickey (2020).

CIC: these include the degree of mastery of the contact languages, the age of onset of bilingualism, the situation in which bilingualism arises, the complete or incomplete acquisition of the other language (Meisel 2011, 2013; Montrul 2008; Müller 2017; Paradis 2011; Paradis and Genesee 1996; Tsimpli 2014; Unsworth 2013, 2016; Unsworth et al. 2014; and many others).

The aim of the *Microcontact* project was slightly different. While acknowledging the relevance of all the factors just listed, the focus was on understanding which role grammar plays in determining the direction of change. One way to do that is by comparing CIC and CID to check whether there are underlying cognitive principles influencing change. To identify them, we need a tool that helps us ascertain whether the witnessed change is due to internal or external causes, and whether it is spontaneous or contact-induced, as stated above. The initial significant challenge is, therefore, to pinpoint appropriate diagnostic tools that are applicable in both contexts: diachronic and contact scenarios. As noted above, building upon existing literature is challenging, due to factors such as limited knowledge regarding speaker attitudes and the circumstances of language acquisition and transfer in diachronic change.

Another significant aspect to be taken into account is that, when the languages in contact belong to the same family, it becomes exceedingly difficult to ascertain whether a specific change is a consequence of contact, or it occurred spontaneously. For instance, we know that the Romance languages tend to overextend the auxiliary HAVE to cover the verb classes selecting BE (Ledgeway 2012; Tuttle 1986; a.o.). If we witness the extension of HAVE in contact across Romance varieties, we cannot conclusively attribute this change to contact more than we can attribute it to a spontaneous trend in the family. To quote Aikhenvald (2006: 9),

> if languages are genetically related, we expect them to develop similar structures, no matter whether they are in contact or not. And if genetically related languages are in contact, trying to prove that a shared feature is contact-induced and not a chance result of Sapir's drift may be next to impossible.

This is certainly true if we conceptualize contact as involving languages dyads (Scontras and Putnam 2020). The whole narrative changes, though, if we take a cross-comparative perspective, whereby one language is checked in contact with many others. The *microcontact* methodology is described as follows:

(1) **Microcontact methodology**

"A language A with a feature X should be studied in contact with a set of languages: B, C, D, and F, which all have an identical feature X and are almost equivalent grammatically BUT for the features Y and Z, strictly related to X. If the feature X in [A|B] (read, A in contact with B) undergoes exactly the same change as [A|C], and [A|D] and [A|F], despite Y and Z are different, we can conclude that the change of X is spontaneous, and that B, C, D, and F did not have any impact on A. If, instead, for instance XB (X in language B) and XD are identical, while XC and XF are completely different, and we see that XA changes in the direction of XB = XD but not of XC or XF, we can conclude that this change is contact-induced, and that possibly Y and Z have a role in this change (depending on how similar they are in B, D)" (D'Alessandro 2015: 18–19).

This original formulation can be broken down into several parts. As stated above, to ascertain whether change is due to contact or it happens spontaneously, the best path is to compare the output of contact with the spontaneous evolution of the element under examination. This means that we need to identify a feature F in one language and follow its evolution in diachrony and compare it to its evolution in contact. Take, for example, differential object marking (DOM henceforth), exemplified in (2) for eastern Abruzzese, an upper-southern Italo-Romance variety. DOM in Abruzzese is restricted to first- and second-person objects:

(2) Eastern Abruzzese, Arielli (D'Alessandro 2017: 8)

a. So vistə a mme/ a tte.
be.1SG seen DOM me DOM you
'I saw me/you.'

b. Semə vistə a nnu/ a vvu.
be.1PL seen DOM us DOM you
'We saw us/you.'

c. *So vistə a Marije/ a jissə/ a quillə.
be.1SG seen DOM Mary DOM them DOM them

available. As shown in Ledgeway (2009), DOM in Neapolitan emerged first in object left-dislocated structures, as exemplified in (3), while it was much less frequent on animate/definite objects in situ (as in (4)).

(3) Neapolitan, 14th c. *Romanzo di Francia* (Ledgeway 2009: 792)

E	a	mene	me	'de	volleva mandare
and	DOM	me.1SG	me	hence	wanted to.send

in outramare co' Ramirante
in overseas with Ramirante

'And he wanted to send me overseas with Ramirante'

(4) Neapolitan, 14th c. *Romanzo di Francia* (Ledgeway 2009: 834)

Vóy	fare	morire	mene
want.2SG	make	die	me.1SG

'You want to make me die'

Under the rather uncontroversial assumption that eastern Abruzzese DOM has undergone the same developmental trajectory as Neapolitan,[2] we can conclude that DOM started in topicalized, or clitic left-dislocated position, and gradually extended to the objects in situ.

Going back to the *microcontact* methodology, we need to ascertain whether that is also the case in contact. However, as repeatedly pointed out, not every contact situation is equal, and in particular looking at dyads in contact will not take us too far, as the result of contact could be accidentally similar or dissimilar to that of CID because of the particular language pair checked. The *microcontact* methodology requires therefore a cross-check between different dyads: one phenomenon needs to be observed in Language A in contact with many other languages. Crucially, these languages need to be syntactically identical (for a given domain) to A except for that phenomenon. The easiest way is to compare languages of the same family, which are structurally very similar but may vary at featural level. For the case of DOM just illustrated, ideally DOM in Eastern Abruzzese needs to be checked in contact with Spanish (which does have DOM), French (which does not have it), Brazilian Portuguese (which does not have it), and so on. If a uniform change is observed, for instance the extension of DOM starting from topic position, then it is licit to conclude that CIC and CID follow the same path, and specifically

[2] This path of emergence of DOM has also been extensively documented by Iemmolo (2009, 2010) for over 100 languages worldwide.

that change has happened spontaneously and independently of the specific dyad involved. If, on the contrary, the output of change moves in different directions for different dyads, we can conclude that change is due to contact, and it is not spontaneous. We will return to DOM in Section 1.4.4 with the results of our inquiry.

Not many languages and language families offer this wealth of contact situations, as well as a thoroughly documented history which allows us to follow the spontaneous development of a feature X according to the *microcontact* methodology. An important exception, as just observed, is the Romance language group, and in particular the Italo-Romance subgroup. Many Italo-Romance languages have a long literary and written tradition. Some of them, like Sicilian, Neapolitan, or Venetan, were used for centuries in little kingdoms or republics in Italy. These languages were never completely isolated, though their extensive contact with superstratal Italian took place systematically and penetrated all social strata only after World War II. This means that all historical records allow us to observe the relevant phenomena in their diachronic evolution, from the Proto-Romance phase to the 1950s–1960s.

The languages chosen for the project were variants of the Romance languages in the Americas.[3] It is well known that Italy has a long history of emigration to both North and Latin America, which happened in waves, starting in the 19th century. The target group was that of speakers that emigrated around the 1950s–1960s and their descendants. This choice was made with the aspiration to track the first-generation emigrants. This, in turn, would enable us to verify the input that heritage speakers of subsequent generations were exposed to, attrited though they might be.

In the 1950s–1960s, a significant population of Italian emigrants, who were primarily monolingual speakers of various Italo-Romance dialects, found themselves in contact with other Romance languages in a situation of microcontact like that described in (1) above. This created the ideal conditions to check the configuration of multiple contact of features. The *Microcontact* project was designed to factor out as much inter-speaker and intra-speaker variation. First, the conditions in which the first-generation speakers that we interviewed entered in contact with the host languages are rather similar: they all left Italy with the intention of relocating to the new country; the host country was distant enough from their homeland as not to allow for frequent contacts with it. Furthermore, the speakers' profiles are rather homogeneous: we selected first-generation speakers with little education and

[3] For the rationale behind this choice the reader is referred to Andriani et al. (2022a).

scarce mastery of Italian and with a comparable socio-economic profile. While for first-generation speakers it was somewhat possible to control for the uniformity of these extra-linguistic factors, this was not always feasible for heritage speakers of subsequent generations, whose profiles varied from holding no school education to university degrees, and whose exposure to their HL was quite heterogeneous. In fact, the classification into "generations" is itself inaccurate, as in many cases third-generation heritage speakers learnt their HL directly from their grandparents, while the generation of their parents was skipped altogether (Andriani et al. 2022b). In general, while we are aware that the variation in the speaker profile as well as many external factors have a huge impact on the outcome of contact, we also could see clear tendencies and identify some generalizations that hold across the board, as those that we are going to discuss in the next sections.

1.3 Generalizations on change in contact

A primary objective of the *Microcontact* project was to enhance our comprehension of some syntactic phenomena by examining them within the context of language contact. Isolating a feature and controlling its development in contact *and* in diachrony provide some important insights into its nature, into the elements directly interacting with it, and ultimately into the syntactic structure. The next sections are dedicated to presenting some case studies of syntactic microcontact, specifically auxiliary selection, demonstratives, null subjects, and DOM.

Before reviewing those case studies, it is necessary to introduce some of the generalizations that have been derived by comparing outcomes across various linguistic phenomena. Building on the results of our research and on its comparison to CIC, the following generalizations can be made:

G1. Elements react to external solicitations of the grammatical system (i.e., to contact) in different ways depending on their nature. Specifically, purely functional elements behave differently than discourse-related ones (see also Sorace 2011 for similar results in bilingual speakers).

G2. Sequences of functional features change in predictable and consistent fashion. Specifically, if they are core grammatical features (such as φ-features) they will follow a specific direction of change, both in diachrony and in contact. This direction is determined by what Terenghi (2021, 2023a) calls the *monotonicity bias*.

G3. Elements that are discourse related, like null subjects or DOM, follow two different paths, possibly depending on the possible mapping of the structure of L1 on L2 and in turn depending on the perception of the locus of variation by the speakers. In this respect, see the concepts of grammatical similarity (Schwartz and Sprouse 1996), perceived typological similarity as in the Typological Primacy Model (TPM) (Rothman 2010, 2011, 2013a,b; Rothman and Cabrelli Amaro 2010).

Some additional background to this latter generalization, and particularly to the issue of structural mapping, is further discussed in the remainder on this section.

1.3.1 Discourse-related elements – structural mapping

That the acquisition of some language features may be facilitated by the other languages present in the speaker's mind is a rather uncontroversial assumption. To begin with, the CEM (Cumulative Enhancement Model; Flynn et al. 2004) states that those elements present in any of the languages spoken by an L3 learner that can facilitate acquisition will be used to build the L3 grammar. Slightly differing from that, perceived typological similarity, as conceived by Rothman in his Typological Primacy Model (Rothman 2010, 2011, 2013a,b), is a key factor in linguistic transfer. Specifically, like for the CEM, the order of acquisition of L1 and L2 will be totally irrelevant for transfer. However, because of principles of economy of language acquisition, underlying structural (and therefore typological) similarity between one of the languages and the L3 will condition the acquisition, and not only facilitate it. The crucial point is that transfer of the entire system is required here, including those features that are dissimilar in the languages. In a well-known study, Rothman (2010) examined bilingual Spanish-English (in either order) Brazilian Portuguese (BP) learners, observing how, even if for some specific structures, like null subjects, BP is closer to English than to Spanish, Spanish structure was massively transferred into L3 BP. Rothman takes therefore a holistic view of transfer, with further refinement. From a microvariation viewpoint, Westergaard (2009a,b, 2013) takes the opposite view, showing that acquisition happens via micro-cues, and therefore transfer is a side-effect of stepwise acquisition (and transfer, in the case of bilinguals) of the various featural values.

From a strictly structural viewpoint, it has been shown that structural similarity facilitates transfer in contact situations, both when typologically similar languages, like Spanish and Catalan, are involved and when typologically divergent languages, like Asháninka and Spanish, are involved (Mayer and Sánchez 2021). The study by Mayer and Sánchez (2021) is particularly interesting in that it shows that in L2 Spanish spoken in Peru by Asháninka L1 speakers DOM is retained, while in L2 Spanish spoken by Shipibo L1 speakers in contiguous regions it is lost. Meyer and Sánchez attribute this different outcome to structural similarity in a wide sense: Asháninka is a nominative-accusative language with SVO order, while Shipibo is an ergative language with SOV order. Structural overlap allows DOM to be retained, while structural mismatch (within the VP domain) causes its disappearance.

Our study shows that something similar is indeed at work, but we move one step further along this line of reasoning. When two languages are in contact, in dyadic contact situations, speakers can usually map given parts of the structure of one language onto the other. If the languages are typologically or structurally similar, like in the case of Asháninka and Spanish, the speakers will have no trouble mapping for instance the VP of one language onto that of the other language. This will result in some sort of "perception" of the locus of variation, meaning that the speaker will be able to perceive the difference and adapt their structures, mostly along the line proposed by Sorace in her Interface Hypothesis. In the case of Asháninka and Spanish, DOM will be retained as the VP structures of the two languages are similar, despite the difference in vocabulary and phono-morphology. Hence, syntactic similarity plays a role, and in fact a much more central role than previously acknowledged. In the case of Shipibo and Spanish, instead, the speakers will identify the locus of variation, perceive the difference, "understand" that the VP in one language is organized differently from the other, and choose for simplification of the structure by dropping the DOM, possibly because of the greater structural difference between the languages. In any event, in the case of dyads in contact, language change will *not* be due to the inability to perceive or track the variation point, but to a grammar-internal resolution mechanism. The perceived typological similarity will be active in this context, as the speaker's mental representation of the two grammars is sufficiently distinct but also identifiable.

In general, if speakers unconsciously perceive the locus of variation, in the sense that the grammars, for the relevant domain, can be easily mapped onto each other, the exact point of variation within a given domain can be identified. Importantly, though, 1:1 mapping of L1 to L2 (in a given domain) allows

the speaker to identify the complexity or redundancy in their own language, and very often reduce it. Note that we are not talking here about simultaneous bilinguals.

The proposal here is not that, whenever direct mapping is possible, simplification (i.e., choice of a system with "fewer points of choice") will ensue, but that it *might* ensue, and in fact that it is most likely to ensue. Conversely, when 1:1 mapping is not possible, the outcome of contact is much more unpredictable.

1.3.2 Discourse-related elements: No mapping possible

When heritage and L2 speakers cannot pinpoint the locus of variation, the conflict between two possible grammatical systems (or options) needs to be resolved in a different manner. This situation occurs, for instance, when the grammars in contact are two dialects of the same language, or two languages of the same family that are structurally typologically identical. In this case, both grammatical similarity (in the sense of Mayer and Sánchez 2021) and perceived typological similarity (in the sense of Rothman 2010 *seqq.*) are very high, too high in fact, up to the point that the speakers are no longer able to identify the point of variation, which will consist, for instance, in the value of a specific feature rather than the presence of the feature itself in the functional spine. This is crucially different from the case of Asháninka-Spanish contact, where the speakers can map one VP onto another, but they still have the clear perception that they are speaking two separate languages.

An example of the impossibility in identifying the locus of variation is a speaker of Abruzzese (a language with DOM only on first- and second-person pronouns, as seen in (2)) in contact with Rioplatense Spanish (a language with definiteness and animacy-driven DOM): this speaker will not be able to perceive the variation point as that is too minimal. Variation resides indeed in the value of a feature; compare (2) to the Rioplatense Spanish counterpart:

(5) Rioplatense Spanish

a. Me vi a mí mismo
me.1SG.CL saw.1SG DOM me.1SG self
'I saw myself'

b. Te vi a vos
you.2SG.CL saw.1SG DOM you.2SG
'I saw you'

c. los vi a ustedes
you.2PL.CL saw.1SG DOM you.2PL
'I saw you'

d. (la) vi a María/ a la chica
her.3SG.CL saw.1SG DOM María DOM the girl
'I saw María/the girl'

The structure of the XP that hosts the DOM marker (the VP) is essentially the same in both languages; structurally identical languages can be easily mapped onto each other, leaving the impression of a non-existent structural difference between them. In other words, functional head by functional head mapping yields a perfect one-to-one mapping. Yet, the feature values represented by the functional heads differ. In this case, the TPM requirement, as well as the CEM requirements, will be vacuously met, but no structural resolution of the difference is envisaged. We maintain that in this case the speakers will have to resort to some basic cognitive strategies, and "reset" the phenomenon along the lines of its spontaneous emergence, as in L1 acquisition.

An excellent example of this is the DOM-cycle described in D'Alessandro (2022a): while languages with DOM in macrocontact tend to lose DOM on in situ objects, languages in microcontact overmark DOM on dislocated elements, as also illustrated in Section 1.4.4. This difference in the strategy of DOM resolution reflects the generalization just exposed: HL speakers in macrocontact situations resort to a different strategy compared to HL speakers in microcontact. Specifically, the locus of variation (in this case, the VP) is easily identifiable through structural mapping (obviously, unconsciously) and the system moves toward simplification; in the latter case, instead, the locus of variation does not easily emerge from mapping, and the resulting outputs are therefore significantly different from those attested in macrocontact.[4] Crucially, not only is DOM emergent in microcontact, but it also emerges in the same contexts in which it is shown to emerge diachronically, namely in object dislocation constructions (see Andriani et al. 2022b; Ciconte 2018a,b; D'Alessandro 2022a; Iemmolo 2010; Sorgini 2022). The strategy of marking dislocated topics is acquired early in L1 acquisition (Belletti 2017, 2018; Belletti and Guasti 2015) and, we maintain, can be used by HL speakers to track long-distance dependencies (see D'Alessandro 2022b). We will not go into specific details concerning this marking strategy here, which would take us

[4] This is a simplification, obviously, which does not take the differences between speakers into account.

too far afield. What matters in this context is that while DOM is observed to emerge in microcontact, it seems to be lost or impoverished (with some exceptions, as we saw) in macrocontact. We attribute this different behavior to the impossibility of speakers in microcontact to identify the locus of variation, the difference between the structures of the two languages in contact. Speakers in microcontact, not being able to identify the structural locus of change, thus resort to some "default" mental representation of the construction at hand, or to what UG makes available, in the sense of Bickerton (1989, 1990).

Microcontact is not the only situation in which the speakers cannot easily map L1 structures onto L2 structures to be able to unconsciously identify structural differences between the two systems. The same difficulty arises in creole languages, where more than two languages, usually structurally very different, enter in contact suddenly and simultaneously. Little mapping between the constituents in the contact languages is possible, so the speakers (just like in the case of microcontact) do not resort to grammatical strategies for simplification of the same kind as those found in macrocontact (Aboh 2017; Aboh and DeGraff 2017; McWhorter 2018; Mufwene 2020; and many others). Creole languages present instead phenomena that are similar to those which arise in microcontact and in L1 acquisition, such as the marking of dislocated elements (D'Alessandro 2021a,b). At the moment, these cannot be presented as more than preliminary observations, based on the tendencies we observed in microcontact compared to macrocontact. We leave this therefore for further research.

Even if our interpretation of these empirical observations were entirely flawed, we could still safely claim that the study of *microcontact* offers an important contribution to the theory of language, and to syntax in particular, as it makes it evident that generalizations based on contact often overlook the different kinds of contact from a strictly structural viewpoint. In fact, no general conclusions can be drawn from the fact that some phenomena (such as DOM or null subjects) weaken as to the conditions that determine this disappearance. Even if the results are similar, the causes underlying these outcomes may be different.

These generalizations regarding the mapping are only relevant to discourse-related elements. As per G2 (see again Section 1.3 above), whenever fixed sequences of functional heads are involved, like in the case of pronouns and demonstratives, structural mapping is irrelevant and other factors are instead at play. We will return to this in Sections 1.4.1 and 1.4.2.

1.4 Four case studies

In what follows, we present four case studies on microcontact, showcasing how it enriches our understanding of syntactic theory. The first case is that of auxiliary selection in heritage Italo-Romance. This study shows that varieties with auxiliary selection tend to reduce the auxiliaries to only one, and that the direction of CIC is exactly the same as that followed by CID (Section 1.4.1). A similar behavior is found in demonstratives, where larger demonstrative systems (three contrastive forms) likewise reduce to smaller demonstrative systems (two contrastive forms), in (micro)contact and diachrony alike (Section 1.4.2). These facts differ starkly from those uncovered by the study of syntactic phenomena that also strongly rely on discourse (hence involving at least two different grammatical modules), as clearly shown by our studies on null subjects and DOM. The null subject data demonstrate that discourse features hinder our predictive power when it comes to the development of pronominal subjects in contact contexts (Section 1.4.3). Finally, DOM in microcontact is shown to behave very differently from DOM in contact, with L1 strategies at work for the resolution of microcontact (Section 1.4.4).

Table 1.1 reports the number of speakers interviewed (with G referring to generations). In the table, it is assumed that the speakers in Argentina are in contact with Spanish, those in Brazil with Brazilian Portuguese, and so on.

1.4.1 Auxiliary selection in heritage Venetan

In the Romance literature (cf. Loporcaro 2016 and references therein), auxiliary selection has received much attention due to the variety of patterns that it displays and to their implications for our understanding of agreement phenomena involving verbal arguments. Apart from the synthetic form of the lexical verb (moved from within the VP; cf. Pollock 1989), Romance languages typically lexicalize T using the auxiliaries BE and HAVE. Although there is no universally consistent correspondence between the choice of the auxiliary and the verbal class (see D'Alessandro and Roberts 2010), transitive and unergative verbs usually select HAVE while unaccusatives and passives usually select BE (but see Sorace 1995 *ff.* for a more detailed description). The different mechanisms of auxiliary selection can provide us with valuable insights into the TP domain.

Table 1.1 Number of speakers interviewed per location and heritage variety

	Argentina	**Brazil**	**Quebec**	**NY**	**(Belgium)**	**Total number of speakers**	**Total minutes**
ABRUZZESE	9 speakers • 7 G1 • 2 HS	–	2 • 2 GI	3 • 1 G1 • 2 HS	1 • 1 G1	15 • 11 G1 • 4 HS	> 150' • 110' G1 • 40' HS
CALABRIAN	9 • 8 G1 • 1 HS	3 • 2 G1 • 1 HS	5 • 5 G1	–	–	17 • 15 G1 • 2 HS	>170' • 150' G1 • 20' HS
FRIULIAN	8 • 4 G1 • 4 HS	7 • 7 HS	7 • 7 G1	6 • 4 G1 • 2 HS	–	28 • 15 G1 • 13 HS	>280' • 150' G1 • 130' HS
SICILIAN	13 • 10 G1 • 3 HS	2 • 2 G1	5 • 4 G1 • 1 HS	10 • 8 G1 • 2 HS	5 • 4 G1 • 1 HS	35 • 28 G1 • 7 HS	>350' • 280' G1 • 70' HS
TRENTINO	3 • 3 G1	7 • 7 HS	–	–	–	10 • 3 G1 • 7 HS	>100' • 30' G1 • 70' HS

VENETAN	16 • 12 G1 • 4 HS	28 • 2 G1 • 26 HS	8 • 7 G1 • 1 HS	–	–	52 • 21 G1 • 31 HS	>520’ • 210’ G1 • 310’ HS
OTHER	15 • 14 G1 • 1 HS	3 • 1 G1 • 2 HS	9 • 9 G1	39 • 19 G1 • 20 HS	2 • 1 G1 • 1 HS	68 • 44 G1 • 24 HS	>680’ • 440’ G1 • 240’ HS
TOTAL	73 • 58 G1 • 15 HS	50 • 7 G1 • 43 HS	36 • 34 G1 • 2 HS	58 • 32 G1 • 26 HS	8 • 6 G1 • 2 HS	225 • 137 G1 • 88 HS	–
TOTAL MINUTES	>730’ • 580’ G1 • 150’ HS	>500’ • 70’ G1 • 430’ HS	>360’ • 340’ G1 • 20’ HS	>580’ • 320’ G1 • 260’ HS	>80’ • 60’ G1 • 20’ HS	–	>2250’ • 1370’ G1 • 880’ HS

Note: Unless specifically indicated otherwise, all the data presented in this chapter pertain to the speakers who were interviewed during fieldwork and are reported in this table.

Source: Andriani et al. (2022b: 14)

In heritage Italo-Romance, we detected some incipient changes affecting auxiliary selection, which seem to point in the direction of a *gradual simplification*, i.e. impoverishment of auxiliary selection and generalization of one auxiliary in certain contexts. This is reminiscent of the CID outcomes mentioned in Section 1.2.1, although we should mention that the direction of change does not always completely mirror the one observed in diachrony. The actual situation is rather that of optionality: speakers start extending the auxiliary HAVE to the verbs that usually select BE; in other words, HAVE is sometimes found in contexts in which the baseline features BE. This means in no way that the change is complete, and that BE has disappeared from the heritage Venetian grammar though (see D'Alessandro and Frasson 2023 for details on the selection and specialization of BE).

Among the varieties in our corpus, the only HL which offered a set of data fully suitable for comparison with the homeland counterpart is that spoken by the Venetian community in Argentina and Brazil. Despite the heavy microvariation, auxiliary selection in homeland Venetian varieties mainly follows the typical split-intransitive pattern (cf. Cordin 2021: §2.3.3), which is as follows:[5]

- BE with unaccusatives → [+ φ-Agr] with S_O;
- HAVE with transitive/unergatives → [– φ-Agr] with A/S_A; [+(/–) φ-Agr] with O

Crucially, and unlike Venetan, Argentinian Spanish and Brazilian Portuguese have generalized the synthetic past form; analytic past forms underwent instead specializations for other (language-specific) temporal and aspectual functions. Prior to these specializations, Spanish and Portuguese displayed the generalization of auxiliary HAVE (Spanish *haber*, Portuguese *ter*) to all contexts but passives.

A parallel expansion of auxiliary HAVE was detected for heritage Venetan in Brazil by Faggion (2013) and Dal Picol (2013, 2014);[6] importantly

[5] However, Venetan varieties do present a few lexicalized instances of HAVE-selecting unaccusative verbs, e.g. ***ga*** *calà* '(s)he's gone down' (cf. Cordin 1997: 93–95). Moreover, these varieties additionally show person-driven auxiliary splits for reflexives; this situation cannot be discussed here for reasons of space. For further details on the heritage Venetan, see D'Alessandro and Frasson (2023); for details on baseline Venetan, the reader is referred to Lepschy (1984); Belloni (1991); Benincà (1994); Loporcaro and Vigolo (1995); Marcato and Ursini (1998); Manzini and Savoia (2005); Cennamo and Sorace (2007); Loporcaro (2007).

[6] *Vêneto sul-rio-grandense* (Stawinski 1987; cf. also Cordin 2021: §1.3, Frosi and Mioranza 1983), the Venetan koine spoken in and around Caxias. The koine has been argued to be based on central Venetan varieties, with occasional Trentino and Lombard elements. In general, cross-dialectal leveling seem to have occurred in many of the areas considered.

for our purposes, these scholars regard the expansion of HAVE as the consequence of contact with Brazilian Portuguese. Faggion (2013: 140) claims that younger speakers (born around 1970) transfer the BP auxiliary *ter* into their heritage Venetan and consequently start alternative *gavèr* (HAVE) with BE with unaccusatives (6a) and reflexive predicates (6):

(6) Heritage Venetan in Brazil, young generation speakers (Faggion 2013: 140)

a. mi **go** **ndato** a Garibaldi
I have.1SG gone.M to Garibaldi
'I went to Garibaldi (town).'

b. **me** **go** **desmentegà**[7]
me have.1SG forgotten
'I forgot.'

In contrast, older speakers retain *esser* ('BE') in the same contexts, while systematizing the different auxiliary with incorporated clitic forms in a different fashion with respect to the baseline (D'Alessandro and Frasson 2023): while baseline Venetan features three separate BE + subject clitic auxiliaries that are diatopically distributed, heritage Venetan in Brazil has developed a very articulated system whereby the choice of the BE auxiliary depends on the position of the subject and the agreement between the past participle and the object:

(7) Heritage Venetan (D'Alessandro and Frasson 2023: 14)

a. **L'è** rivà i bisnon-i
is arrived-M.SG the-M.PL great-grandparents-M.PL
qua tal Brasil.
here to.the Brazil
'The great-grandparents arrived here in Brazil.'

b. La so mare **ze** nasest-a in Italia.
the his mother-F.SG is born-F.SG in Italy
'His mother was born in Italy.'

c. I non-i **è** vegnest-i de navio.
the grandparents-M.PL be.3 come-M.PL by ship
'The grandparents came by ship.'

[7] Notice that this form is also attested in several varieties of Venetan (Benincà 1994; and data from ASIt, http://asit.maldura.unipd.it).

Heritage Venetan in Brazil has organized BE auxiliaries according to a system that is not found in the baseline spoken in Italy (though it is reported for some older varieties). While *l'è* is selected with postverbal subjects and non-agreeing past participle, *ze* is selected in sentences where the subject is preverbal. Furthermore, *è* is almost exclusively selected with preverbal masculine plural subjects. This system shows an incredible complexity: speakers are sensitive to subject position and agreement patterns, showing that handling syntactic complexity and long-distance agreement is not really an issue for them.

To sum up, in Brazilian Venetan we see at least two developments with respect to auxiliary selection. Although BE is still very much present in the auxiliary system of heritage Venetan and BE/HAVE alternations may be found, some of the tendencies we observe evidently point to an ongoing change in the auxiliary selection patterns in Brazilian (and Argentinian) Venetan. In these varieties, HAVE has started to gradually take over some of the slots reserved for BE (with the relevant agreement patterns) and, because of this extension, the BE/HAVE auxiliary selection is getting confined to fewer contexts. At the same time, the BE auxiliary distribution is also diversifying.

Importantly, this incipient shift toward HAVE seems to coincide with what we observe in diachrony, i.e. that many languages (e.g., Spanish and Portuguese, as well as English) lost auxiliary alternation and generalized HAVE, a less computationally costly option. A pure transfer account (i.e., importing the pattern found in the contact language, as maintained by Faggion 2013 and Cordin and Degasperi 2020 for Trentino) could only be invoked for the case of Argentina, where the exposure to the (specialized use of the) HAVE + past participle construction in the local Spanish is higher compared to that of Brazilian Venetans to the *ter*-past participle construction. This lack of exact correspondence between structures in the input and outcome of change in the HL leads us to infer that this slight restructuring of the auxiliary selection patterns in heritage Venetan is not directly triggered by the presence of a structurally similar construction in the dominant language. However, some sort of reduction is at work, which is hardly quantifiable at this stage because of lack of data.

This state of affairs falls under G2: auxiliaries do not interact with discourse, and can be considered core syntax. Their evolution is therefore predictable, and independent of the contact language. A similar behavior is found in demonstratives, which also follow a predetermined path that is the same in contact and in diachrony. The next section will expand on this topic.

1.4.2 Demonstratives in microcontact

Demonstrative systems constitute another domain in which a reduction from larger to smaller systems, like the one just reviewed in Section 1.4.1, is attested; this reduction proceeds in one and the same fashion in (micro)contact and diachrony alike. This case study focuses on the demonstrative forms in heritage southern Italo-Romance varieties spoken in microcontact.

Demonstrative forms, such as *this* and *there*, are used exophorically if they denote the location of a referent in the external world with respect to a deictic center, which serves as a point of reference (see Diessel 1999; Lyons 1977; a.o.). Languages exhibit significant diversity when it comes to the range of deictic centers they employ (Diessel 1999; Levinson 2018; a.o.), as exemplified by the Romance varieties in (8):

(8) a. Italian

questo	quello
'this near me'	'that far from me'

b. Brazilian Portuguese

esse	aquele
'this near me and/or you'	'that far from us'

c. Spanish (some varieties)

este	ese	aquel
'this near me'	'that near you'	'that far from us'

Cross-linguistically, the most common deictic center is the speaker (Diessel 1999: Chapter 3), which yields demonstrative systems as the one in (8a). These systems include two forms that stand in a two-way speaker-oriented opposition: near the speaker as opposed to far from the speaker (speaker-based binary systems). Other possible deictic centers include: the participants (8b), such that referents are defined as near the participants (i.e., the speaker and/or the hearer) as opposed to far from the participants (participant-based binary systems); and the hearer, besides the speaker, whereby a three-way deictic contrast is yielded among referents near the speaker, referents near the hearer, and referents far from both (ternary systems, (8c)).

Our study focused on heritage Italo-Romance varieties whose homeland counterparts are of the latter type (for further details, see Terenghi 2022, 2023a). More concretely, we considered heritage Sicilian (n = 7), heritage Abruzzese-Molisano (n = 4), and heritage Calabrian (n = 1) speakers in

different immigration countries, that is, in contact with different majority languages: in microcontact with Argentinian Spanish ($n = 6$) and with (Quebecois and Belgian) French ($n = 2$). Furthermore, we investigated contact with US English ($n = 6$), included as a control group to test whether microcontact follows the same change mechanisms as other contact types in this domain.

While standard Argentinian Spanish has a ternary system similar to the one displayed by the southern Italo-Romance varieties under investigation (see again (8c)), its informal, spoken variety employs a speaker-based binary system as the one in (8a); the system in (8a) is also in use in French and in English.[8] Given these differences across heritage and majority languages, our purpose was to assess whether the encoding of deictic features in the demonstrative systems of the former undergoes CIC, i.e. whether the organization of the demonstrative systems of the HLs spoken in microcontact undergoes any change driven by the organization of the demonstrative systems of the relevant majority languages.

The results of our research show that ternary systems in microcontact and macrocontact alike tend to undergo a reduction toward a binary system: this is clearly shown by the fact that, while the speaker-oriented and the non-participant-oriented domains systematically get each its own exponent (as in (8c)), the hearer-oriented deictic domain has its dedicated (target) exponent only in roughly half of the elicited data points. This is shown in Table 1.2, where microcontact and macrocontact data from all three varieties are reported jointly, as no significant differences were detected across them.

The high percentage of 'this near me'- and 'that far from us'-type demonstratives in the hearer-oriented domain suggests that erstwhile ternary systems are moving toward binary systems in the heritage southern Italo-Romance varieties under investigation. Specifically, 'this near me'-type forms are compatible with a participant-based binary demonstrative systems (as in (8b)), while 'that far from us'-type forms are compatible with a speaker-based binary demonstrative systems (as in (8a)). It should however be noted that, despite the at face-value clear-cut results, our informants displayed a significant level of intra-speaker variation (for its full extent, see Terenghi 2023a), suggesting that change is still in progress in these varieties.

Importantly, our data consistently suggest that the attested incipient change is not driven by the contact varieties also in this case, despite the surface similarity between the reduced heritage systems and the contact systems: in fact,

[8] Note that, strictly speaking, French has an even smaller inventory of demonstrative forms, with the only demonstrative form *ce* 'this/that' which combines with locative adverbs *-ci* 'here' and *-là* 'there' to disambiguate between the proximal interpretation (*ce … -ci* 'this') and the distal one (*ce … -là* 'that').

Table 1.2 Speaker-, hearer-, and non-participant-oriented deictic domains in heritage southern Italo-Romance

	'This near me'-type	*'That near you'-type*	*'That far from us'-type*
Speaker-oriented domain	53/53 (100%)	–	–
Hearer-oriented domain[*]	13/53 (24.53%)	26/53 (49.01%)	10/53 (18.87%)
Non-participant-oriented domain	6/53 (11.32%)[**]	2/53 (3.77%)	45/53 (84.91%)

[*] Furthermore, 4/53 (7.55%) elicited answers show optionality between the 'this near me'-type and the 'that far from us'-type.

[**] Note that the high percentage of non-target-like responses in this case is an effect of the picture-sentence matching task, as the non-target-like 'this near me' forms decrease to just 1/23 (4.35%) when the production task is considered. For a discussion of this task effect, see Andriani et al. (2022a). Importantly, this effect is not prominent for the hearer-oriented domain, where the percentage of 'this near me' forms in fact increases to 8/30 (26.66%).

Shaded cells represent non-target-like answers (e.g., 'this near me'-type used in the hearer-oriented domain, which would require a 'that near you'-type under (8c)).

non-target 'this near me' forms are not expected under any of the contact situations and yet are the most elicited forms; likewise, target 'that near you' forms are attested in contact with Spanish (as expected), but even more so in contact with French in Belgium and with English in the United States (contrary to expectations). Thus, the patterns of reorganization in contact cannot be explained as the result of transfer from the dominant language.

Instead, the simplification of demonstrative systems attested in heritage Italo-Romance varieties spoken in (micro)contact is consistent with the wider tendency toward the reduction of demonstrative systems, as also attested in diachrony. As shown by Terenghi (2023a,b) on the basis of Romance data collected by Ledgeway and Smith (2016), ternary demonstrative systems in diachrony tend to lose the dedicated 'that near you'-type exponent for the hearer-oriented domain, and hence to turn into speaker-based or participant-based binary demonstrative systems, as well. This strongly implies that these paths of change in heritage and non-heritage varieties should be explained in a holistic fashion. A unified account for this simplification is put forth by Terenghi (2023a) and relies on featural and structural considerations. Here we only review the former as one of the major contributions of the *Microcontact* project to the understanding of language change.

The main assumption that underlies the featural account proposed is that demonstrative forms encode deictic information by means of person features and, more specifically, two binary features: [±author] and [±participant] (see the account proposed by Harbour 2016 for person systems). As such, the deictic information encoded in the different demonstrative forms of ternary systems can be derived as follows:

(9) a. Speaker-oriented deictic domain ('near me'): [+author, +participant]
b. Hearer-oriented deictic domain ('near you'): [–author, +participant]
c. Non-participant-oriented deictic domain ('far from us'): [–author, –participant]

Based on (9), we advance the claim that the derivation of 'near you' forms constitutes the locus of complexity within the grammar of deixis, as their featural specification is the only one that involves non-matching feature values (i.e., one + and one –, as opposed to the matching feature values of the other two forms). Non-matching feature values are identified as complex by virtue of a general monotonicity bias active within the person domain (D'Alessandro and Terenghi 2024; Terenghi 2021, 2023a).

In line with this proposal, Terenghi (2023a) proposes that one person feature could be lost in the derivation of demonstrative forms to ease the complexity of (9). This leads to the loss of the dedicated exponent for the hearer-oriented form and concurrently to the reduction of original ternary systems to speaker-based binary systems (these are derived by [±author] alone) or to participant-based binary systems (derived instead by [±participant]). As the monotonicity bias is conceived as a general cognitive bias, its effects are expected to hold across diachrony and contact, in line with the demonstrative data discussed in this section and captured by G2.

We now turn to the presentation of contact phenomena in HLs involving discourse and more generally more than one grammatical module. They are shown to differ radically from what we just saw, and to conform with G3 instead. We start by presenting a case study on null subjects and subject clitics in microcontact.

1.4.3 Subjects in microcontact

Changes in the way the subject is expressed are extremely frequent in HLs; the distribution of null and overt subjects in null subject languages represents one of the most discussed topics in the literature on bilingualism and

HLs, as already mentioned in Section 1.3.1. Most studies analyze the effects of contact of a dominant non-null subject language with a null subject language (e.g., Montrul 2004; Sorace and Filiaci 2006). As we saw above, such studies highlight a tendency to extend the use of overt subjects to discourse contexts in which null subjects would be otherwise expected. Non-null subject languages have only one option (the overt subject pronoun), null subject languages have more than one option; besides, the interpretation of null subjects is not immediately available to the hearer, as it is based on a complex interaction between grammatical and discourse factors. Therefore, change in this domain is attributed to a preference in bilingual speakers for more transparent systems.

Frasson (2022) investigates cases in which new structures are introduced or additional constraints are added to existing structures in heritage null subject languages. This is the case of subject realization in heritage Venetan in Brazil and Argentina; the number of possible configurations of different types of subject pronouns is bigger than in the homeland variety of the language. Notice that Venetan has two overt types of subject pronouns, strong and clitic, as well as null subjects:

(10) Heritage Venetan (Frasson 2022a: 10)

a. Lu no me ga vedesto.
   he not me has seen

b. Lu no'l me ga vedesto.
   he not=he.SCL me has seen

c. No'l me ga vedesto.
   not=he.SCL me has seen

d. El no me ga vedesto.
   he.SCL not me has seen

e. *pro* no me ga vedesto.
   pro not me has seen
   'He did not see me.'

In (10a), a subject pronoun *lu* is overtly realized. In (10b), two subject pronouns appear: the subject pronoun *lu* and the subject clitic *(e)l*. In (10c), only a subject clitic is realized after preverbal negation *no*. In (10d), the same subject clitic is realized before preverbal negation *no*. Finally, in (10e) there are no overt subject pronouns. Options (10d) with a subject clitic preceding preverbal negation, and (10e), without any overt subject pronouns, are not grammatical in homeland Venetan. In heritage Venetan, however, all these possibilities are

attested, evidencing a complexification of the system. While BP shows a partially null-subject system, Frasson (2022) shows extensively that the contexts in which this complexification happens are not the same in which BP expresses the subject overtly. In other words, a direct transfer analysis of the patterns in (10) is not feasible.

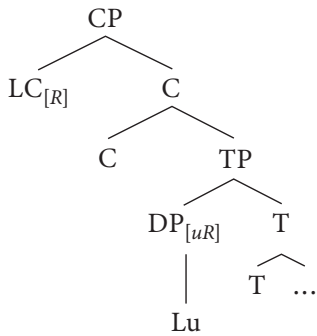
The microcontact data discussed here show a resolution of null subject facts that is, apparently, the exact opposite of that uncovered for other types of contact, as discussed in Section 1.3.1 building on Sorace and Filiaci (2006); in the latter case, we saw that whenever a 1:1 mapping between the language structures is possible, multilingual speakers tend to solve the apparent optionality in the realization of the subject by extending the use of overt subjects, hence simplifying the system. However, in microcontact, null subjects keep constituting one among the multiple possible options to realize the subject: this can be traced back to the difficulty, in microcontact contexts, to exactly pinpoint the locus of variation; therefore, speakers resort to some basic cognitive strategies. In this case, we believe that the relevant cognitive strategy is related to how information is structured.

To account for this divergence, we follow Frasson (2022) in capitalizing on the role of discourse features. In particular, he showed that operations on features taking place in narrow syntax, such as agreement, may involve (besides formal features, such as φ-features) discourse features; these depend on the way speakers structure information, according to the meaning and interpretation that they want to convey. The distribution of discourse features may change in situations of language contact, affecting the realization of subject pronouns as in (10) and the way syntax interacts with information structure. This approach builds on Miyagawa (2005) and Aboh (2010), who propose that agreement for discourse properties is associated with specific discourse features that are assigned to lexical items already in the numeration (see also Wiltschko 2021). The discourse feature ruling the distribution of subject pronouns and their interpretation is defined in Frasson (2022) as [R] (referential): when subject pronouns encode this feature, they are overt and referentially specific enough to obviate or switch reference; when subject pronouns lack this feature, they refer to the most salient discourse antecedent and are normally not phonologically realized.

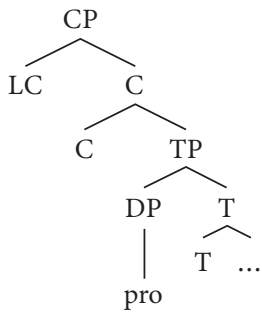
The distribution of subject pronouns in heritage Venetan does not depend on differences with respect to their internal structure (see also Pescarini 2021 for a wider discussion of this possibility): all pronouns have the structure of a DP but are distinguished by the presence or absence of a [uR] feature. The presence of [uR] makes some pronouns referentially specific enough to

switch the reference to a non-salient discourse antecedent (11). Conversely, the lack of [uR] implies that the pronoun will refer to the most salient discourse antecedent. In null subject languages, subject pronouns lacking [uR] can undergo PF deletion under φ-feature identity with finite T, resulting in a phonologically null pronoun (Roberts 2009; Sheehan 2006) as shown in (12).

(11) subjects (Frasson 2022a: 211)



(12) Null subjects (Frasson 2022a: 165)

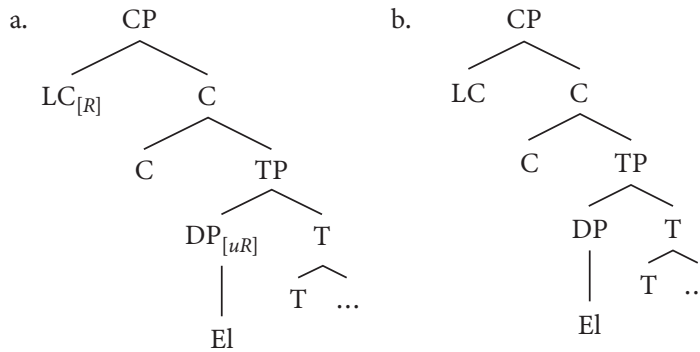


This difference captures the distribution of Venetan strong pronouns (11) and null subjects (12) and defines it in terms of the availability of interpretations allowed by the presence of [uR] in strong pronouns and the lack of such interpretations in null subjects.

However, it was shown that heritage Venetan allows for a third possibility: subject clitics (13). These elements have previously been analyzed in homeland varieties of Venetan as agreement markers, and not as real pronouns, because of their distribution and position inside the sentence (Rizzi 1986); this is shown, for instance, by the fact that homeland Venetan does not allow for subject clitics to be separated from the verb. However, this distribution

is grammatical in heritage Venetan, as shown in (10). Analyzing subject clitics from the perspective of information structure, it is possible to account for their typical pronominal behavior. Frasson (2022) proposes that subject clitics are pronouns that may or may not encode [uR]. Specifically, when they encode [uR], they behave as regular overt pronouns (13a); when they lack [uR], they are phonologically realized counterparts of *pro* (13b).

(13) Subject clitics



The difference in the distribution of [uR] in subject clitics across different varieties of Venetan depends therefore on a rearrangement of the distribution of features known in the literature on second language acquisition as feature-reassembly (Lardiere 2008). Subject pronouns are sensitive to syntactic and discourse factors alike: the interaction of these factors leads to precise conditions of antecedent selection. In contact situations, these conditions are not lost, but rearranged and mainly affect agreement with discourse features: as such, it is not possible to establish a predefined path for the development of pronominal subjects in contact varieties, precisely as they involve a discourse feature whose implementation depends on grammar-external causes (G3). On the contrary, φ-agreement (and thus formal features) does not undergo change in heritage Venetan, in line with G2 (see also Sections 1.4.1 and 1.4.2).

1.4.4 Differential Object Marking in microcontact

We turn now to DOM (Bossong 1985, 1991), the phenomenon whereby a subset of direct objects ("DOs") display a different morphological realization by virtue of their semantic or pragmatic features, as well as of properties of the

verb, as illustrated in Section 1.3.2. DOM in Italo-Romance languages displays a high degree of microvariation: the phenomenon is pervasive especially in central and southern varieties and affects different ranges of DOs. Specifically, it is generally found more often on objects that are high on the Animacy Hierarchy proposed by Silverstein (1976) than on objects in the lower part of the scale, which is reported in (14):

(14) *Animacy Hierarchy*, based on Silverstein (1976) and adapted from Dixon (1979):
1st and 2nd person pronouns > 3rd person pronoun > proper names > human common nouns > animate common nouns > inanimate common nouns

DOM has been the topic of much contact research, especially in the field of heritage linguistics. Restricting the focus to Romance languages, several studies investigate heritage Spanish spoken in the United States (Luján and Parodi 1996; Montrul 2004; Montrul and Bowles 2009; Montrul and Sánchez-Walker 2013; Montrul et al. 2015; Silva-Corvalán 1994) to show that DOM is rather unstable and weakens already in the first generation of heritage speakers. This is in line with the discussion in Section 1.3.1: whenever there is considerable structural difference across the languages in contact (as is the case for macrocontact between Spanish and English), the multilingual speakers tend to simplify their grammar by moving from a two-choice system (in this case: some objects are marked and some are not) to a one-choice system (in this case, no object is marked). In this section, instead, we present the results of our research on DOM as attested in HLs spoken in microcontact: following the rationale outlined in Section 1.3.2 and already validated by the subject data of Section 1.4.3, we expect that also in this case heritage speakers will fall back on general cognitive strategies. In what follows, we show that this prediction is borne out.

In the interest of space, here we limit the discussion to two Romance varieties from Italy, namely Friulian and Abruzzese, spoken as HLs in Argentina and Brazil. The setting of the investigation is such that a non-DOM variety (Friulian) and a DOM variety (Abruzzese) are investigated in contact with both DOM (Argentinian Spanish[9]) and non-DOM varieties (BP). The results of our investigation are partially antithetical with respect to what has been

[9] In particular, most of our data come from the Rioplatense variety of the language, where DOM is more extended than in European Spanish, as noted in Saab (2018).

described for the homeland varieties: the HLs under discussion display different patterns of DOM, and more precisely we see that the rate of DOM increases, rather than decreasing, both in heritage Abruzzese and heritage Friulian. On the one hand, heritage Abruzzese exhibits DOM on referents that are typically not affected by the phenomenon in the homeland variety, as shown in (15); on the other hand, DOM emerged in heritage Friulian (16):

(15) Heritage Abruzzese in Argentina (Andriani et al. 2022b: 18)

Lu	lopə	s'a	magnatə	**a**	**nu**	**gnillə.**
the	wolf	si=has	eaten	DOM	a	lamb

'The wolf ate a lamb.'

(16) Heritage Friulian in Argentina (Andriani et al. 2022b: 18)

Tu	as	fât	un	sium.	Tu	as	bussât
you.SCL	have	made	a	dream	you.SCL	have	kissed

a	**to**	**fie.**
DOM	your	daughter

'You had a dream. You kissed your daughter.'

These data point to an additional conclusion, namely that topicality conditions are involved both in the emergence of DOM in heritage Friulian, and in the extension of DOM in heritage Abruzzese. We identified therefore the effects of a new factor in the attribution of the marking, different from that which is found in the homeland variety, where the position of the object is irrelevant for DOM (Ledgeway et al. 2019). In other words, in both HLs DOM is accepted more often in fronted topic position than in situ. Importantly, this result also differs from what has been noticed for heritage Spanish (Luján and Parodi 1996; Montrul 2004; Montrul and Bowles 2009; Montrul and Sánchez-Walker 2013; Montrul et al. 2015; Silva-Corvalán 1994). In (17a) and (17b), we show that DOM was accepted for one and the same type of referent by one out of four heritage Friulian speakers when that object was in situ, and by four out of four heritage Friulian speakers when it was instead topicalized. Similarly, for heritage Abruzzese, one and the same referent was never differentially marked when in situ, but both our informants accepted it with DOM when in topic:

(17) Heritage Friulian in Argentina

a.

O	ai	copat	(a)	chel	sarpint	neri.	DOM: 1/4
I.SCL	have.1SG	killed	DOM	that	snake	black	

'I killed that black snake.'

b. (A) chel sarpint neri, îr, lu=avin
DOM that snake black yesterday it=have.1PL
copât. DOM: 4/4
killed
'That black snake, yesterday, we killed it.'

Thus, DOM is not only preserved, but even strengthened in both varieties at issue spoken in Argentina, in clear contrast with the outcomes of macrocontact as discussed by the available literature (Spanish in contact with US English). Moreover, we found that heritage Friulian speakers in Brazil also accept DOM to some degree. Similar to what was described for Argentina, DOM acceptance is favored by topicalization of the DO. Since Brazilian Portuguese does not display DOM, this strongly suggests that this change does not follow from transfer, but is in fact endogenous, i.e. it mirrors the diachronic pattern of DOM emergence. This is further in line with findings from diachronic studies of DOM in Romance varieties (Iemmolo 2009, 2010). Another argument for the endogenous nature of the change at issue lies in the fact that the distribution of DOM in heritage Friulian displays different features with respect to the contact languages: on the one hand, Brazilian Portuguese lacks DOM altogether; on the other hand, Argentinian Spanish features a system where DOM affects also non-animate referents, independently of their position within the sentence (Di Tullio et al. 2019). In other words, topicalization of the objects does not affect the presence of DOM in Argentinian Spanish. As such, we can conclude that change in the domain of DOM is in conformity with G3.

1.5 Microcontact and syntactic theory

In Section 1.4, four case studies were presented on microcontact. Their results can be summarized as follows, in structural terms.

1.5.1 Auxiliaries

Auxiliary selection in microcontact yields two separate outcomes: on the one hand, HAVE auxiliaries overextend; on the other, a new rule emerges for the distribution of BE auxiliaries. As broadly discussed in Andriani and D'Alessandro (2022), this indicates a simplification of the paradigm, as the

alternation between two forms changes into a mono-selectional process. This result parallels those that have been discussed in Section 1.3.1 regarding the reduction of null subjects and the overextension of overt subjects and is very likely dictated by economy considerations. However, as Andriani and D'Alessandro also argue, selecting HAVE means opting for the most informative and structurally complex auxiliary (Kayne 1993; Tuttle 1986). The fact that simplification is not the only direction taken by auxiliation is further witnessed by the appearance of specialized BE auxiliaries in many Venetan varieties in Brazil. There, a distribution not existing in the Italian-based counterparts and bearing on the position of the subject and the anti-agreement effect originating with postverbal subjects is at work. This distribution suggests the complexification of the auxiliary selection system.

The conclusions that we can draw, based on the observation of the CIC of auxiliary systems, is that economy principles might be at work in featural systems overall, but that we are not authorized to conclude that they will always be at work, especially as far as features are concerned. The existence of specialized emerging forms suggests in fact the opposite: an increase in the complexification of the system. However, following Polinsky (2018), the emergence of specialized forms resulting in one-form-to-one-function mappings can also be interpreted as reduction of complexity (see also Lohndal and Putnam 2021). Given the scarcity of data, it is not possible to draw firm conclusions regarding auxiliary selection in contact. We limit ourselves to identifying two different paths, one of overextension and paradigmatic reduction, and another one of specialization.

1.5.2 Demonstratives

A clearer view is instead offered by demonstratives in contact. In that case, a rather large data set (see D'Alessandro and Terenghi 2024; Terenghi 2023a) supports the observation that systems featuring mismatching values which break the monotonicity of the functional sequence contained in demonstratives tend to be reduced in a very systematic way: the functional sequence disrupting monotonicity is reduced. In Terenghi's study, a direct comparison is made between the developments of demonstrative systems in HLs, standardized (monolingual) varieties, and creoles. Interestingly, these systems exhibit analogous patterns of behavior. The conclusion to be drawn is that change affecting φ-features is largely predictable and uniform (see also Sorace 2011 for similar conclusions). Regarding the internal structure of demonstratives,

this study allowed us to understand and pinpoint the exact universal sequence of functional heads composing them, by observing the elements that get eliminated first and their consistent peripheral position in the syntactic spine.

1.5.3 Null subjects and DOM

The same cannot be said when it comes to phenomena involving discourse-related interpretations, which are much more unpredictable. As we have seen, both null subjects and DOM evolve in a very different fashion than they do in simple dyadic contact. However, both null subjects and DOM in microcontact mirror the diachronic development of the same elements: they both rely heavily on topicality. While null subjects in dyadic contact tend to be replaced by overt subjects, this is not the case in microcontact. Likewise, DOM in dyadic contact mostly weakens (Montrul 2004), while in microcontact it gets reinforced and extended.

The conclusions that we can draw here are of two kinds: the first one is that these results cannot be interpreted easily through the TPM, as this model is too coarse-grained and vacuously effective, given that we are comparing different dialects of the same Romance macro-grammar. The second one is that, as mentioned in Section 1.3.2, when two varieties enter in microcontact with one another it is difficult to pinpoint the locus of variation, as structural mapping is identical, and what varies is usually only the value of a feature, but not their structure overall (as extensively shown in Frasson's study on subject clitics and null subjects). In this case, the grammars behave as they do in diachrony in most cases: while DOM in situ can be weakened, DOM on topics systematically reappears. While agreement clitics become pronouns, pronouns start getting dropped in topic continuation situations. This, we maintain, reflects a general cognitive principle at work also in L1 acquisition, linked to the development of the clause and on the importance of topicality in the construction of the sentence (Belletti 2017, 2018; Jourdain et al. 2020; a.o.).

Regarding the structures themselves, microcontact helped us identify two important facts. The first one is that, based on their behavior in microcontact, sensibly different from that of dyadic macrocontact, there is no evidence for the distinction between subject clitics, full pronouns, and null subjects other than the value attributed to the R(eferential) feature. The second one is that DOM in situ and in topic are two rather distinct phenomena, and that the former is an extension of the latter (pretty much in line with Iemmolo's proposal,

but only partially with the analyses considering DOM as phenomenon involving just the checking of an animacy or definiteness feature, see López 2012, Torrego 1998, and others).

In general, the results that *Microcontact* yielded have identified some shortcomings in both the current methodology in use in HL studies and the empirical base against which generalizations are drawn. Methodologically, drawing generalizations on the basis of dyads in macrocontact seems to be the wrong way to go, as it shows only a part of the picture. We have systematically shown that change in microcontact deviates from what is usually considered to be the standard behavior of languages in contact. This indicates further directions for the study of CIC, targeting also creoles and koines, with the aim of capturing the cognitive underlying principles of HL syntax, of which what we see could be just a little part. In fact, the behavior of dyads in contact could constitute an exception rather than the norm as far as the cognitive strategies for resolving contact are concerned.

Another crucial contribution of *Microcontact* to linguistics has been the enlargement of the database, to include dialects of the same family, of underdocumented varieties. As also underlined by Scontras and Putnam (2020), enlarging the pool of dyads (and of more than two languages) in contact is essential for a clear and more accurate understanding of CIC, and of HL syntax in particular.

1.6 Conclusions

This chapter highlights the significance of incorporating the micro-dimension and cross-checking as analytical tools for examining CIC. Methodologically, if we were to examine language dyads alone, it might lead to a mistaken conclusion that several observed changes, such as the emergence or expansion of DOM, are primarily a result of transfer. Examining the same microphenomenon in multiple contact situations, with all other factors held constant, is the only approach that can definitively address questions about the conditions that license change.

Empirically, we hope to have shown that we cannot easily conclude that CIC and CID are perfectly parallel (as proposed by Kupisch and Polinsky 2022), but that change targeting fixed sequences of functional heads is easier to predict. Conversely, phenomena involving the interfaces are much more complex and unpredictable. In that case, we have shown that, while it is still possible to identify some triggers for change, it is not easy to identify a unique direction

of change. The configuration in which the languages enter in contact plays an important role here, unlike for core syntactic features, in that the perception of the locus of variation is easier in language dyads, especially if the languages are sufficiently distant structurally. If structural mapping identifies a point of variation, the conflict between structures can be resolved mainly toward simplification. If this identification is more difficult, like in the case of minimally differing languages, the speakers resort to basic cognitive strategies for the resolution of grammar conflict, like topic marking in dislocation, a strategy which is also learned very early on by L1 acquirers and at work diachronically.

Microcontact studies have yielded valuable insights into the structure of demonstratives, null subjects, subject clitics, and DOM. If anything, *Microcontact* has underscored the significance of the micro-dimension when attempting to formulate generalizations. However, it has also demonstrated that generalizations remain attainable; they simply require a comprehensive consideration of all potential contact configurations, specifically including microcontact.

2

Systematic and predictable variation in heritage grammars

The role of complexity, diachronic change, and linguistic ambiguity in the input

Esther Rinke and Cristina Flores

2.1 Introduction

Previous research on heritage speakers of various languages has shown that their linguistic production and comprehension show specific particularities (Montrul 2016; Polinsky 2018). In addition, heritage speaker populations show generally high levels of individual variation (Benmamoun, Montrul, and Polinsky 2013; Rinke and Flores 2014).

However, as we have argued in previous work, the structural particularities of heritage languages should not easily be equated with deviance or lack of linguistic knowledge (Flores and Rinke 2020b). In particular, we have claimed that deviance has to be distinguished from linguistic variation that is already present in the input to the heritage speakers and visible in other language acquisition settings (particularly in monolingual acquisition). As shown by a growing body of research on heritage languages, variation in heritage grammars is not random and not the result of deficient grammars (D'Alessandro 2023; D'Alessandro, Natvig, and Putnam 2021; Łyskawa and Nagy 2019; Requena 2023; Rinke and Flores 2014; Shin, Cuza, and Sánchez 2022).

One major difficulty in the study of heritage grammars consists in distinguishing variation from deviance or lack of knowledge, especially because linguistic variation comes in different shapes and has different and partially overlapping dimensions. We may, for instance, differentiate between intra-linguistic variation (on different linguistic levels), register variation, diachronic variation, and cross-dialectal variation, including at the level of

Esther Rinke and Cristina Flores, *Systematic and predictable variation in heritage grammars*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro et al., Oxford University Press. DOI: 10.1093/9780191987731.003.0003

microvariation (D'Alessandro 2021; see also D'Alessandro et al. 2025, Chapter 1 of this volume). In our view, divergent structures are only those structures that are not present in and not predictable from the input that heritage speakers receive in their heritage language. A case in point would be, for example, structures transferred from the dominant environmental language into the heritage grammar, which are not present in non-contact varieties of the target language (e.g., deviant case morphology in heritage Russian under the influence of Hebrew or Dutch as reported by Meir and Janssen 2021; or fusion of the gender and number heads in the nominal domain in Spanish under the influence of English, as discussed by Scontras, Fuchs, and Polinsky 2018). In contrast, register or dialectal variation, ongoing diachronic change or (structural or lexical) ambiguity in the input to heritage speakers may predictably lead to variation in heritage grammars, even though this does not necessarily mean that the options found in the input will be retained in exactly the same way in the heritage language. Nevertheless, specific properties of heritage languages can be regarded as a deterministic outcome of variable and reduced input in first language acquisition. In the following sections, we will argue that (i) variation can be amplified or more pronounced in heritage speakers under the specific acquisition constraints that shape heritage grammars, and that (ii) it is indispensable to take into consideration the variation in the colloquial varieties of the input language in order to be able to differentiate between deviation/language transfer and existent or predictable linguistic variation which is independent from the contact language (see Alexiadou and Rizou 2025, Chapter 5, for a case in point in heritage Greek in Germany and the United States). Of course, this does not exclude that some linguistic areas may be resilient to variation and change, as, for example, shown in Flores, Rinke, and Wagner 2022 (see also Lohndal and van Baal 2025, Chapter 8 of this volume).

In the following sections, we discuss a number of concrete examples for enlarged linguistic variation in European Portuguese (EP) as heritage language in contact with German and other environmental languages. Concretely, we will focus on three different dimensions of linguistic variation: (a) variation caused by linguistic complexity (Section 2.2, clitic placement), (b) variation caused by ongoing diachronic development (Section 2.3, null objects), and (c) variation caused by lexical and grammatical ambiguity (Section 2.4, interpretation of overt pronouns; multiple functions of *que* in EP).

2.2 Variation caused by linguistic complexity: Clitic placement in EP

The first showcase of systematicity of variation in heritage grammars comes from clitic placement in EP. Object clitic placement in EP can be viewed as a complex phenomenon since enclisis and proclisis depend on a range of syntactic constraints.

In contrast to other Romance languages (e.g., French or Spanish), proclisis is restricted to a number of well-defined contexts, namely to (i) subordinate clauses introduced by a complementizer (1a); (ii) *wh*-clauses (1b); (iii) aspectual adverbs (1c); (iv) sentential negation and negative adverbs (1d); and occurs (v) whenever a non-specific indefinite Quantifier Phrase (QP), a negative QP, a universal QP, or a DP modified by a focus particle precedes the verbal complex within the minimal CP that contains it (1e) (Barbosa 2000).

(1) a. A Maria afirma que **o** viu [Complement clause]
the Maria claims that CL saw
na rua.
on+ the street.
'Maria claims that she saw him on the street.'

b. Quem **o** viu? [*wh*-phrase]
who CL saw
'Who saw him?'

c. A Maria já/ ainda **o** viu. [Aspectual adverbs]
the Maria already/ still CL saw
'Maria already/still saw him.'

d. A Maria não **o** viu. [Negation]
the Maria not CL saw
'Maria did not see him.'

e. Todos **o** viram na rua. [Subject QP]
everybody CL saw at+the street
'Everybody saw him at the street.'

Enclisis occurs in verb-initial declarative main clauses or sentences introduced by a subject DP or pronoun (2a) or in topicalization structures (2b).

(2) a. A Maria/Ela viu-**o**. [Declarative main clause]
the Maria/she saw-CL
'Maria/She saw him.'

b. O Jorge, a Maria viu-**o.** [Topicalization]
the George, the Maria saw-CL
'George, Mary saw him.'

Given the different conditions and contexts, it is not clear whether proclisis or enclisis is the default or the more basic option, in particular because the frequency of the enclitic and proclitic patterns are very similar and cannot be taken as indicators of a default option (Flores and Barbosa 2014).[1] We argue that a look at heritage speakers of EP may clarify this question and reveal that enclisis is the more basic pattern.

Our first evidence comes from a study on child heritage speakers of EP (7–15 years of age) living in Germany. Flores and Barbosa (2014) analyzed clitic placement in Portuguese by means of an oral sentence production task. The results of the study showed that the heritage children overgeneralized enclisis in contexts where proclisis was required (see (3)), but not proclisis in contexts requiring enclisis.

(3) Agora já o vi.
now already CL saw
'Now I already saw it.'

Two subsequent studies on adult heritage speakers of EP living in Germany (Flores, Rinke, and Azevedo 2017; Rinke and Flores 2014) confirm that heritage speakers accept and produce more enclisis in contexts of proclisis than the other way around. The study by Rinke and Flores (2014), based on a Grammaticality Judgment Task, shows that heritage speakers are much more prone to accept enclitic placement in sentences requiring proclisis than proclisis in enclisis contexts. The same is true for production. Flores, Rinke, and Azevedo (2017) revealed that in spontaneous speech, adult EP heritage speakers did not frequently diverge from the expected pattern. However, if they did, enclisis in contexts of proclisis (5.2% of all proclisis contexts) was twice as frequent as proclisis in contexts of enclisis (2.4% of all enclisis contexts, see Table 2.1).

[1] Flores and Barbosa (2014) carried out an automatic search on the Linguateca Speech Corpus Museu da Pessoa. In a total of 6501 occurrences of clitics, 3380 tokens of enclisis (52%), and 3121 (48%) tokens of proclisis were attested. Enclisis was only slightly more frequent than proclisis, showing that frequency is not the relevant factor in determining a basic order.

Table 2.1 Clitic placement (raw counts, % of deviant placement; G2-Bil refers to the heritage speakers)

	G1_BIL	G2_BIL	G1_MON	G2_MON
Proclisis contexts – use of enclisis	369 7 (1.9%)	270 14 (5.2%)	365 15 (4.1%)	321 3 (0.9%)
Enclisis contexts – use of proclisis	418 2 (0.5%)	210 5 (2.4%)	382 1 (0.3%)	240 0

Source: Flores, Rinke, and Azevedo (2017)

The fact that heritage speakers show difficulties with proclisis and overgeneralize enclisis suggests that enclisis may be the more basic clitic position and proclisis the more complex option that is acquired later.

The findings reported by Flores and Barbosa (2014) and Flores, Rinke, and Azevedo (2017) are in line with the results reported in studies on monolingual acquisition of EP clitic placement, which reveal that children up to age 8/9 show the same kind of variation as heritage bilinguals. As demonstrated by Costa, Fiéis, and Lobo (2015), EP children start to overgeneralize enclisis when acquiring the clitic system, i.e. younger children use enclisis instead of proclisis in several contexts but the use of proclisis in enclitic contexts is rarely attested.[2]

The findings from acquisition studies have received different theoretical explanations. In general, it is a common assumption that enclisis is the *default clitic position* in EP (Duarte and Matos 2000). Some researchers link the observed pattern to computational costs associated with derivational complexity assuming that the overgeneralization of enclisis in proclisis contexts is evidence for the default, less complex configuration of the enclitic pattern (Costa et al. 2015; Duarte and Matos 2000; Duarte, Matos, and Faria 1995; Duarte, Matos, and Gonçalves 2005). A different account is proposed by Flores and Barbosa (2014), based on Barbosa (2000, 2008), who claims that the

[2] We would like to note that this may be changing due to the strong presence of Brazilian *YouTubers* in Portuguese social media, including in programs and videos for young children. There is still no scientific research on this topic, but ongoing discussions on social media forums involving Portuguese parents and teachers, and several magazine articles published in the last two years document the increasing tendency of Portuguese preschoolers to develop Brazilian-type patterns at the levels of morphosyntax, phonology, and lexicon, which is particularly visible in the pronominal system. See, for instance, the newspaper article "Há crianças portuguesas que só falam 'brasileiro'" ("There are Portuguese children who only speak 'Brazilian'"), published in November 2021 in *Diário de Notícias*, https://www.dn.pt/sociedade/ha-criancas-portuguesas-que-so-falam-brasileiro-14292845.html.

clitic is inserted to the left of the functional head containing the verb. Enclisis is derived in the post-syntactic component of the grammar by a rule of *Local Dislocation* placing the clitic in postverbal position. In this scenario, the overgeneralization of enclisis derives from a strategy that is in line with the Subset Principle: the acquisition process starts with the grammar that generates the smallest possible configuration that is compatible with the trigger experience: Local Dislocation applies without exception. During the acquisition process (i.e., being exposed to sufficient positive evidence), the child has to learn in which contexts this rule is blocked (an operation similar to the Blocking Principle proposed by Andrews 1990).

Both accounts consider enclisis to be the default option in language acquisition, but for different reasons. In either account, enclisis comes for free and the child has to gather enough evidence to acquire the constraints that lead to proclitic placement. Interestingly, previous research has shown that different proclisis triggers are not acquired at the same time by EP children (Costa et al. 2015). There is an order of acquisition of these constraints (a developmental scale) that goes from contexts that are stabilized first (e.g., sentential negation) to contexts that are acquired latest (e.g., contexts with quantified subjects or certain complement clauses), with several contexts in-between.

What is relevant for the present discussion is the explanation for this developmental scale. Costa et al. (2015) hypothesize that it is the degree of variation associated with each proclisis trigger in the input that delays the acquisition. Hence, the more variable the input the later the EP child acquires the proclisis trigger and, consequently, the more variation is expected in heritage speakers.

One example of variation in the input is found in the context of embedded clauses. Although finite embedded clauses trigger proclisis in the standard grammar, in colloquial EP we do find instances of enclitic placement in certain embedded clauses (Flores, Rinke, and Azevedo 2017), particularly in indicative complement clauses introduced by verbs like *dizer* 'say' or *achar* 'think' (see (5a) below) (Duarte, Matos, and Faria 1995). This is a clear instance of register-based variation. In addition, certain infinitival contexts categorically trigger enclisis (e.g., the infinitival complements of perception verbs like in (4)).

(4) A Maria viu o Pedro sujar-se todo.
the Maria saw the Pedro get dirty-CL all
'Maria saw Peter getting all dirty.'

This means that children acquiring this proclisis context have to figure out the particular embedded contexts which trigger proclisis (or block the rule of Local Dislocation in Barbosa's 2008 model). Determining these variable contexts takes more time than acquiring categorical proclisis triggers that do not show any variation, such as the contexts of sentential negation.

Now let us turn again to the studies on heritage speakers mentioned above and look in more detail into the different proclisis triggers. In the study by Flores and Barbosa (2014) on clitic placement in child heritage speakers, it was, indeed, in the contexts with subordinate clauses (introduced by the bridge verb *achar* 'to think', see (5a)) where the use of enclisis in proclisis contexts was particularly pronounced. Inversely, the contexts with negation (5b) were the most accurate ones (see Figure 2.1).

(5) a. Achas que ela se magoou?
think that she CL hurt.PST
'Do you think she got hurt?'

b. Isso não me agrada.
that not CL like
'I don't like it.'

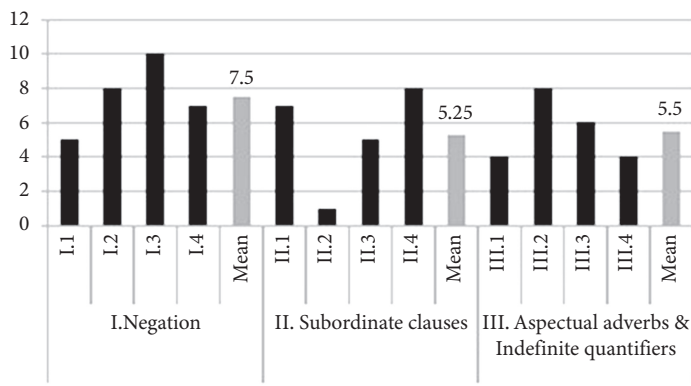


Figure 2.1 Accurate use of proclisis per condition (raw counts).
Source: Flores and Barbosa (2014)

A closer look at adult heritage speakers in Rinke and Flores' (2014) acceptability judgment results in the proclisis conditions confirms the differences between the various proclisis triggers. The heritage speakers were more accurate in rejecting ungrammatical enclisis in negation contexts than in complement clauses (see Table 2.2).

Table 2.2 Clitic placement with negation and subordinate clauses (% of accuracy)

Experimental condition	Monolingual speakers ($n = 18$) mean (SD)	Heritage speakers ($n = 18$) mean SD	Mann–Whitney U	p
Preverbal position with negation (grammatical)	100.(00)	89.22 (19.20)	117.00	<.05
Postverbal position with negation (ungrammatical)	95.56 (8.56)	67.78 (28.40)	42.00	<.001
Preverbal position with subordinate clauses (grammatical)	96.67 (7.67)	86.67 (16.80)	111.000	0.05
Postverbal position with subordinate clauses (ungrammatical)	80.00 (20.58)	52.50 (27.24)	70.500	<.05

Source: Rinke and Flores (2014)

Interestingly, in the same study, we also find differences in the rate of acceptability between negation and subordinate clauses in the monolinguals group (95.5% of rejection of enclisis with negation, but only 80% of rejection of enclisis in subordinate clauses).

What can we conclude from the results of the various studies? We argue that the variation found in the different heritage bilinguals groups tested (children and adults) is systematic and predictable: it is related to complexity (going only in one direction, namely from the more to the less complex pattern) and it is related to variation in the input. These factors are also responsible for the varying paces of acquisition of proclisis in monolingual first language acquisition.

2.3 Ongoing diachronic development: Null objects in EP

In this section, we will provide another example of systematic and predictable variation in heritage speakers of EP: the use of null object constructions. Portuguese is a null object language, allowing for the omission of definite direct objects as in example (6).

(6) [Eles têm o seu carrinhos.]
[they have their car(s).]

Ela vem buscar Ø e pronto e
she goes pick up (it/them) and prompt and

nós lá vamos.
we there go

'She picks it/them up and there we go!' (PSFB corpus, Barbosa 2014)

Null objects are a frequent and common pattern in European (and also in Brazilian) Portuguese. They occur in contexts where the object represents given information and is mentioned in the previous context. Pragmatically, this is the same context in which object clitics occur too. According to Raposo (2004), there is a certain parallelism between null object constructions (7b) and clitic left dislocation (CLLD) topicalization structures (7a) in EP (see also Kato and Raposo 2005), the difference only being that the clitic realizes a D-head whereas the null object construction includes a null determiner.

(7) a. (esse livro) eu só encontrei [$_{DP}$ o pro] na FNAC

b. (esse livro) eu só encontrei [$_{DP}$ D pro] na FNAC
(this book) I only found it/__ in-the FNAC

'This book, I only found it at the FNAC shop.' (Raposo 2004)

Given that clitics and null objects seem to be two sides of the same coin, the question arises what determines the choice between the two options.

There are syntactic, diastratic, and semantic conditions shaping the variation between clitics and null objects in EP. Syntactically, null objects are excluded from or dispreferred in syntactic islands in EP, but clitics are not constrained in the same way. Concerning register, null objects are part of the colloquial speech but are rare in the written/formal register, whereas clitics occur both in the written/formal and the spoken/colloquial register. In fact, there is no register variation concerning the use of clitic pronouns, i.e. clitics are frequent in colloquial and in standard EP. With respect to their semantic features, null objects are more likely to occur with inanimate referents and are avoided with animate referents. By contrast, clitics are not constrained with respect to animacy (Rinke in press). Concerning the semantic features, Cyrino, Duarte, and Kato (2000) propose that the probability of having a null or overt realization of objects is constrained by referential features. The authors propose a referential scale as given in (8) and claim that the likelihood

of using an overt form increases universally from left to right on the scale. According to the authors, this scale, called *Implicational Referential Hierarchy* (IRH), does not only account for synchronic and cross-linguistic variation, but also for diachronic change: if a language develops from a non-null object to a null object language, the development would start with the realization of less referential objects as null objects and proceed from left to right to more referential ones.[3]

(8) *Implicational Referential Hierarchy*

non-argument	proposition	[–human]		[+human]	
		3 p.		3 p.	2 p. 1 p.
			–spec.	+spec.	

[–ref] < ---------- > [+ref.]

If all these observations are on the right track, the following predictions can be made for heritage grammars: If heritage speakers rely predominantly on colloquial input, they may show a diachronically advanced behavior. Consequently, we expect them to use more null objects and to extend their use to more referential referents, proceeding from left to right on the IRH.[4] The alternative is less likely: if heritage speakers rely more on the formal register and are diachronically more conservative in this domain of pronoun realization, we would expect them to realize more clitics and fewer null objects. Actually, the consistent use of clitics would be a very successful acquisition strategy since the use of clitics – in contrast to null objects – is not dependent on semantic and syntactic constraints, nor is it subject to register variation.

As claimed above, systematic variation visible in heritage grammars may be echoed in first acquisition of a target language. Thus, a further look at the acquisition of the EP pronominal system in monolingual settings is due. It is a well-established fact that object omission is cross-linguistically a very frequent strategy in first language acquisition (Pérez-Leroux, Pirvulescu, and Roberge 2017). Children go through a stage in which they omit objects, even in languages that do not employ null objects (Wang et al. 1992). The timeline of this stage of object omission varies from language to language due to language-specific features that make the learning task more or less

[3] Of course, the opposite development is also possible (extension of full pronouns from more referential to less referential), as exemplified by Cyrino et al. (2000) for subject pronouns in Brazilian Portuguese.

[4] This extension may also affect island contexts, but this has not been systematically investigated so far. There are some residual occurrences in Rinke, Flores, and Barbosa (2018) indicating that this may be the case.

time-consuming. Extensive research has shown that this object omission stage is very pronounced in EP (Costa and Lobo 2007; Flores, Rinke, and Sopata 2020; Varlokosta et al. 2016), because the EP pronominal system is argued to be very complex (Costa, Fiéis, and Lobo 2015). In particular, even at school-age, EP children still extend null objects to syntactic contexts that do not license them according to previous research, namely to island contexts (Costa and Lobo 2007). Nevertheless, research also shows that EP children have early pragmatic knowledge, meaning that the children are sensitive from an early age to the pragmatic conditions that determine the distribution of pronominal objects (Costa et al. 2009; Flores et al. 2020).

In sum, the overuse of null objects is prolonged in the acquisition of EP, mainly due to the syntactic and semantic constraints on object realization and the variation between clitic and null pronouns. Now let us turn to the heritage speakers of EP.

In Rinke, Flores, and Sopata (2019), we analyzed the production of different types of direct objects in the heritage language of two groups of child heritage speakers living in Germany: those of Portuguese and those of Polish descent children. Both languages display clitic and null objects, thus we aimed to analyze whether bilingual children acquiring Polish and EP as heritage languages are sensitive to the language-specific determinants of object realization (namely pragmatic accessibility and animacy). Focusing now on our target group, the EP heritage children, the results showed that these children are in fact sensitive to the pragmatic and semantic constraints of EP object realization, but they overgeneralized null objects (and full noun phrases) until later stages of development (age 9) – thus longer than monolingually raised EP children (Flores et al. 2020). Importantly, the heritage children followed the same path of acquisition of object realization as described for homeland children, not showing signs of cross-linguistic influence from the contact language, German.

In a corpus study analyzing the speech of adult heritage speakers of EP living in Germany, we investigated object realization in this speaker group compared to first-generation migrants and to monolingually raised EP speakers in Portugal (Rinke, Flores, and Barbosa 2018). Interestingly, the results revealed higher rates of null objects in the heritage speakers' speech, as well as in age-matched monolinguals, compared to the two older generations. We argued that this tendency might reflect an inter-generational development that favors null objects. What is relevant for the present discussion is a particular tendency observed in the speech of heritage speakers. The analysis of the syntactic and semantic conditions determining the occurrence

of null objects revealed higher rates of animate and non-propositional null objects in this group, differently from the other speaker groups. This is precisely the tendency observed diachronically for object realization, captured by Cyrino et al.'s (2000) referential hierarchy. Thus, the heritage grammar analyzed in this study may reflect a language-internal diachronic pathway, which is independent of language contact. This observation adds to the evidence in favor of systematic variation governing heritage languages, being a showcase of diachronic development (see Frasson 2025, Chapter 7 of this volume for diachronic change of subject clitics in heritage Venetian and object clitics in heritage Bulgarian). The systematicity of this grammatical outcome is further strengthened by the particular development observed in EP first language acquisition, which is characterized by a prolonged stage of object omission.

We conclude that, due to their particular input conditions (i.e., dominant exposure to colloquial registers and reduced input), children acquiring EP as a heritage language may show a delayed overuse of null objects and adult EP heritage speakers extend the use of null objects along a referential hierarchy. In doing so, they foster diachronic development that is ongoing in the target language (Rinke and Flores 2021).

2.4 Lexical and grammatical ambiguity: Overt subject pronouns and multiple functions of *que*

Null subject languages like EP show variation between overt and null subject pronouns. There is generally a strong tendency to associate a null pronoun with the subject of a previous sentence and an overt pronoun with the object referent (see (9)). This tendency of null subjects to express topic continuity and of overt pronouns to express topic shift has been captured in terms of the *Position of Antecedent Hypothesis* (PAH) by Carminati (2002). This predicts that, in ambiguous contexts, the null pronoun refers to the subject in the IP, whereas the overt pronoun tends to select an antecedent lower in the phrase structure.

(9) a. A mãe cumprimentou a avó
the mother greeted the grandmother
quando *pro* entrou na cozinha.
when pro entered in-the kitchen
'The mother greeted the grandmother when she entered the kitchen.'

b. O avô fotografou o menino quando
the grandfather took-a-picture the boy when
ele saiu da garagem.
he came-out of-the garage
'The grandfather took a picture of the boy when he left the garage.'
(Lobo and Silva 2016: 327)

The realization and interpretation of null and overt subjects is a notoriously difficult area for bilingual speakers (Kaltsa, Tsimpli, and Rothman 2015; Laleko and Polinsky 2017; Montrul 2004; Sorace, Filiaci, and Baldo 2009; Tsimpli et al. 2004; see Polinsky 2025, Chapter 3 of this volume). Interestingly, it can be shown that this variation in the heritage language is not due to cross-linguistic influence because it also exists in contexts in which both the heritage and the societal language are null subject languages, as shown by Rinke and Flores (2018) and Sorace et al. (2009).

The study by Rinke and Flores (2018) investigated the interpretation of null and overt subject pronouns in monolingual adults and three groups of children and teenagers in the age span between 8 and 16 years: monolingual EP children, German-EP bilinguals, and Spanish-EP bilinguals. The results reveal that the monolingual and bilingual speakers differentiate between null and overt subject pronouns and preferentially interpret null subjects in terms of topic continuity and overt pronouns as topic shift. Nevertheless, the bilingual speakers show some interesting particularities: first, they show more variation in comparison to the monolinguals with respect to the interpretation of null subjects in cataphoric and intersentential conditions (see (10)), but not in the intrasentential condition, see example (9) above.

(10) A Sónia telefonou à Susana. Como sempre
the Sónia called to+the Susana as always
pro/ela está atrasada.
pro/she is late
'Sónia called Susana. As usual she is late.'

Intersentential contexts like (10) show a weakening of the PAH also in the adult language (e.g., Fedele and Kaiser 2014) and are acquired later by monolingual children than the intrasentential contexts. These children, however, overcome the difficulties more or less by the age of 8–9 (Lobo and Silva 2016). Hence, given the specific input conditions of heritage bilinguals, it is expected that the bilingual children in the study by Rinke and Flores (2018) may need more time to fully master this domain.

More interesting is the second particularity shown by the EP heritage children, namely the variation in the overt pronoun conditions. With respect to overt subject pronouns, all groups of (monolingual and bilingual) children/teenagers show a non-adult-like behavior and do not really differentiate between the different syntactic conditions (anaphoric/cataphoric/intersentential). Rinke and Flores (2018) conclude that the interpretation of overt pronouns represents a particularly difficult acquisition task that is especially demanding for bilinguals. This result is in line with a number of previous studies showing that heritage speakers tend to overuse overt subject pronouns and may have difficulties interpreting them (Keating, Van Patten, and Jegerski 2011). This variation is expected given that overt subject pronouns are not always contrastive and do not consistently behave like strong pronouns in the sense of Cardinaletti and Starke (1999). Crucially, the PAH is not the only criterion determining the variation between null and overt subject pronouns. The likelihood of an overt pronoun referring to a subject constituent increases (or is mandatory) if there is a gender mismatch between the pronoun and the object constituent (11b), if there is only one antecedent available (11a) or if antecedent and overt pronoun occur in two separate main clauses (Alonso-Ovalle et al. 2002; Carminati 2002; Morgado 2013; Rinke and Flores 2018; (11c)). In addition, also contextual and world knowledge interferes significantly (e.g., in (11c) the likelihood that the pronoun refers to the subject antecedent is high because, generally it is the father that buys a car for the son and not the other way round).

(11) a. A Maria acendeu a luz quando *ela* entrou
the Mary turned on the light when she entered
na cozinha
in-the kitchen
'Mary turned on the light when she entered the kitchen.'

b. O avô fotografou a neta quando
the grandfather took a picture the granddaughter when
ele saiu da garagem.
he came out of the garage
'The grandfather took a picture of the granddaughter when he came out of the garage.'

c. O pai chamou o filho quando *ele* lhe comprou
the father called the son when he CL bought
o carro.
the car
'The father called the son when he bought him the car.'

In line with the previous argumentation, we argue that the variation found in the speech of heritage speakers is, again, not random but predicted by the ambiguity of overt subject pronouns in the target language. In this sense, heritage speakers do not diverge from the systematic path observed in the development of EP and other Romance languages. Furthermore, in many child and adult heritage speaker populations this variation is reduced (or similar to monolingual populations), reinforcing the conclusion that it is a language-internal phenomenon that is dependent on the type and quantity of the heritage language input that these speakers receive (see Flores and Rinke 2020a, for adult heritage EP speakers in Germany; Nagy 2016, for adult heritage Cantonese, Italian, and Russian speakers in Canada).

A similar effect of lexical ambiguity leading to variability could be observed in a cloze test study with children acquiring EP as a heritage language in Switzerland (Rinke, Flores, and Torregrossa 2024; Torregrossa, Flores, and Rinke 2023). In EP, the lexical item *que* can have a number of different functions, e.g. it may function as a complementizer of a complement clause (12a), a relative pronoun (12b), or a consecutive adverbial conjunction (12c).

(12) a. Ele pensa **que** pode ir buscar um balão para a sua amiga.
he thinks that can go bring a balloon for the his friend
'He thinks that he can bring a balloon for his friend.'

b. Mas sem querer, o coelhinho larga o balão, **que** voa para longe.
but without wanting the rabbit releases the balloon that flies to far away
'Without wanting it, the rabbit releases the balloon that flies away.'

c. A cadelinha está tão zangada **que** começa a gritar (…).
the little dog is so angry that starts to shout
'The little dog is so angry that she starts to shout (…).'
(Rinke, Flores, and Torregrossa 2024)

The examples in (12a–c) served as stimuli for a cloze test conducted with 180 heritage children in Switzerland (60 EP-German, 60 EP-French, and 60 EP-Italian bilinguals). The test included 40 gaps with a variety of structures (eliciting nominal morphology, verbal morphology, prepositions, complementizers,

(clitic) pronouns, articles, etc.). In the examples given in (12a–c), the functional element *que* represented the gap that should be recovered by the children. Interestingly, the gaps including *que* belonged to the most demanding structures in the test (besides *clitics*) and the children showed a lot of difficulties in recovering these gaps, particularly in the sentences ((12b), subject relative pronoun) and ((12c), consecutive adverbial conjunction). The difficulties that the bilingual children showed with respect to *que* in those structures are mirrored in monolingual acquisition, for which it has been shown that *que* as a complementizer (12a) is acquired relatively early (Santos 2017), whereas the relative pronoun *que* (12b) is relatively late acquired (Soares 1998) and that the same holds true for adverbial subordinate clauses in general.

Once again, this shows that the variation found in the heritage speakers is not random but that their difficulties with the ambiguous element *que* are expected and relate particularly to areas in which the acquisition and mastering of *que* is also a complex task for monolingual children. We believe that lexical ambiguity is one of the factors that contributes to this high complexity and the resulting variation.

2.5 Conclusions

In the last two decades, heritage speakers and their language competences have become a central research target in the study of bilingualism. On the one hand, this is due to increasing migration movements across the globe and its social and pedagogical impact on migrant families and their host countries. Studying heritage language development gives us fundamental insights into the social and pedagogical needs of heritage speakers (Aalberse, Backus, and Muysken 2019). On the other hand, and as we attempt to show in this chapter, studying heritage grammars also gives us deeper theoretical insights into the nature of human grammatical systems more generally, in particular into the systematicity of linguistic variation and change. As Lohndal (2020) claims in his commentary to Polinsky and Scontras (2020), "(i)f we have sufficient information about the variables that determine heritage grammars, their interrelationship, and their weighting, it should be possible to predict a speaker's grammar." One important endeavor in identifying these variables consists in describing linguistic variation in the baseline grammar and the multiple sources of this variation. We predict that the linguistic structures that

are variable in the baseline grammar will be subject to higher degrees of variation in the heritage language – independently of the contact language. We have attempted to show that this holds true for a range of linguistic structures in EP: clitic placement, the realization or omission of direct objects, null and overt subjects, and structures involving the functional word *que*. The source of variation underlying these structures is, however, not necessarily of the same kind. It may derive from variation present in non-standard registers, from ongoing processes of language change, or from structures that are per se ambiguous, making the learning path under reduced input more challenging. Disentangling the different sources of variation is, however, not always possible. For instance, register variation in clitic placement in some subordinate clauses (which require proclisis) may derive from ambiguous input itself (e.g., the existence of infinitival clauses with enclisis).

In sum, we propose that the particularities of heritage grammars cannot easily be ascribed to deviance or lack of linguistic knowledge unless it can be shown that it is not the predictable outcome of variation and variability in the input. Apparent deviance detected in heritage languages is often just apparent and systematically limited by grammar-internal constraints, which can be of a different nature.

PART II
SENTENCE STRUCTURE

3

Heritage language gaps

Maria Polinsky

3.1 Introduction

Researchers of heritage languages (HLs) seek to understand and model the linguistic representations (mental grammars) of heritage speakers. In pursuing this goal, we compare a particular HL to its baseline, the language of exposure that constituted the heritage speakers' input. This comparison should eventually allow us to understand the main mechanisms driving the construction of mental grammars in natural language.

To discuss the nature of underlying representations in HLs, I will focus here on the alternation between overt and silent elements in HLs. In a nutshell, the problem is as follows: as compared to the baseline, HL speakers have difficulty in detecting and interpreting silent grammatical elements. Thus, their comprehension reflects what Laleko and Polinsky (2017) characterize as the Silent Problem. Similar problems are observed in production, but at the same time, the production of HLs can be characterized by a number of missing elements, some of which are not attested in the baseline. Most commonly such omissions in production are found in morphophonology or inflectional morphology, but free-standing clausal constituents also seem to be affected.

Before I go into the details of the Silent Problem, let me present some introductory remarks that have to do with the *inventory of silent elements* and with the *divide between comprehension and production*.

Not all missing material is the same from a theoretical standpoint. Some silent material arises under deletion of individual segments which form part of a movement chain (using the traditional terminology, I will refer to such deleted material as a *trace*). Likewise, the material deleted under identity in ellipsis results in silence, and as I have shown elsewhere, such instances of ellipsis undergo reanalysis in HL (Polinsky 2016, 2018: 263–268). Silent elements can also correspond to dedicated lexical items: null pronominals, which

Maria Polinsky, *Heritage language gaps*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro et al., Oxford University Press. © Maria Polinsky (2025). DOI: 10.1093/9780191987731.003.0004

constitute part of the lexical inventory of a given language. Next, we recognize implied missing subjects of control clauses; depending on the theoretical outlook and particular analysis, these are identified either as PRO, which is a universal lexical category without a corresponding overt counterpart; as material deleted under movement, or as a null pronominal (see Landau 2013 for an overview). In addition to phrasal silent elements, there are also empty (functional) heads, which I will not consider here.

In what follows, I will concentrate on null pronouns (*pro*) in subject and object position, and on gaps in relative clauses. The latter can be interpreted as material deleted under A-bar movement or as null pronouns, depending on the language and analysis of the relevant relative clause (see Bianchi 2002, Cinque 2020, de Vries 2018 for overviews of prominent analyses).

In preparing this chapter, I was astonished to see that as a field, we are missing quite a few data points with respect to HLs (and corresponding baselines). Crucially, as we proceed to fill the empirical gaps in HLs, a key strategy should be to use phenomena rather than languages as a starting point. Volumes of work have been produced on individual HLs and comparisons between them, but it is not uncommon for us (present company included) to lose track of what these empirical observations are supposed to show. Knowing exactly what we do not know is the first step toward plugging information gaps, and it is my hope that some of the points made below can serve as goalposts for future work.

Turning to language performance, linguistic knowledge is deployed in two main ways: people use language to convey their thoughts to others or to themselves (production) and to understand what is said by others (comprehension). In both situations, they must assemble the building blocks of language in a systematic fashion. Accordingly, as we evaluate HL data, it is sensible to combine the data from HL production and comprehension. Of course, this brings in additional complications for a researcher. On the practical plane, more data have to be collected (and indeed, most of the existing work on HLs focuses exclusively on just one of the two aspects of language knowledge and use). On the theoretical plane, paying attention to production and comprehension alike commits one to taking a stand on the relationship between the production (generation) and comprehension (parsing) of language. That relationship is a complex problem, one that has been at the forefront of linguistic theorizing, especially as experimental work feeding back into linguistic theory has become a norm in linguistics (see Momma and Phillips 2018 for an overview). Roughly, we can identify two main approaches to production and comprehension:

(i) Production mechanisms and comprehension mechanisms are distinct, but they interact heavily during a single act of understanding or speaking (see Huettig and Hartsuiker 2010 for an overview).
(ii) Production mechanisms and comprehension mechanisms share the same underlying representations (Momma and Phillips 2018).

Approach (ii) represents a more nuanced variation on what is commonly referred to as a *single-mechanism view*. Note that both types of approaches, (i) and (ii), allow for interaction between production and comprehension but differ in the understanding of the root mechanisms of these two processes. Here, I adopt approach (ii), both in light of growing evidence in its favor and because of analytical parsimony: unless we have evidence to the contrary, it is more economical to assume that production and comprehension are based on the same mechanisms of structure building.

At the same time, there are interesting differences between production and comprehension in general and in HLs in particular (see, for example, Bowles 2011, Frasson et al. 2021, Montrul et al. 2008, with respect to silent elements). Taken at face value, such differences can be evidence in support of approach (i), which implies differences between comprehension and production. If so, such discrepancies in data need to be examined closely.

The structure of the chapter is as follows. In Section 3.2, I consider the use of null pronouns in HLs, and in Section 3.3, I examine gaps (or absence thereof) in relative clauses. In considering both types of silent elements, I distinguish between comprehension and production data. In Section 3.4, I offer my main conclusions and identify areas where further work is needed.

3.2 Null and overt pronouns

3.2.1 Comprehension

The fate of null pronouns, primarily in the subject position, has been a popular topic in bilingual research. Null pronouns need to be licensed and identified/recovered (Rizzi 1986b). *Licensing* ties a null pronoun to a particular structural position, which in principle could include the subject or object position. *Recovery* allows the reference of a null pronoun to be identified in one of two ways: (i) through agreement (in a morphologically rich language; see Rizzi 1986b and subsequent literature), or (ii) through an interpretive rule linking the null element with a current discourse topic, especially in languages like

Mandarin or Japanese (compare Huang 1984: 550: "there is a rule of coindexation, in the discourse grammar of a discourse-oriented language, which coindexes an empty topic … with an appropriate preceding topic").

Since null pronouns do not fully disappear from HLs (an issue I will return to below), the question arises as to whether "bilinguals have native-like knowledge of the syntax of null and overt subjects more generally" (Keating et al. 2016: 39), that is, whether bilinguals retain the baseline licensing and recovery conditions for null arguments. In considering this question, researchers often ask whether the null/overt contrast is sufficiently represented in the input to a given heritage language and whether heritage speakers of that language actually produce null forms.

Few studies address the question of whether bilinguals maintain the contrast between overt and null pronouns (many researchers concentrate on possible biases in the linking of a pronoun to an antecedent). Keating et al. (2011, 2016) asked whether bilinguals maintain the null/overt contrast and respond differently to null versus overt pronouns during online processing. Using a series of comprehension experiments, they investigated this question for heritage speakers of Spanish, a full pro-drop language, who are dominant in American English. They examine the bilinguals' knowledge of impersonal null pronouns as well as their knowledge of the contrast between referential null and overt pronouns. In comparing their bilinguals to the Spanish baseline, they conclude that "both groups were similar in that they responded differently to null versus overt pronouns during online processing" (Keating et al. 2016: 45). In other words, the bilinguals' lexicon still includes null pronominals, arguably with the same specifications and licensing requirements as in the monolingual lexicon.

Rinke and Flores (2018) examined the interpretation of null/over subjects by German-Portuguese and Spanish-Portuguese bilingual children (thus, heritage speakers of Portuguese living in Germany and in Andorra respectively).[1] Their results are similar to the ones by Keating and co-authors, with the main conclusion being that heritage speakers had the same specifications and licensing requirements as in the baseline.

In a similar vein, Dubinina and Polinsky (2013) noted sensitivity to the null/overt pronoun contrast in the subject position for heritage speakers of Russian. The preservation of this contrast is particularly striking because in Russian, unlike in Spanish, pro-drop is limited and occurs only in certain embedded clauses, something which is unlikely to be common in the input.

[1] In this chapter, I discuss some data from European Portuguese and Brazilian Portuguese. Unless noted otherwise, *Portuguese* refers to the European variety.

Laleko and Polinsky (2017) noted that Heritage Korean speakers also maintain the null/overt pronoun contrast. Thus, although the empirical data are still limited, the balance of evidence suggests that the null/overt contrast is maintained in both rich-agreement pro-drop languages (Spanish) and languages with topic-determined pro-drop (Korean).

Given that null pronominals are in principle available in the heritage grammar, how are they recovered and identified? With respect to null and overt pronouns, interpretive biases have been explored extensively, in both the theoretical and psycholinguistic literature (see Keating et al. 2016 for an overview). The overall conclusion is that monolinguals preferentially link null pronouns to subject antecedents and link overt pronouns to antecedents in lower structural positions (Carminati 2002, 2005; Keating et al. 2011, 2016). Note that there are two correlations here: one is between subjects as antecedents of null pronouns, and the other between non-subjects as antecedents of overt pronouns.

To illustrate, in the Spanish example below, the null pronoun in the second clause is preferentially interpreted as referring to the subject, and the overt pronoun *él*, as referring to the object:[2]

(1)	a.	Juan$_i$	pegó a	Pedro$_k$.	*pro*$_{i>k}$	está	enfadado.	[Spanish]
		Juan	hit PRP	Pedro		be.PRS.3SG	angry.M	
	b.	Juan$_i$	pegó a	Pedro$_k$.	Él$_{k>i}$	está	enfadado.	
		Juan	hit PRP	Pedro	He	be.PRS.3SG	angry.M	

'Juan hit Pedro. He is angry.' (Keating et al. 2016: 38)

However, it is worth noting that trends observed in the baseline are only preferences, not a hard rule. Depending on the test, these preferences are observed in 50–75% of cases (Carminati 2005; Keating et al. 2016, and references therein).

How do heritage speakers deal with these preferences? It appears that the two-way association between subjects as antecedents of null pronouns and non-subjects as antecedents of overt pronouns is not pronounced in HLs (compared to the baseline). In their extensive comparison between monolingual Mexican Spanish speakers and English-dominant bilinguals, Keating et al. (2016) observed that "only the monolingual group showed a referential bias with overt pronouns," linking these pronouns to non-subject antecedents (Keating et al. 2016: 45).

[2] The interpretive preferences are indicated by subscripts.

Heritage speakers appear to generalize the recovery strategy that links any pronoun, whether null or overt, to the highest structural position available. To reformulate the hypothesis proposed by Carminati (2002: 57) for Italian, we can then observe the following principle:

(2) *Position of antecedent strategy (PAS)*
Choose the highest structural argument as the antecedent.

Three factors are likely to play a role in the observed generalization of PAS by heritage speakers: (i) the reduction in the scope and use null pronouns in HLs; (ii) the *scalar nature* of PAS, and (iii) the *long-distance nature* of the anaphoric dependency.

It is not a big surprise that heritage speakers may have more restricted lexical resources than baseline speakers. In particular, their inventory of null pronominals seems to be more limited than the inventory in the corresponding baseline. On the flip side, looking at the heritage grammar, we observe some changes in agreement, both on goals and probes. In terms of the content of phi-features, their array is less rich than in the baseline and/or they may become more opaque (see Scontras et al. 2018 which shows that number and gender features get bundled together in heritage Spanish, but not in the baseline). It may also be the case that fewer structural elements probe for phi-features in HLs, as compared to the baseline, which again contributes to a more restricted set of agreeing elements. These changes in agreement contribute to the decline in null pronominal elements, at least in rich-agreement languages. The result is the reduced range of null pronominals in an HL, and ultimately the loss of these elements.

Perceptual salience, or rather lack thereof, may also contribute to these putative changes in the grammatical system. Functional material is often unstressed and low in perceptual salience. If heritage speakers do not perceive functional elements, they may also produce them less and less, which is yet another pathway to the loss of agreement. To replace the material that is low in perceptual salience, heritage speakers may start producing more overt forms. The combination of licensing restrictions and rearrangement of vocabulary items in favor of the more salient elements creates a perfect storm that leads to changes in the agreement system and the abundance of overt pronouns.[3]

Let us now turn to factor (ii). In the baseline languages, the linking of overt pronouns to non-subject antecedents is a preference, not a categorical

[3] I am grateful to Tanja Kupisch for suggesting to me the role of salience in this instance.

rule. Heritage speakers disfavor scalar principles more generally, not just in application to PAS. This aversion to scalar principles in HL gets realized as the elimination of privative oppositions in favor of equipollent ones. In a privative opposition, one of the members is fully specified, and the other either lacks the relevant feature or does not have it fully specified (while the absence of a feature vs. lack of specification is generally important, it is not crucial for the discussion here). In an equipollent opposition, both members of a grammatical opposition are equally specified (see Polinsky 2018 for an overview, and Laleko 2011, 2015, for the elimination of privative oppositions in favor of equipollent ones in the aspectual grammar of HLs). It is tempting to view the difference between the two types of opposition in terms of economy. Full specification appears to be less economical, but at the same time, lack of specification leads to greater uncertainty, something that heritage grammars tend to avoid.[4] This avoidance of uncertainty underlies the switch from scalar to discrete principles. If this is on the right track, the erosion of the scalar linking principle associated with pronouns is unsurprising.

With respect to factor (iii) above, associating pronouns (null or overt) with discourse antecedents requires significant processing effort. This ties into a general challenge for heritage speakers. Lower proficiency heritage speakers (and other speakers differing from the monolingual baseline, as explored in the work of Sorace and her colleagues) have problems connecting elements of structure that are separated by distance. If the lexical content is more clearly specified, that allows for an easier construction of long-distance dependencies. Here I would like to underscore that factoring in processing difficulty requires a transition from processing to grammar. Building long-distance dependencies may start out as a challenge in real-time processing only, but if strategies deployed to deal with such a task become entrenched, they become part of the new grammar. How such entrenchment happens needs to be worked out, otherwise the connection between processing effort and grammar is merely a desideratum.

In sum, the problems heritage speakers encounter with null subjects arise from the confluence of three factors: difficulty with elements that are low in perceptual salience – nothing is perceptually lower than a silent element;

[4] A question may immediately arise here: how can we reconcile the avoidance of underspecification with the lack of full feature specification on some probes that I discussed in relation to factor (i)? There is no contradiction here, because the lack of full feature specification on probes eventually leads to the loss of particular features, something that can be amply observed in the loss of gender features or clusivity (which, according to Harbour 2016, is a privative feature, hence another challenge to heritage speakers) under contact.

difficulty with scalar representations; and difficulty integrating linguistic material at a distance.

In the next section, I will examine the influence of these (and additional) factors in a different grammatical domain: relative clauses. But before I do that, let me also consider the comprehension of null objects.[5] Superficially, languages with null objects can be divided into two types: languages where null objects alternate with overt noun phrases and clitic pronouns and languages without object clitics, where the alternation is only between overt noun phrases and null objects. Structurally, not all null objects are the same, and they may receive different analyses (see Flores et al. 2020, Rinke et al. 2019 for a short overview).

The comprehension of null objects in HLs has not received much attention. Chou et al. (2020) stands out as a study where such comprehension was closely investigated with an eye to a comparison between monolingual and English-dominant heritage speakers of Mandarin. In their picture-verification experiment, they provided subjects with a context and then asked them to evaluate if the test sentence, with a transitive predicate and a null object, appropriately described the context. Consider the following test stimuli:

(3) a. Context sentence (with accompanying picture):
'It is a cold day. Cuiyi has just received a box in the mail.'

b. Test sentence (with accompanying picture):

Cuiyi	yaohuang-le	*ec*	yixia.	[Mandarin]
Cuiyi	shake-PRF		once	

'Cuiyi shook (the box).' (Chou et al. 2020: 234–235)

The authors did not find any difference between monolinguals and heritage speakers; both groups were able to correctly interpret the null objects in context. However, as the authors themselves point out, "being able to accept null objects in contextually appropriate contexts is different from being able to reject null objects in prohibited contexts" (Chou et al. 2020: 252). The latter include, for example, situations of A-binding, where Mandarin null objects are impossible in the baseline, possibly for minimality reasons:

[5] I will only consider definite null objects here, since the licensing and interpretation of indefinite (weak) null objects is less clear to begin with.

(4) Zhangsan$_i$ shuo [Lisi$_j$ bu renshi $ec_{k/*i/*j}$] [Mandarin]
Zhangsan say Lisi not know

'Zhangsan$_i$ said that Lisi$_j$ does not know (him$_{k/*i/*j}$).' (Chou et al. 2020: 252)

At this stage, it is not known whether baseline and heritage speakers respond differently to such structures. If both baseline and heritage speakers have the same interpretation of (4), that may be a signal of the shared grammar of null objects, and concomitantly, of the fact that heritage Mandarin allows such null objects. If heritage speakers are accepting of sentences like (4), and at the same time are accepting of null objects more generally, that would be an indication of a different grammatical representation in heritage Mandarin.

3.2.2 Production

In principle, monolinguals and bilinguals may differ in both their licensing and recovery of null pronouns, but there is no evidence that the licensing conditions differ across the two populations. In particular, we do not find null pronouns in adult HLs appearing in structural positions that are different from the structural positions found in the baseline. As there is no evidence of different licensing conditions in HLs, compared to the baseline, by elimination, all the issues in the variability of null pronouns have to do with interpretation and use. The general intuition is as follows: in cases where an overt form alternates with a null one, the overt form is easier to interpret, which may make it preferable. That would explain the greater number of overt expressions in HL in those positions where the baseline variety has null ones.

However, the preference is not categorical; heritage speakers still drop overt elements, and the absence of categorical differences between the baseline and HLs is a particularly noteworthy characteristic of heritage production. Further still, as I mentioned in the introduction, some instances of heritage production are ridden with omissions, both at the morphological and even clausal level, which creates an intriguing tension between the apparent need for an overabundance of overt material for comprehension and the observed dearth of overt material in production.

Numerous studies document the increased use of overt subjects in pro-drop languages, like Spanish and Italian, when these languages come into contact with non-pro-drop languages, such as English or German (Lipski 1996; Montrul 2002, 2008, 2016; Müller and Hulk 2001; Otheguy et al. 2007;

Serratrice 2007; Serratrice et al. 2004; Silva-Corvalán 1994; Sorace and Filiaci 2006; Sorace et al. 2009; Tsimpli et al. 2003, 2004; among others). This is not surprising, as overt subjects are not ungrammatical in the baseline, and they allow heritage speakers to avoid any ambiguity in production. The overall use of these subjects (which is never categorical) is consistent with the general tendency for overmarking observed in heritage language (see Polinsky 2018 for an overview).

The use of null pronouns may already be diminished in the speech of first-generation immigrants, whose language serves as input to heritage speakers, although the empirical data vary. Several researchers document a smaller proportion of null subjects in first-generation immigrant Spanish (Montrul 2016; Otheguy et al. 2007; Raña-Risso and Barrera-Tobón 2018). For Spanish, one could argue that the overuse of overt subjects is due to transfer from English (Otheguy et al. 2007). On the other hand, in their production study of two generations of Portuguese speakers in Germany, Flores and Rinke (2020b) did not find any decline in the use of null subjects by first-generation immigrants from Portugal. Note that in both cases the societally dominant language (English, German) is not *pro*-drop, but it is hard to estimate how much of the dominant language is actually used by first-generation immigrants in each case study. Some first-generation immigrants may have very low proficiency in the societal language, and if so, transfer from that language is unlikely. Additionally, immigrant speakers converse with interlocutors of different backgrounds and may feel uncertain about how well they are understood, which may lead to the need for explicitness. The use of overt pronouns is a particular case of explicitness, and this speech behavior may become entrenched.

A more telling test case for the emergence of overt pronouns in heritage language would be a setting not driven by transfer: a bilingual environment where both the dominant and baseline language allow null arguments, particularly null subjects. The Spanish-Catalan dyad presents such a case: de Prada Pérez (2009, 2019) examined the contact between these two languages in Minorca and found no significant difference in the rate of overt subjects among Spanish monolinguals, Catalan monolinguals, Spanish-dominant bilinguals, and Catalan-dominant bilinguals. She did, however, observe a difference between the omission of first- and third-person pronouns; the bilinguals in her study had about the same rate of null first-person pronouns accompanied by a decline in pro-drop with third-person pronouns, an issue I will revisit shortly.

The case I now turn to, heritage Friulian, is more nuanced. Heritage Friulian is a northern Italian variety spoken in Latin America (Frasson et al. 2021). Frasson and colleagues examine Friulian as spoken by first-generation immigrants and next-generation heritage speakers in Argentina and Brazil.

While Brazilian Portuguese has restricted pro-drop, Spanish is a fully pro-drop language, so it is not unreasonable to expect the growth in pro-drop in the Friulian of Argentina, supported by both languages in the dyad.

Heritage Friulian presents an intricate pattern, one that requires a separate consideration of each person. Second-person pronouns do not drop at all, and are in fact excluded from consideration by Frasson and colleagues. In the third person, pro-drop seems to be more restricted than in the baseline, and lexical subjects are widely expressed – a predictable pattern consistent with the other observations presented here.

In the first person (singular and plural), however, overt subject pronouns appear to be omitted more than in the homeland variety. This result may appear surprising, on a closer investigation, what appears to be the rise of pro-drop is actually due to the realignment in the pronominal system. As Frasson et al. (2021) argue, subject clitics in the Friulian immigrant varieties have been changing from agreement markers to weak pronouns.[6] If a subject clitic is in fact reanalyzed as a pronoun, it is expected to be infelicitous in combination with another pronoun.

The result is a new system with three options for first-person pronouns: strong tonic pronouns (possibly used contrastively); weak overt pronouns, the ones that develop from clitics, and regular null pronouns. Strong pronouns (*jo* for 1sg, *on/noaltris* for 1pl) are not used together with weak overt pronouns. However, the drop of a strong pronoun does not entail pro-drop; instead, that space is taken by overt pronouns of a different type.

On a more general level, the Friulian data show that the change in the heritage pronominal system does not proceed at the same pace in all grammatical persons. In the heritage Friulian case, the development of a new first-person pronoun is ahead of potential changes in the other persons. This is important for considerations of pro-drop in other heritage languages, where more fine-grained changes should at the very least be expected. As I already mentioned, the examination of Spanish-Catalan bilinguals also points to staggered development, with a stronger decline in third-person pro-drop as compared to first-person pro-drop, where more fine-grained pronominal distinctions can be observed (de Prada Pérez 2019).

Since the data from heritage Friulian and Spanish-Catalan bilinguals are based on production, it remains to be seen if the staggered reanalysis of the pronominal system is also observed in comprehension.

[6] Such a change is consistent with the characteristic growth in analyticity observed across different HLs. It also constitutes a particular way of sidestepping the problems with inflectional morphology that are typical in HLs as well. It is common to see inflectional morphology just disappear, but Friulian offers an example of reanalysis, not loss.

Another important test case would be a bilingual dyad where the dominant language is pro-drop and the heritage language is not. If heritage languages avoid null arguments as an overarching rule, in this context, there should not be any transfer effect from the dominant language to the heritage language. Heritage English (in contact with Brazilian Portuguese, Hebrew, or Japanese) or heritage Russian (in contact with Greek, Turkish, or Hebrew) could offer such test cases. Heritage English speakers in Japan or Israel, dominant in Japanese and Hebrew respectively, two pro-drop languages, do not show any transfer effects with respect to pro-drop, and they correctly produce overt subjects (Polinsky 2018), but our empirical knowledge of the relevant dyad is still scant.

Turning to the production of null objects, the data are rather scarce, just as they were with the comprehension of this category. Rinke et al. (2019) studied the production of null objects in heritage Polish and heritage Portuguese as spoken by German-dominant school children; for heritage Portuguese, Rinke et al. (2018) also conducted a corpus study of spontaneous speech. In all the instances, they actually found more object omissions than in the baseline, contrary to the trends I have reviewed so far. One of the reasons behind this pattern may have to be as follows: null objects in Portuguese are highly sensitive to discourse factors, and since both studies targeted spontaneous connected discourse, rather than eliciting production without much context, this may have affected the results (the speakers had a set of particular referents in mind and omitted them more freely).[7]

3.3 Relative clauses: Gaps or no gaps?

Studies of relative clauses in monolingual and bilingual speakers have been no less popular than studies of null subjects. While relative clauses come in a number of flavors, attention in bilingual studies has mainly been paid to *restrictive externally-headed relative clauses*, whose internal gap corresponds to the head noun of the clause, as in the example below:

(5) the girl [RELATIVE CLAUSE that the boy photographed ___ near the Smithsonian]
HEAD NOUN GAP

[7] I am grateful to Cristina Flores and Esther Rinke for pointing out the relevance of discourse factors to me.

The structure of relative clauses is not uniform across languages (see Bhatt 2002; Bianchi 2002; Cinque 2020; Hulsey and Sauerland 2006; Keenan and Comrie 1977; Lehmann 1984; Reinhart 1998; among others). The first cut one could make is in terms of the syntactic analysis: between those relative clauses whose gap results from movement and those where the gap is a null pronominal bound by an operator; I will address some diagnostics of these analyses below. The use of a null pronominal (as opposed to a copy deleted under movement) is a particularly likely option in languages that independently have null pronominals in their inventory (consider Matsumoto 1989 or Reinhart 1998, scholars who approach the issue from very different perspectives but nevertheless reach similar conclusions).

Relative clauses are relevant for processing studies because they force the speaker to create a connection between the antecedent and the gap at a distance. The long-distance dependency established in relative clauses is potentially challenging from two different standpoints: relating elements of structure (including the silent ones) and a general memory load.

3.3.1 Comprehension

3.3.1.1 Current state of knowledge

In numerous studies, using different techniques and sampling different populations, researchers have shown that the comprehension of relative clauses with a subject gap (subject relatives; SR) is more accurate and faster than the comprehension of those with an object gap (object relatives; OR);[8] consider Lau and Tanaka (2021) for a recent overview of the subject advantage and all the factors that may contribute to it – their number only keeps growing.

Subject advantage is quite visible in the comprehension of relative clauses by HL speakers (and also by L2 learners and some child populations). O'Grady et al. (2001) investigated Korean heritage speakers' comprehension of subject and object relative clauses, comparing the heritage speakers (16 learners in an accelerated second-semester Korean course at a US university) to 45 non-heritage L2 learners of Korean in second-semester and fourth-semester courses at the same university. Participants completed a picture-selection task

[8] The gap in a relative clause may correspond to other constituents, not just subject or object (see Keenan and Comrie 1977), but there has been little comparative work including other types (but see, for example, Diessel and Tomasello 2005, who do consider other types). In experimental studies, the main emphasis has been on comparing SRs and ORs, and other gap positions are still awaiting investigation.

in which they had to circle the picture corresponding to the sentence they heard; all the pictures represented reversible actions. The comparison was between subject relatives, (6a), and object relatives, (6b):

(6) a. SR [Korean]

[___$_i$ namca-lul cohaha-nun] yeca$_i$
man-ACC like-ADN woman

'the woman who likes the man'

b. OR

[namca-ka ___$_i$ cohaha-nun] yeca$_i$
man-NOM like-ADN woman

'the woman whom the man likes'

HL speakers and L2 learners performed similarly; everyone did better with subject relatives than object relatives and often misunderstood the latter as the former. Similarly, heritage Russian speakers show a strong subject bias in relative-clause comprehension (Polinsky 2011), regardless of the order in the relative clause (SV or VS).[9] The four possibilities, all of which are licit in Russian, are presented below:

(7) a. SR, VX order [Russian]

sobak-a$_i$ [kotor-aja ___$_i$ ukusila košk-u]
dog-NOM which-NOM bit cat-ACC

b. SR, XV order

sobak-a$_i$ [kotor-aja ___$_i$ košk-u ukusila]
dog-NOM which-NOM cat-ACC bit

'the dog that bit the cat'

(8) a. OR, VX order [Russian]

sobak-a$_i$ [kotor-uju ukusila ___$_i$ košk-a]
dog-NOM which-ACC bit cat-NOM

b. OR, XV order

sobak-a$_i$ [kotor-uju košk-a ukusila ___$_i$]
dog-NOM which-ACC cat-NOM bit

'the dog that the cat bit'

[9] The syntactic derivation of these word orders is complex, but the details are not critical to the points made here. See Polinsky (2011) for more details.

The native-speaker adults, monolingual children, and bilingual children (average age about 7;0) performed extremely well on all types of relative clauses, surpassing 90% accuracy on each. Difficulties appeared only in the heritage-adult group; these speakers did well with subject relatives but performed below chance on object relatives with both word orders. The fact that bilingual children performed indistinguishably from their monolingual peers suggests that the performance by adult heritage speakers cannot be due to a fossilization of the child language. Instead, this is a striking instance of reorganization later in life.

Transfer from English cannot explain the data, since the heritage Russian speakers performed equally well (or equally poorly) on all word orders and did not prefer orders that mirror English (i.e., subject relatives in VO order and object relatives in SV order). Instead, the explanation seems to follow from the intersection of two effects: First, heritage speakers seem to lose their sensitivity to case morphology, the essential clue that guides native speakers (and, arguably, bilingual children) to determine who did what to whom. This decline in sensitivity to morphological distinctions mirrors the general neglect of inflectional morphology so typical of HLs (Lohndal 2021; Putnam et al. 2021). Second, in the absence of case-differentiating morphology, the universal preference for the subject relative interpretation, which can be interpreted as a version of PAS (presented above in (2)) kicks in, causing heritage speakers to perform perfectly on subject relatives and below chance on object relatives. The result is a rearrangement of relative-clause syntax in heritage Russian: it remains in the heritage grammar, but it is limited to the highest structural position in the clause.

In heritage Spanish, the subject advantage is also present, but with an additional effect from word order: ORs with the embedded subject in preverbal position are easier to comprehend than those with the embedded subject in the postverbal position, both for children and adults (Sánchez-Walker 2012), and this pattern seems to mirror what is observed in the comprehension of relative clauses by child language acquirers (e.g., Prévost 2009; Sevcenco et al. 2013).

The target-like comprehension of relative clauses by bilingual children has also been noted for Turkish (see Marinis and Özge 2020, on Turkish–English bilinguals, age range 4;7–9;2) and Mandarin (see Jia and Paradis 2018, on Mandarin Chinese–English bilinguals, age 6;7).[10] In other words, the bilingual

[10] In trilingual children, however, the results are not as even; for example, Chan et al. (2017) show that Mandarin–English–Cantonese speaking children have difficulties identifying non-SRs in Cantonese, despite the slight preference for ORs demonstrated by monolingual Cantonese children (Chan et al. 2018).

children are similar to the monolinguals (whose data I will discuss shortly later in this section). But adult heritage speakers seem to deviate from this target-like comprehension.

Although more data are needed, these comprehension results point to a coherent picture, one that is consistent with the observations about coreference presented in Section 3.2.1. We again see a strong preference for interpreting the gap in the relative clause as the subject, and this preference is stronger than what is found in the baseline. This is what transpires with the data on heritage Russian (Polinsky 2011) and heritage Korean; heritage Spanish data show more variation (Sánchez-Walker 2012), which could be related to differences in the proficiency and HL exposure levels within the study pool.

Another factor that may determine the preference for the subject interpretation deserves mention: literacy, which in turn may best be understood as the stand-in for quality of exposure. Heritage speakers rarely have experience with literacy (assuming the relevant language has written registers at all), and they often lack schooling in their HL. Complex structures, including relative clauses, are of course more common in written language (or in more elaborate registers if no written language exists for the baseline). "To the extent that increased exposure results in stronger representations, speakers with more exposure to written language might be expected to acquire stronger representations of these more complex structures" (Dąbrowska et al. 2022).

While there is no arguing that literacy/exposure plays an important role in the setting of grammatical representations, it is worth noting that the SR/OR distinction is well acquired in pre-literate children; their production may be skewed in favor of SRs, a point that I will revisit in the section below, but the comprehension of relative clauses is quite target-like around age 4;0 in monolinguals. For instance, Diessel and Tomasello (2005) consider relative clauses in child English and German and show that the main difficulties arise with the relativization of indirect objects and possessors. Hu et al. (2016) show target-like performance on Mandarin relative clauses among monolingual children, and Chan et al. (2018) document a slight preference for ORs in monolingual Cantonese children. Recall also the data from bilingual children addressed earlier in this section; bilingual children, who lack literacy in the home language, perform at ceiling in the comprehension of SRs and ORs. Thus, exposure-qua-literacy is an unlikely culprit in the deficits in relative clauses that we observe in HL speaking adults.

3.3.1.2 Gaps in our knowledge

One of the analytical uncertainties concerning relativization in HLs has to do with the nature of the gap in the relative clause and its relationship to the head noun and relative pronoun (if available). On the analytical level, we can recognize two main strategies of relative-clause formation, one with a syntactic dependency (created by movement), the other with an anaphoric dependency between the head noun, the binding operator, and the position of the relativized constituent in the relative clause.

Simplifying things somewhat, the movement analysis of an externally-headed relative clause looks like this: An operator or the pronominal representation of the head is fronted and forms a movement chain with the element in the base position; the head noun is coindexed with these but it is generated externally to the relative clause:[11]

(9) Movement analysis of relative clause

[DP DPHEAD-NOUN [CP Relative-pronoun$_i$/Op$_i$ (C)[TP … *trace*$_i$ ….]]]

coindexation movement

This analysis is not the only one that posits movement; the head-raising analysis of relative clauses provides an alternative (Bianchi 1999; Cardoso 2017; De Vries 2002; Kayne 1994; among others). On that analysis, the head noun is generated inside the relative clause (in its thematic base position) and then undergoes A-bar movement to the CP-domain; the relative pronoun, overt or abstract, is an operator, just as in (9). Next, the head NP is moved to its surface position left of the relative pronoun, where it can be associated with the external determiner.

For the purposes of this discussion, the difference between the A-bar movement analysis and the A-movement analysis is not critical; therefore, I will refer to this family of analyses generically as *movement*. Taking this more general view allows for a more productive review of the data, since the two possible analyses retain a number of unresolved issues, and may not even be mutually exclusive (see de Vries 2018; Hulsey and Sauerland 2006).

In addition to the movement analyses, however implemented, another alternative has been proposed: the operator inside the relative clause binds an element in the base position, and these two pieces of structure are coindexed

[11] In the discussion below, I represent relative clauses as adjuncts to the head noun; for the purposes of this chapter, nothing hinges on this particular representation.

with the head noun (which is generated outside the relative clause, as in (9) above); consider Nishigauchi (1986); Reinhart (1998); a.o. Another flavor of this analysis, one that I will refer to as coindexation, is that there is simply coreference between the head noun, the relative pronoun (if any), and the gap or resumptive pronoun in the base position in the relative clause. The schematics for this analysis are as follows:

(10) Coindexation analysis of relative clause

[DP DPHEAD-NOUN [CP Relative-pronoun$_i$/Op$_i$ (C)[TP … *pro*$_i$ ….]]]

coindexation

Both movement and coindexation are compatible with the silent or overt realization of the relativized constituent in the base position inside the relative clause. In other words, this base position can be filled in two different ways: by what we can represent atheoretically as a gap, or by an overt constituent (a resumptive pronoun). The overtness criterion alone does not distinguish between A-bar movement and coindexation; in principle, both links in a syntactic chain can be spelled out, or the lower one can be spelled out exclusively (see Bošković 2007, 2014; Nunes 2004). Likewise, pronouns can come in overt or silent form as well (McCloskey 2017).

Since the differences between movement and coindexation may not be directly visible from the surface structure of a given relative clause, the learner/speaker needs to assess additional evidence in favor of one analysis or the other. Signs of A-bar movement include island sensitivity, reconstruction, parasitic gaps, or weak crossover effects (see McCloskey 2017 for an overview and some other diagnostics). Languages vary with respect to the availability of such effects, so the resulting syntactic accounts often have to rely on extremely subtle data.

Of course, all analyses are as good as the data that they account for and predict, but assuming these competing accounts are on the right track, the following question arises: does each individual HL maintain the syntactic structure of the relative clause that is available in the baseline, or do we observe a reanalysis? For example, relative clauses in the baseline may be built via movement, but does that entail that the corresponding HL is going to keep the same structure?

To the best of my knowledge, this question has not been addressed before. The main options, summarized in Table 3.1, include the preservation of the same strategy as in the baseline (movement, of the A-bar or A-type, or

Table 3.1 Relative-clause syntax, baseline vs. heritage language

Baseline	HL
Movement	Movement (no reanalysis) Coindexation (reanalysis)
Coindexation	Coindexation (no reanalysis) Movement (reanalysis)

coindexation via unselective binding) versus reanalysis. Under each of these options, the copy of the relativized constituent inside the relative clause can be silent or overt (to simplify the exposition, I do not mark the (non-)overtness of the base position in the table).

In addition, the syntax of relative clauses in a particular HL may also be affected by transfer from the bilingual's dominant language. If the home and the dominant language structure their relative clauses differently, it is reasonable to consider structural transfer effects (so-called abstract transfer, as opposed to the transfer of actual linguistic material) whereby the syntax of relative clauses in the dominant language gets imposed on the relative clauses of the corresponding HL. As with the general question raised above, we are simply on unchartered empirical territory here, but that is all the more reason why we should be aware of potential influences on the structure.

We still lack empirical facts needed to evaluate the possibilities outlined here with respect to HLs. Nevertheless, several conceptual and empirical considerations can still be offered. On the conceptual side, three properties are characteristic of HLs. The first of these is the preference for Merge over Move. This finds its reflection in the elimination of movement dependencies, which are often replaced by dependencies whose links are all base-generated and coindexed. At this stage of our knowledge of HLs, this preference for Merge over Move does not go beyond empirical observations; a principled explanation for this preference is still outstanding.

Two other characteristic properties observed in HLs have a rather direct connection to performance limitations, but it is worth noting that although these properties may be rooted in performance, they are represented in grammar. I have already touched upon the strong preference for overt forms over silent ones in HLs. That would favor all kinds of overt elements in relative clauses and could work well with the preference for coindexation between the head noun of a relative clause and an (overt) pronoun inside the relative clause. Accordingly, we may also expect the absence of island effects in such relative clauses. That in turn may skew the representation of relative-clause syntax in favor of coindexation, away from movement. And finally, we often

find that HLs, compared to the baseline, use shorter dependencies, something I return to in Section 3.3.2.2.

Turning to empirical facts, I conducted a pilot study, testing SRs and ORs in baseline and heritage Russian with respect to weak crossover (WCO), which has been claimed as a characteristic property of relative clauses in Russian and, according to some researchers, in English. WCO arises when a variable co-refers with a pronoun that does not c-command it (cf. Lasnik and Stowell 1991; Postal 1993; Ruys 2000; Safir 1984, 1986, 1996; a.o.). The reasoning behind the pilot study is as follows. If speakers are sensitive to WCO violations in relative clauses that would indicate that their grammar of relative clauses likely involves A-bar movement (other A-bar movement effects, such as reconstruction and parasitic gaps, are not observed in Russian). The absence of WCO effects in the respective relative clauses can either point to the suspension of crossover effects (as in Lasnik and Stowell 1991) or to the absence of movement. Although the absence of WCO effects may be equivocal, we stand to gain from comparing the baseline group and the heritage group; if the two groups are different with respect to WCO, that may help us understand the relevant underlying representations.

The choice of English as the dominant language for heritage speakers is also important; English shows WCO effects,[12] so we can expect that effect to be amplified in heritage Russian under abstract/structural transfer.

The following examples illustrate WCO in English restrictive relative clauses (where the relativized noun phrase cannot cross over the pronoun with which it is coindexed, in this case, a subject possessive pronoun), (11a), and in a *wh*-question, where the *wh*-expression cannot cross over the coindexed possessive pronoun, (11b):[13]

(11) a. *the kid$_i$ [who his$_i$ sister called ___$_i$ a moron] (Postal 1993: 540)

b. *Who$_i$ does his$_i$ mother love __$_i$?

WCO effects in Slavic, Russian in particular, are easy to observe in *wh*-questions, where the violation is very similar to what we see in the English examples such as (11b). Consider:

[12] As I already alluded to, this is a simplification. While WCO effects in *wh*-questions are robust, researchers disagree as to whether WCO effects in English relative clauses are present (see Lasnik and Stowell 1991, Safir 2017, for some discussion). If we go beyond the pilot study presented here, WCO in English relative clauses also needs to be tested in a systematic manner.

[13] I represent the base position of the moved constituent as a simple gap with the relevant index. Both examples in (11) are grammatical without coindexation.

(12) *Kogo$_i$ ego$_i$ mam-a ljubit ___$_i$ [Russian]
who.ACC his mother-NOM loves
('Whom does his mother love?')

Outside *wh*-questions, however, WCO effects in Slavic are not as visible as in English; while generally present, they get suppressed under A-scrambling when the object linearly precedes the subject (Bailyn 2003, 2012; Slioussar 2007; Witkoś 2007; a.o.). Based on these additional considerations, only certain word orders inside a relative clause are informative with respect to WCO. In particular, if we need to test for WCO effects in Russian restrictive relative clauses with the object gap (ORs), the subject of that relative clause has to precede the verb, as in (13a):

(13) a. OR, WCO context [Russian]
ženščin-a$_i$ [kotor-uju$_i$ jejo$_i$ sobak-a ukusila ___$_i$
woman-NOM which-ACC her dog-NOM bit
za ruku]
at hand
'the woman whose dog bit her on the hand (lit.: which her dog bit on her hand)'

b. OR, no WCO context
ženščin-a$_i$ [kotor-uju$_i$ zlaja sobak-a ukusila ___$_i$
woman-NOM which-ACC angry dog-NOM bit
za ruku]
at hand
'the woman whom the/a fierce dog bit on the hand'

In the mini-experiment, I asked baseline and heritage speakers to evaluate 10 pairs of such relative clauses, using the 1–7 Likert scale (1: impossible, 7: fine). The baseline speakers (n = 15) all left Russia as adults and settled directly in the United States (their average age is 48; average time outside the homeland, 12 years). The heritage speakers (n = 14), average age 24;1, were all born in the United States. Typically, heritage speakers receive their main input in the home language from their parents or other older caretakers, so the choice of older baseline speakers allows us to better approximate the language of home exposure.

The distribution of responses according to conditions and participants' groups is presented in Figure 3.1.

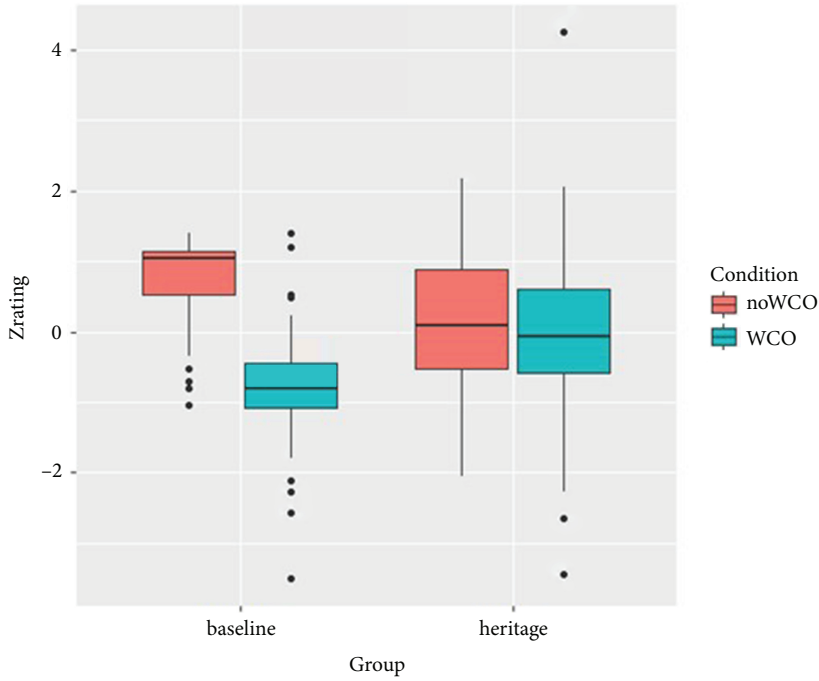


Figure 3.1 Ratings (z-score) by condition and participant group.

We observe the effect of both Group and Condition. The baseline group has a clear contrast between the WCO and non-WCO condition; the Bayesian model suggests a clear negative effect of the WCO condition (β^ = −1.41, 95% CRI = [−2.00, −0.76]), such that participants rated sentences in which the object crossed over a possessor less natural than those in which the object crossed over an adjectival modifier. The WCO effect was reduced in heritage speakers, as suggested by the positive interaction between Group and Condition (β^ = 2.19, 95% CRI = [−1.99, −3.28]).[14] Thus, WCO violations were judged natural by heritage speakers as compared to the baseline.

Finally, the baseline speakers are not uniform; they present a striking bimodal distribution shown in Figure 3.2. Some speakers are clearly sensitive to WCO violations in ORs; for nine speakers, the raw rating of such relative clauses was 2.2, as compared to 5.8 on non-WCO condition. In the other group (six baseline speakers), WCO violations received the average raw score of 5.75, as compared to 6.1 on the non-WCO condition. Crucially, the ratings from

[14] I am grateful to Zuzanna Fuchs and Utku Turk for help with the statistical analysis.

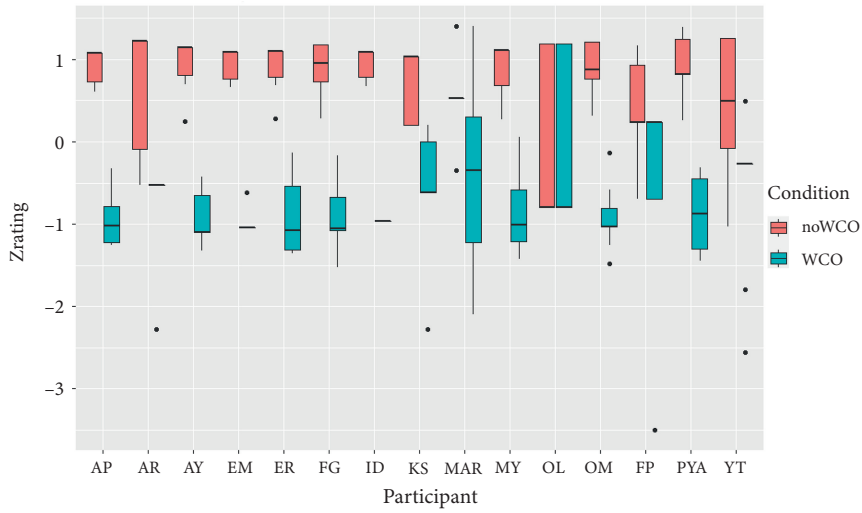


Figure 3.2 Ratings (z-scores), baseline speakers.

each individual baseline speaker are consistent; they are either about the same as the ratings of the non-WCO stimuli or they are quite low. In other words, there are no speakers who alternate their ratings for individual stimuli with WCO between high and low.

Let me first address the implications of these results for our view of *relativization in the baseline* and then for the understanding of *relativization in heritage Russian*.

It is standardly assumed that Russian restrictive relative clauses are built via A-bar movement, as diagrammed in (9) above (see Bailyn 2012: 106–108). However, the new empirical data suggest that this account is only partially correct. The bimodal pattern points to two grammars of relativization in baseline Russian, one with A-bar movement, (9), where WCO violations are found, and the other with coindexation (unselective binding), as in (10) above. In the latter case, there is no difference in rating between examples where the subject of the restrictive relative clause is modified by a possessive pronoun and those where it is modified by an adjective. Hence, WCO effects are absent.

Two other considerations lend support to the coindexation analysis. First, base-generated topics in Russian can be coindexed with null pronominals in the object position; thus, null pronominals in the object position are independently available. Second, data from the Russian National Corpus also provide additional evidence in support of the structure in (10); out of 100 restrictive object RCs with the feminine head noun, I found nine with WCO violations; for example,

(14) PVO Iraka upravljalas' francuzskoj sistemoj$_i$,
ADA Iraq.GEN was.controlled [French system].INS
kotoruju$_i$ ee$_i$ "xozjaeva" smogli distancionno
which.ACC her.GEN masters.NOM managed remotely
otključit'.
turn.off.INF
'The Iraqi air defense system was controlled by a French system that was successfully turned off remotely by its "owners."'

Thus, relativization in baseline Russian allows for at least two grammatical analyses, each observed in a subset of baseline speakers.

The co-occurrence of multiple grammars of relativization in one and the same language is not unheard of. For example, Barker et al. (1990) document two different strategies for the formation of Turkish relative clauses, distributed across speaker groups in a consistent manner. In an experimental study of English restrictive relatives, we also observed two different grammatical strategies, one with head raising, the other with operator movement (see Scontras et al. 2014; Tsai et al. 2014). Researchers have also documented variability compatible with multiple grammars in other aspects of language structure; for example, Han et al. (2007, 2016) uncover bimodal distribution among Korean speakers with respect to verb raising (as evidenced by scope readings under negation). Just as with the case presented here, speakers do not maintain such multiple grammars simultaneously; each speaker has a definite grammar and does not change it, but in the entire population of speakers, more than one option is present.

Unlike the baseline speakers, the heritage speakers were consistent in rating the WCO and non-WCO condition equally high (see Figure 3.1), which suggests that they all entertain the structure sketched out in (10): coindexation between the operator and the gap in the relative clause, rather than movement. The shift from syntactic to anaphoric dependency in heritage Russian is even more striking given that WCO effects are observed in the dominant language; the expected abstract transfer from English does not take place, at least not in this construction. It is a common observation that heritage speakers are a very varied group, ranging from highly proficient to rather weak speakers. Yet, contrary to all the expectations of variance among heritage speakers, we find that their relative clauses are built in a more uniform way than those in the baseline.

As striking as this result may seem, several cautionary notes are in order. The noteworthy difference between a subset of the baseline speakers and all the

heritage speakers may be due to a confluence of factors, non-discrimination of case morphology the first among them. If the heritage responses reflect the strong SR preference presented in Section 3.3.1, then these results may just mean that the speakers interpreted all the stimuli as having subject gaps, leaving us with no WCO to worry about. This explanation is unlikely since the stimuli did not denote reversible actions and had a clear plausibility bias in favor of the correct parse (e.g., "the woman that her dog bit"). Further still, such an explanation cannot be extended to baseline speakers, who are presumably sensitive to case morphology. One then needs two different accounts for the converging result.

Next, given the bimodal distribution in the baseline speakers, one can imagine that all the heritage speakers tested here were exposed exclusively to those baseline speakers whose relative clauses were built via coindexation, not movement. This is unlikely, but to rule out this possibility definitively, one has to test each heritage speaker together with the speakers who provided the primary input.

To build on this mini-experiment, one would need data on island effects in the relative clause.[15] It would also be helpful to compare WCO effects among first-generation immigrant speakers and age-matched speakers in the homeland; however, WCO violations found in the Russian National Corpus (see above) suggest that this pattern is found in the homeland.

To summarize, the patterns observed with relative clauses are comparable to those observed with coreferential null pronouns and point to similar trends, with the strong preference to interpret the gap as the subject, and the particular ascent of this tendency in adult heritage speakers. We also see indications of

[15] The expectation is that speakers who did not notice WCO violations would also accept scrambling (topicalization) out of the relative clause, as in (i), or would at least rate (i) higher than the group that followed the syntactic dependency. In (i), the scrambled expression and its base position are shown in bold, and the relative pronoun and the gap of the relativized argument, in italics.

(i) **Takuju** **zadaču$_i$** ja ne znaju ni odnogo čeloveka$_k$ [Russian]
[such problem].ACC 1SG not know not [one person].GEN

[*kotoryj$_k$* *__$_k$* by rešil *__$_i$* bez podgotovki]
which.NOM SBJV solved without preparation

'Such a math problem, I don't know a single person that would have solved it without special preparation.' (lit.: … that would have solved *ec* without preparation)

I have tested such examples with two speakers from the WCO group, who rejected (i); two speakers who allowed WCO violations accepted it. However, not all Russian speakers accept extraction out of finite clauses. Complicating matters further, other diagnostics of A-bar movement are not reliable in Russian. Finally, examples such as (i) are quite complicated, which may make them degraded for independent reasons, and particularly difficult to test with heritage speakers.

a syntactic dependency (movement) being replaced with the anaphoric one. This finding is based on just one data point I presented here, but if confirmed, that type of change would be consistent with other observations on HLs where the range of syntactic dependencies is more restricted than in the baseline.

3.3.2 Production

3.3.2.1 Gaps vs. resumption

In Section 3.3.1.1, I discussed the subject gap processing advantage observed in the comprehension of restrictive relative clauses. This advantage is well-attested in monolingual populations and is even more pronounced in HL speakers. What about the production of relative clauses? As I already discussed in Section 3.2.2, most studies indicate that children acquire relative clauses relatively early, around age 4;0–5;0 (see papers in Kidd 2011; also Belletti and Guasti 2015: Chapter 5, Diessel and Tomasello 2005, Lakshmanan 2000, among others). Nevertheless, they continue to make occasional errors up to ages 7;0 or 8;0 (Friedmann and Novogrodsky 2004). Further still, several researchers note the asymmetry between the comprehension and production of relative clauses; for example, Friedmann and Novogrodsky (2004) observe that relative-clause comprehension lags behind the production of relative clauses. The general bias in child language is toward relative clauses with intransitive predicates (as in *the dog* [*that*__ *ran away*]). While this is technically a sign of subject advantage, there are no competing referents in such clauses, and the relativization is arguably easier than in clauses with a transitive predicate where one has to choose between the subject and the object.

Given these observations, we would expect the production of relative clauses in HLs to proceed in a relatively unencumbered manner. The data are still quite scarce, and more empirical work is needed, going in two directions. First, it would be helpful to establish quantitative data on the types of relative clauses in heritage speakers' production, both across languages and within a single language, especially if such a language forms a bilingual dyad with different dominant languages (such as Spanish in the United States, in Germany, or in Brazil). Second, structured production data, which would allow comparisons across subjects and across languages, are highly desirable, covering both adult and child bilingual populations. The methodology for structured production can involve a preference task, a picture-description task (cf. Novogrodsky and Friedmann 2006), or elicited production (the latter is accepted but has many critics).

The picture-description task by Novogrodsky and Friedmann (2006) was designed for special-language impairment populations, but such and similar tasks can be easily modified for typically developing children and adults. For instance, in the preference task the experimenter presents two options and asks the participants to choose which one they would like better. The task is constructed in such a way that the choice would have to be expressed as a relative clause, as shown in (15):

(15) There are two children. The father combs one child, the barber combs another child. Which child would you rather be? Start with "I would rather be …" or "The child that …" (Novogrodsky and Friedmann 2006: 367)

In a pioneering study, Avram et al. (2022) applied this elicitation task in their study of French-dominant Romanian speakers ranging in age from 4;0 to 11;07, whom they then compared to an age-matched monolingual cohort. Both groups showed a preference for subject relative clauses, thus matching the results of comprehension studies. In addition, the older bilingual children produced a significant number of French-like subject relative clauses, which could be interpreted as a sign of structural transfer (as well as an indication that they are trying to avoid more complex relative-clause structures). On a more general level, the heritage speakers in this study were more likely to avoid using a relative clause than the monolingual group; instead, they used a free-standing root clause or a fragment.

The main finding that is relevant for the purposes of the current discussion had to do with the use of resumptive material in place of the gap in the relative clause, for example,

(16) […] copilu' [care încălzeşte duşu' **pe copil]** [Romanian]
child.the that warms shower.the ACC child
'the child for whom the shower warms up' (Avram et al. 2022: ex. (12))

Such use of a resumptive phrase is unacceptable in baseline/homeland Romanian, and it was not attested in the monolingual group in this study. While still tentative, these production results point to the same tendency to avoid silent elements as noted in comprehension studies. The data presented by Avram et al. (2022) suggest that monolinguals are comfortable with silent gaps. Likewise, Diessel and Tomasello (2005), who examine the acquisition

of relative clauses primarily in English, comment on the virtual absence of resumption in their monolingual data. Consistent with the comprehension data, the Silent Problem is thus not limited to referential pronouns.

The production of resumptive expressions is probably motivated by several factors, some of which match the factors influencing comprehension (see Section 3.2). First off, establishing a long-distance dependency between the antecedent (the head noun, in this case) and the corresponding element in the relative clause requires computational effort, and it is arguably easier to connect two overt expressions than an overt and a silent element, for the simple reason that the overt expressions provide more specific instructions on accessing their denotations than the silent ones. Next comes the uncertainty of interpretation; such uncertainty is characteristic of different kinds of silent elements. It is typically observed in comprehension but it may be also implicated in production; thus, heritage speakers, whose own metalinguistic experience is more limited than that of more fluent speakers, use fully specified lexical material to avoid uncertainty and to make sure they can bring their message across to their interlocutor. And finally, the difficulty of establishing a long-distance dependency increases with structural and linear distance. Avram et al. (2022) note that resumption is more common in object relatives, confirming the effect of structural distance.[16]

3.3.2.2 Truncation?

So far, I have considered only those head-external relative clauses that have full sentential structure, from CP all the way down. However, head-external relative clauses can also be smaller than a CP: so-called truncated relatives. In addition to CP-relatives, the head noun can in principle be modified by a TP- or by a vP-level nominalization. Thus, the truncated option is provided by the grammar and supported by some cross-linguistic data. For instance, nominalized TPs as relative clauses have been proposed for Semitic languages (Ouhalla 2004), where they may have developed from paratactic constructions (Givón 2009). Although we do not have data on truncated relative clauses in HLs, *child language* facts and data from *endangered/shifting languages* point to the possible relevance of truncation in heritage grammars.

Truncated relative clauses have been documented in child language production. In particular, Pérez-Leroux (1993) found that Spanish children often

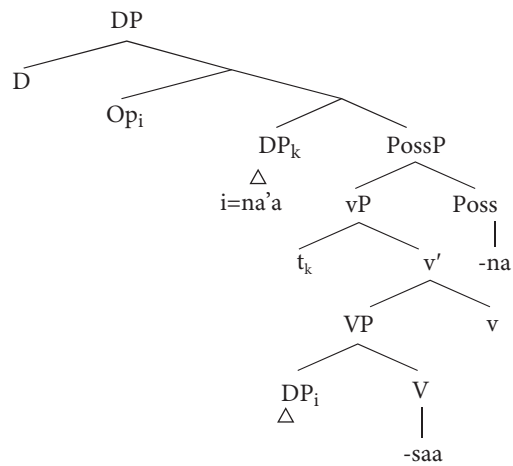
[16] Frequency of different types of relative clauses is also likely to play a role, and this is where corpus data would be helpful.

produced truncated relatives after the main clause, paratactically connected to the main clause. Parallels between properties observed in child language and HLs are often brought up in research, and it is often the case that HL language development follows the L1 developmental route. On these assumptions, the presence of truncation in child language suggests that it should be explored in HLs as well.

Truncated relative clauses are also found in a number of endangered or shifting languages. For instance, a particularly clear and well-researched case of truncated relative clauses is reported for northern Paiute, a severely endangered Uto-Aztecan language (Toosarvandani 2014). Toosarvandani provides a careful syntactic and semantic analysis of these clauses showing that they are nominalized vPs (they can include vP-level adverbials but not TP- or CP-level ones). He also shows that the syntax of these nominalized clauses does not involve A-bar movement; rather, an operator in the D-field binds a (silent) resumptive pronoun (DP_2 below), which contributes the variable (the diagnostics he uses to distinguish deletion under movement from a null resumptive pronoun are the same as discussed in Section 3.3.1.2 above). The structure of one of his examples (presented in (17a)) is shown in (17b). The subject of the truncated clause still undergoes short movement to Spec,Poss, but there is no movement associated with relativization.

(17) a. Nɨɨ ka=i=naa'a saa-na tika. [Northern
1SG.NOM DEF.ACC=1SG.GEN=father cook-NMLZ eat Paiute]
'I ate the thing that my father cooked.' (Toosarvandani 2014: 809, ex. (60a))

b.



(Toosarvandani 2014: 810)

What does this reduced structure have to do with representations of relative clauses in HLs? Just like heritage speakers, speakers of many endangered or shifting languages had limited exposure to the home language as compared to the monolingual baseline; for both types of speakers, the dominant language becomes their primary language, associated with greater societal value than the home language. In both bilingual situations, we observe the transition from early and naturalistic immersion in the ancestral language to the use of the societally dominant language. The result is different degrees of unbalanced bilingualism. Reduced exposure to the home language also leads to reduced knowledge of registers or lack of literacy. Socio-economic and socio-historical circumstances of language endangerment and HL development are different, but structural consequences of reduced exposure are quite similar in many respects (see Polinsky 2018: Chapter 7, Sasse 1992, for more discussion). Further still, for quite a few (not all) endangered languages, the only remaining speakers are heritage speakers, who are different from so-called "traditional speakers" (Bloomfield 1927; Kehayov 2017; Polinsky 2018: Chapter 1; a.o.).

In the case of HLs, the target language is spoken by homeland speakers or first-generation immigrants, and those varieties are the baseline. In the case of endangered languages, there is often no traditional language left. That's why in examining structural consequences of unbalanced bilingualism, we find a richer empirical base in HLs, and findings drawn from HLs can be extended to endangered languages. However, in case of truncated relatives, things can go in the opposite direction: while we do not have direct evidence of relative-clause truncation in HLs, the possibility of such truncation should be considered in light of data from a different set of bilinguals.

I would like to underscore here that this scenario involves the use of truncated relative clauses as the only option for relativization; if a language under contact has both full and truncated relatives, that creates additional learning problems and may interfere with the economy of resources needed under reduced exposure. In other words, we would expect an HL to have either a fully tensed relative clause or a truncated relative clause, not both.[17] Building on the parallel with endangered languages that I introduced above, it is worth noting that although truncated relative clauses are not unique to endangered languages, it is often the case that truncated structures are the only possible relative-clause type in an endangered language.

[17] Elena Soare (p.c.) informs me that while baseline Romanian has both types, no truncated relative clauses are observed in heritage Romanian production.

For example, in widely spoken languages truncated relatives co-occur with other types of relatives, e.g. in Akan or in Quechua, whereas in endangered languages they seem to be the only possible type of relative clause – that appears to be the case in Uto-Aztecan, Chimariko (Jany 2008), and Wappo (Thompson et al. 2006).

If these considerations are on the right track, we can expect to find relative-clause truncation in HLs. Its use may be yet another strategy for creating a smaller size structure (and possibly one without A-bar movement), yet without sacrificing the internal complexity of the nominalized clause.

All told, the predictions we can make regarding the division of labor between full and truncated relative clauses reflect the pattern of the division of labor between anaphoric and syntactic dependencies in a restrictive relative clause (see Section 3.3.1.2), where several grammatical mechanisms are available in the baseline, but the respective HL narrows down the choices to fewer options.

3.4 Concluding remarks

This chapter presented and analyzed several cases where overt and null expressions can appear and alternate in heritage language. In earlier work (e.g., Laleko and Polinsky 2017; Polinsky 2018), we have referred to the avoidance of silent categories and some difficulties related to their interpretation as the Silent Problem. It is worth noting that the Silent Problem extends over different types of silent categories; in this chapter, I have considered null referential pronouns in the subject/object positions and gaps in relative clauses, which in turn could be traces of A-bar movement or (null) pronominals (depending on the language and on the particulars of a syntactic analysis). Several threads get woven together in the discussion of the Silent Problem, and in this conclusion, I will try to summarize each of them.

3.4.1 Lack of data

As we start to address new, more sophisticated questions related to HLs, gaps in our empirical database also become more apparent. In the discussion throughout this chapter, I have identified several areas where we lack certain basic empirical facts needed to go forward. The first has to do with the comprehension and production of null objects, especially in the contexts where such objects are impossible (see Section 3.2.1 on Mandarin Chinese). Null objects need to be evaluated both in languages where they are cross-referenced

on the verb (e.g., Romance, Georgian) as well as in so-called topic-prominent languages (e.g., Mandarin, Korean).

Next, we do not have enough information on the actual underlying representation of restrictive relative clauses in HLs; it has been a tacit assumption in the field that the size of relative clauses in a particular HL should be identical to their size in the baseline. However, the widespread use of truncated relative clauses in child language and in endangered languages, which bear a structural resemblance to HLs, points to the need to re-evaluate this tacit assumption. In a similar vein, I have shown that the underlying representation of relative clauses may differ between the baseline and the HL, which again points to the need to produce more work in this area. Relative clauses have been reasonably well studied, and syntacticians have a wealth of diagnostics at their disposal to tap into their structure in different populations.

3.4.2 Where's the change?

An obvious answer to this question is that silent categories should not be considered holistically. Yes, we see preference for overt material, but that alone does not necessarily reflect deep changes in the underlying representations. One could claim that the abundance of overt forms has to do with easing the conditions on interpretation and avoidance of ambiguity. The language of first-generation immigrants (which constitutes the main language of exposure for heritage speakers) appears to show some overextension of overt material, albeit on a smaller scale than the corresponding HLs.[18] This points to possible extra-linguistic motivation for overt forms: speakers lack the linguistic confidence typical of a more homogenous speech community, and that compels them to avoid ambiguity or uncertainty as much as possible. There is no arguing against such speech behaviors, but they are a trigger, not the mechanism that underlies the changes we observe. I contend that the core issue of HL grammars has to do not with the licensing of silent elements, but with the dependencies in which they appear.

In at least two cases the grammar of silent categories undergoes restructuring in HLs. First, the data discussed here suggest that the subject advantage in long-distance dependencies is stronger in HLs than in the baseline. This in turn may be related to the more general property of HLs, which tend to

[18] The observations on overextension of null elements in immigrant varieties are based on the use of overt subjects in contrast to null pronominals. No such observations are available for relative clauses.

favor equipollent oppositions, where each member of the opposition is fully specified, over privative ones (where one member of the opposition represents the elsewhere condition). That means that in the overt/null expression dyad, each comes with full featural specification, which then determines its use in the syntactic structure.

Another case discussed here where HLs manifest change in underlying representations has to do with the grammar and reanalysis of relative clauses. To begin with, the baseline group surveyed in the mini-experiment that I reported here was heterogeneous with respect to the syntax of relative clauses; some speakers build their Russian relative clauses via A-bar movement (as is claimed in most analyses, e.g. Bailyn 2012), while others show evidence of coindexation. This heterogeneity points to multiple grammars of relativization within a single language. At the same time, all the heritage speakers in this study had a more uniform grammar of relative clauses, one that avoids movement. This finding points to the leveling of the grammatical patterns in an HL. A larger question that looms behind these results has to do with the direction of grammatical reanalysis: do we expect reanalysis to go in both directions, from movement to coindexation, and from coindexation to movement, or are the pathways to reanalysis more constrained? And on a more general plane, what motivates the preference for Merge over Move that we observe in heritage grammars? Here, one could imagine two families of answers. On the one hand, as some researchers have proposed, the motivation may lie entirely in performance; HL speakers are trying to minimize working-memory load, and somehow merging items is easier in that regard than displacing them. However, whether one moves or merges, a long-distance dependency is the end result, and one then needs to explain why anaphoric dependencies are somehow easier for processing. An alternative explanation may rely on the lack of featural specification that drives movement. HL grammars demonstrate the opacity of grammatical features or their attrition (e.g., Fuchs et al. 2024; Scontras et al. 2018), and this explanatory route deserves to be explored further.

3.4.3 Beyond HLs

I have noted several parallels between HLs and patterns observed in endangered languages. The reason to pursue these parallels has to do with structural consequences of reduced exposure. If endangered-language grammars show similarities to those of the more familiar HLs, then we have evidence that

HLs share universal properties that reflect restricted input in language acquisition, irrespective of the circumstances that shape them. In a similar vein, we can expect structures that have different underlying representations in heritage grammar to correspond to structures that appear late in acquisition (e.g., Montrul 2016). Assuming that different underlying representations in heritage grammar map to structures that appear late in acquisition, we can determine what properties are responsible for the late acquisition and whether they come from the limited input or from Universal Grammar.

The connection between HLs, endangered languages, and child language can be seen as a two-way street; observations from one empirical domain can be tested on the others. These general considerations also suggest that HLs are interesting not solely due to their unusual kind of bilingualism, but because they help us identify vulnerable aspects of grammar and eventually contribute to theories of learning in monolinguals.

3.4.4 Loose ends

The full list of unresolved questions would probably be longer than this entire chapter, but I wanted to zero in on one question in particular. The patterns discussed throughout this chapter are uniform in that heritage speakers tend to interpret overt forms more effectively than the corresponding silent ones. Some of these patterns also appear in production. However, heritage production is much more varied than comprehension and, crucially, is the showcase for the omission of material, especially in spontaneous speech. Whether or not this is due to real-time pressures still remains to be understood. In other words, the interplay between the overmarking and omission of material in HLs remains a challenge for future studies. Accordingly, the variation in the production by heritage speakers (which in turn has to do with the great variability of proficiency in such populations) is rich material for constructing models of comprehension and production. I started this chapter by presenting two such models, dual and single. The data considered here may simply not be enough to fully support the single-model approach to comprehension and production, but they could clearly make an impact on our decisions about the models.

Acknowledgments

This work was supported in part by NSF grant BCS-1941733 and was partially completed during my visit to the Lab “Structures formelles du

langage" at CNRS/Paris-8. I am grateful to Guglielmo Cinque, Cristina Flores, Zuzanna Fuchs, Matthew Hewett, Tanja Kupisch, Oksana Laleko, Terje Lohndal, Ora Matushansky, Eric Potsdam, Esther Rinke, Greg Scontras, Elena Soare, and two anonymous reviewers for helpful comments. All errors are my responsibility.

4

Word order and prosody in the expression of information structure

Oksana Laleko

4.1 Introduction

Word order variation in Slavic languages has been recognized as "the primary overlapping issue of interest for both syntactic theory and Slavic syntax for the foreseeable future" (Bailyn 2000, quoted in Berger et al. 2009: 654; see also Stepanov 2021 for a more recent discussion in a similar vein). Much ahead of its time, Henry Weil's 1844 astute comparative analysis of French, German, and English to Latin and Greek first drew attention to systematic differences between fixed and free word order languages, observing how word order may not only serve a role in the encoding of grammatical functions, but also be employed as a marker of discourse relations. In Weil's (1844) own conceptualization, word order in some languages is aligned more closely to the natural flow of thoughts, where that which serves as the main contribution of the sentence comes after that which is perceived as carrying less importance. The study of sentence structure from the perspective of information units contained therein – i.e., its information structure – has since formalized these early insights in continuously evolving ways, yielding various terminological systems all converging on the partition of an utterance into a relatively less informative part, labeled in various traditions as the psychological subject, theme, or topic, and relatively a more informative part, termed as the psychological predicate, rheme, or focus (Gundel and Fretheim 2004; Lambrecht 1994).

Notwithstanding the remarkable advances (particularly of the 20th-century functionalist linguistics) achieved in our understanding of how these information structural primitives are mapped onto linguistic structures in typologically distinct linguistic systems, the relative freedom in the placement of

Oksana Laleko, *Word order and prosody in the expression of information structure*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro, et al., Oxford University Press. © Oksana Laleko (2025).
DOI: 10.1093/9780191987731.003.0005

sentence constituents in languages designated as displaying flexible word order (i.e., scrambling), as illustrated for Russian in (1a–f), remains a continuous challenge for modern syntactic theorizing. Within formal generative studies, the pivotal question has been the extent to which discourse-oriented factors can be integrated into models of grammar not originally designed to account for syntactic variation arising from grammar-external, contextually imposed requirements. Indeed, as frequently noted in the sizable literature on Russian word order, every logically possible combination of the major clausal constituents is licensed to occur, and in fact does occur, under particular discourse conditions, yielding multiple construals of the same situation rendered in mutually equivalent truth-conditional terms, but differing markedly with respect to the speaker's communicative goals. To wit, the tendency of given (i.e., previously mentioned or otherwise already known) entities to precede newly introduced entities, which in turn serve as the communicative focus of the utterance, accounts for all of the word order permutations in (1), where the sentence-final constituent, which is also the constituent that bears sentence-level stress in Russian, is most easily interpreted as the main informational contributor of the entire syntactic unit:

(1) a. Babushka poteryala ochki. SVO
b. Babushka ochki poteryala. SOV
c. Ochki babushka poteryala. OSV
d. Ochki poteryala babushka. OVS
e. Poteryala ochki babushka. VOS
f. Poteryala babushka ochki. VSO
'Grandma lost (her) glasses'

To complicate matters further, the movement operations that bring the surface orders into alignment with their pragmatically determined functions are on an optional nature: as illustrated in (2) below, the speaker of a flexible language may not need to utilize scrambling but may instead opt for the syntactically unmarked, "basic" order (assigned to SVO in Russian for reasons discussed further on), relegating information structural partition solely to prosodic prominence[1] (and, by extension, deaccenting). Thus, the prosodically stressed element in each of the following sentences receives a focused interpretation in situ, i.e. without changing its base-generated syntactic position. The fact

[1] Where relevant, prosodic prominence is marked by small capitals.

that (2a–c) indeed differ in their information structural content can be easily demonstrated with the standard "question-answer" diagnostic of new information focus, whereby the constituent that contributes the crucial element of the response to a corresponding *wh*-question is said to be presentationally focused (e.g., Van der Wal 2016):

(2) a.	Babushka poteryala OCHKI.	'What happened?' or 'What did grandma lose?'
b.	Babushka POTERYALA ochki.	'Why isn't grandma wearing her glasses?'
c.	BABUSHKA poteryala ochki.	'Who lost their glasses?'
	'Grandma lost (her) glasses'	

Among the plethora of unresolved issues pertaining to the apparent optionality in the word order domain of Slavic, the question of disentangling the effects of syntax and prosody in the encoding of information structural distinctions has been at the forefront of much recent theoretical and experimental work. Experimental studies with Russian monolinguals, for example, have shown that linear constituent order and the acoustic prosodic expression work in tandem to convey distinctions associated with the partition of sentences into units of information (Luchkina and Cole 2021; Sekerina 2003). However, as we will see in greater detail in the section that follows, the division of labor between these mechanisms is subject to extensive theoretical debates. While traditional, Prague School inspired proposals continue to advocate for the primacy of word order in the expression of information structure in Slavic, alternative accounts assign a more significant role to prosody, with word order serving only as a secondary mechanism employed to modulate the placement of sentence stress. As research into these issues continues, investigations of information structure marking in colloquial and spoken registers of discourse-prominent word order languages will be increasingly important, as it is in these varieties that we observe the fullest extent of grammatically permitted variation in both domains.

Within the burgeoning literature on heritage languages (D'Alessandro et al. 2021; Montrul and Polinsky 2021), word order change has, somewhat unsurprisingly, also emerged as one of the central research avenues in recent years. Such rapidly growing interest in the expression of sentence- and discourse-level relations in heritage languages, following nearly two decades of work targeting their inflectional properties, stands to reflect a new phase in the overall development of the field, which has undergone a significant shift from

a framework sometimes characterized as a "deficit-based" or "reductionist" approach to the study of heritage grammars to a more targeted analysis of areas of structure-preserving change, complexity trade-offs, and redistribution of functions in these developing varieties (see, e.g., D'Alessandro and Frasson 2023; Laleko and Scontras 2021; Polinsky and Putnam 2024). The need for a more nuanced understanding of complexification processes in heritage language change has led researchers away from the narrow focus on paradigmatic reduction toward an exploration of rules and principles underlying syntagmatic relations and structures, where the effects of structural loss are less discernable. Inquiries into the encoding of information structural relations by heritage bilinguals are particularly promising in this regard. Since these fundamental communicative distinctions are encoded in some way in the grammars of all human languages, we are likely to observe their manifestations at various stages of heritage language change; and since the particular means for their expression tend to be especially rich and varied at the level of spoken registers, the predominantly oral linguistic experience of early naturalistic bilinguals provides optimal conditions for the study of change without loss (Laleko 2021).

Along with the ongoing epistemological shift in modern heritage language research, the last few years in particular have seen an increasingly expanding volume of work conducted with advanced, communicatively competent heritage bilinguals, capable of producing large units of connected discourse (and not only in oral but also written modality) and able to carry out sophisticated experimental tasks aimed to elicit nuanced grammatical distinctions. A quick glance at the existing studies on heritage Russian in contact with English serves as a handy illustration of this tendency. Up until the end of the first decade of this century, most published work drew on the linguistic knowledge of limited-proficiency speakers, including nearly dormant bilinguals with restricted opportunities for language use, bringing into focus areas of overt representational restructuring in adult grammars affected by language shift. Studies with fluent and actively practicing heritage bilinguals, including fieldwork research in integrated Russian-speaking communities (e.g., Laleko and Miroshnychenko 2022) and investigations of grammar-building processes in child heritage speakers, only took off during the last decade, drawing researchers' attention to areas of stability and processes of non-reductive change in heritage language transmission.

The chapter aims to contribute to both lines of research by probing the interactions between word order and prosody in the formal encoding of information structure in homeland and heritage varieties of Russian, with an

additional distinction made based on the proficiency of heritage speakers. Departing from much prior work on grammatical competence of heritage language speakers, and building on the growing recognition of heritage bilinguals as native acquirers of their home language regardless of the attained – or maintained – proficiency in it, the approach undertaken here does not aim to identify problematic areas in heritage speakers' knowledge or frame differences between bilinguals and predominantly monolingual homeland speakers in terms of deficits to be overcome with the larger goal of bringing the underlying grammatical systems into a closer alignment. Instead, assuming that the natural outcomes of language change are bound to vary in its varying ecological environments, the chapter offers a conceptualization of heritage language word order change in terms of change in the degree of prosodic and word order plasticity (Vallduví and Engdahl 1995) in ways that broadly parallel microvariation observed across Slavic languages (cf. Šimík and Wierzba 2017 on Czech, Slovak, and Polish). An important advantage of this approach is that it is more explicit than most traditional accounts of heritage language change in approaching all grammars under analysis as internally consistent, even if differently configured, systems. In this regard, the chapter advocates for and advances the position that heritage languages are a valid and essential empirical resource for linguistic theory building – not only on the grounds of what gets eliminated or streamlined in these grammars, revealing the bare bones of linguistic structure in ways that X-ray images represent the human body, but also, more broadly, as sources of rich data on the evolutionary mechanisms of variation, differentiation, and change that drive the formation of linguistic diversity over time.

4.2 The expression of information structure in Russian

Prior to reviewing the competing positions that have been put forward in the Slavic linguistic literature regarding the roles of word order and prosody in the expression of information structure in syntactically flexible systems, we must first discuss the place of these systems within the existing word order typologies. Since Greenberg's (1963) original classification of languages into types based on the most neutral ordering of the subject, direct object, and verb, it has become commonplace to find Russian included in the SVO class (Bailyn 2002; Dryer 2007). However, it has also been widely acknowledged that Slavic languages come out as highly exceptional when compared to other Indo-European languages within this class (cf. English, north Germanic, or

Romance): once the full range of their syntactic properties is considered, SVO seems merely "the least inappropriate" label (Siewierska and Uhlířová 2010). In fact, some recent analyses have employed the Slavic data to propose a new word order type, T3, not yet sufficiently integrated into existing grammar theories, where the verb–object relationship is parametrically unspecified and thus ambidirectional (Haider and Szucsich 2022). While compelling on many grounds, these accounts are certain to require extensive empirical testing, as they must ultimately be reconciled with the experimental data that have consistently pointed to SVO serving as the unmarked, discourse-neutral order in processing and acquisition studies on Russian (Bailyn 1995; Kallestinova 2007; Sekerina 1997, 2003), with corpus studies that find SVO to predominate in spoken and especially written production (Miller and Weinert 1998), and with most existing theoretical models of scrambling, which explicitly build on an underlying SVO tree in deriving all remaining orders (Dyakonova 2009; Junghanns and Zubatow 1997; Kondrashova 1996; Slioussar 2007; but see King 1995 for the position that SVO is derived from the more basic VSO).

4.2.1 Syntactic expression of information structure

Since the earliest formalizations of information structure encoding through word order (Firbas 1964; Mathesius 1947), the Slavic syntactic tradition has been heavily dominated by what may be termed "the primacy of word order" approach to information structure. In this account, word order variation in discourse-configurational languages like Russian is taken to follow directly from the need to satisfy discourse requirements by bringing the surface order of constituents into alignment with the information structural template shown in (3) (Brun 2001; Dyakonova 2009; Gundel 1988; King 1995):

(3) (Topic) – (discourse-neutral information) – focus

Setting aside variations in prosody, which we will attend to in the next section, this underlying structuring has been shown to account for the subtle nuances in the interpretation of sentence sets like those in (1a–f) and hence posited as the motivation for the syntactic processes necessary to derive the non-basic orders in (1b–f) from (1a). Uncovering the nature and mechanics of these derivations has been a focal point of nearly three decades of research on Russian, yielding a vast array of proposals striving to deliver a

unified account of scrambling along its many dimensions: what moves and where, at which point in the syntactic derivation, and by what mechanism. Unfortunately, as attested by the easily observable "disorder" in the available word order literature, such a unified account has not been attained to date: even the most fundamental issues seem to resist consensus, such as whether topical and focused elements are housed in dedicated syntactic projections (Dyakonova 2009; Junghanns and Zubatow 1997; King 1995) or are marked solely with features (Kondrashova 1996), whether all movements are of the same type (and of which type, A or A') and directionality (Bailyn 2002; Sekerina 1997; Slioussar 2007), and what additional interpretational effects, if any, each of the movement operations may bring about (Bentzen 2015; Titov 2012).

In a very general sense, the proposed transformations may be summarized as falling into two basic types: object fronting, manifested most typically in Russian as SOV and OSV orders, and subject-postposing, yielding (O)VS structures; other, less frequent, orders could then be derived by some combinations of these processes. Assuming the overarching principle summarized in (3) above, the unifying factor for the first set of structures in most accounts involves the need to remove the object from its base-generated postverbal domain, which happens to be the focal domain; depending on its discourse status (i.e., topic or discourse-neutral material), the object moves to the clause-initial position (OSV) or to the middle-field (SOV), yielding structures illustrated in (1c, b) above. Conversely, when the subject functions as the exponent of focus, it appears at the right edge of the clause, often triggering the displacement of the object to the left-edge position in accordance with the principle of progression from given to new information.

Various analyses have been put forward to account for the above processes. For example, in Kučerová's (2012) analysis of Czech, Russian, and Serbo-Croatian, word order variation reflects a grammatical system of givenness marking that relies on a specialized covert G-operator that identifies expressions as given, much in the same way as the definite article marks definite DPs in English. However, other researchers have argued against integrating information structural notions into the grammar (cf. Chomsky 2008), either as features or as heads that attract XPs with the relevant values (Rizzi 1997): since these notions are construed in context, the relevant interpretations could conventionally piggyback on syntactic variability driven by independent principles (Fanselow and Lenertová 2011; Neeleman and Szendrői 2004).

4.2.2 Prosodic expression of information structure

Unlike the word order-based accounts discussed earlier, the prosodic accounts of information structure encoding in flexible order languages take word order variation to be a consequence of a combination of two prosodic requirements: first, that sentences are expected to bear stress on their final constituents; second, that given expressions cannot be stressed. The "given-before-new" principle is hence derived from the interaction of these constraints without the need to posit additional rules in the grammar pertaining to information structure-related constituent movement (Zubizarreta 1998; for applications to Slavic see Szwedek 2011 for Polish and Šimik and Wierzba 2015 for Czech).

The role of accentuation as a trigger of syntactic movement has been at the center of accounts arguing against a direct link between the notions of information structure and syntactic derivation. In these models, the primary means of information structure realization is prosodic prominence: focused elements are associated with the main sentential stress, which they may receive in situ via the shift of stress from its default position or, alternatively, by a constituent order change that aligns focus and stress in a way that preserves the prosodic shape of the utterance. As with the purely syntactic accounts discussed earlier, the literature is rich in the proposals about the nature of the associated transformations, operationalized either in terms of focus movement (Reinhart 1995) or as the evacuation of non-focused material from positions associated with focus (Neeleman and Reinhart 1998).

Within the Russian linguistic tradition, researchers have long recognized the contribution of prosody in information structural partition, drawing a distinction between neutral sentences, in which the primary (i.e., nuclear) pitch accent falls on the sentence-final constituent in accordance with the Nuclear Stress Rule (Chomsky and Halle 1968), and sentences with a non-neutral (i.e., emotive, emphatic, expressive) intonation (Bryzgunova 1981; Kovtunova 1976; Krylova and Khavronina 1988; Svetozarova 1998; Yokoyama 1986). The established linguistic classification makes use of the notion "intonational contour" (IK) to distinguish sentences of the first type (IK-1) from sentences of the latter type (IK-2); the following summary in Table 4.1, from Jasinskaja (2016), outlines the particular prosodic features associated with each pattern.

The position of the nuclear accent has further been linked with the distinction between broad and narrow focus. Under broad focus (also known as sentence focus), the domain of focus is said to project from the accented sentence-final constituent to the entire sentence, marking the entire utterance

Table 4.1 Summary of prosodic features of "neutral" (IK-1) and "non-neutral" (IK-2) intonation in Russian

	"neutral"	**"non-neutral"**
alignment of the nuclear fall	early (HL*)	late (H*L)
position of the nuclear accent	on the last word	on any word
prenuclear pitch accents	on all other words	few if any
downstep	yes	no
nuclear fall excursion	low	high
intensity of the accented syllable	low	high

Source: Jasinskaja (2016)

as new information in discourse (and hence a suitable answer to the question What happened?). Under narrow focus, the domain of focus is restricted to the constituent bearing the accent, such that the utterance is only felicitous as an answer to a question about that constituent, as demonstrated in (2b–c) above, or in corrective contexts. Notice that for the SVO sentences such as the one in (2a), both broad and narrow focus possibilities are available; however, as predicted by the summary of features in (4), their prosodic contours under the two readings are not identical (see also Alter et al. 2001).

In scrambled sentences, too, the effects of prosody distinguish structures with different pragmatic properties. For example, the object-preposing constructions SOV and OSV, discussed previously only as instances of relocating given objects in order to satisfy (3), are not limited to these contexts; in fact, both scrambled orders can also be produced by focus preposing, noted to be particularly well represented in colloquial Russian (Krylova and Khavronina 1986; Yokoyama 1986). Example (4), from Dyakonova (2009: 64), illustrates various landing sites of the preposed focus, including the middle field 4(a,b) and the left periphery of the clause (4c):

(4) (Context: What did his parents give Nikita for his birthday?)

a.	Oni	emu	ŠĆENKA	podarili.
	they-NOM	him-DAT	puppy-ACC	gave-PL
b.	Oni	ŠĆENKA	emu	podarili.
	they-NOM	puppy-ACC	him-DAT	gave-PL
c.	ŠĆENKA	oni	emu	podarili
	puppy-ACC	they-NOM	him-DAT	gave-PL

'They gave him a puppy.'

Middle-field focus and left-peripheral focus have been shown to differ, both from each other and from sentence-final focus, in conveying different degrees of discourse accessibility; despite this interpretive difference, and contrary to prior accounts in which only sentence-final focus could project beyond the element bearing stress, preposed focus constructions too have been suggested to have projection properties (Dyakonova 2009: 69–82; Jasinskaja 2016), making them compatible with a much wider set of contexts than previously thought. If on the right track, these observations paint a picture of a highly complex system of interactions between prosody and word order in Russian and call for more experimental studies, which have been slower to come than the quickly expanding theoretical work.

4.3 Prior work on word order in heritage Russian

Early explorations of heritage Russian in the United States, framed as investigations of American Russian, contain only occasional references to its word order properties; however, even from these mostly descriptive accounts one is able to identify a rather consistent pattern of overextension of the basic SVO order to contexts where a non-canonical order would have been more felicitous. For example, the "American Russian" and "Full Russian" versions of the same sentence in (5a–b) below, presented in Polinsky (2006: 212–213) as an illustration of English interference in the domain of lexical constructions (namely, the English calque "start the car"), demonstrate differences in the placement of the subject, with the SV order in the heritage Russian sentence (5a) corresponding to the VS order in its baseline equivalent (5b):

(5) a. Sedognja moja mašina ona ne načinalas'.
today my car she$_{RP}$ NEG began

b. U menja segodnja ne zavodilas' mašina.
by me today NEG wound car
'My car wouldn't start today.'

Building on earlier characterizations of American Russian as displaying "frozen SV/SVO order" (Polinsky 1995: 386), Polinsky (2006: 237) notes that "American Russian speakers clearly limit the range of word orders available to them," but finds that VS orders are not completely absent across the board. Rather, they appear to be restricted to existential and presentational constructions and surface most frequently under negation or when preceded by an overt adverbial (Polinsky 2006: 237).

Further evidence for the growing frequency of SVO constructions in heritage Russian comes from a series of elicited production studies with Russian-English bilinguals. A comparative investigation of narratives from college-age heritage learners of Russian and monolingual speakers by Isurin and Ivanova-Sullivan (2008) revealed that VS orders were used three times as often by baseline Russian speakers compared to heritage learners (6% versus 2.1%), and that their use among the six heritage language speakers surveyed appeared to be linked to the participant's length of exposure to English (Isurin and Ivanonva-Sullivan 2008: 99). These overall tendencies were confirmed in a more large-scale study by Ivanova-Sullivan (2014), who once again documented under-production of VS orders in heritage speakers' oral narratives compared to those of baseline controls (5–8% versus 16%) and noted their correlation with heritage language proficiency (pp. 105–110). However, since none of the above studies analyzed factors governing the occurrence of SV and VS structures relative to their context of use or examined other non-canonical orders, the implications of these findings concerning an at-large typological change toward SVO taking place in heritage Russian under the influence of English, a position maintained at least tentatively by the authors of the reviewed studies, remain subject to further substantiation.

In fact, some experimental investigations targeting the acceptability of inverted structures under subject focus in heritage Russian have challenged claims about word order linearization in this domain. A bimodal acceptability ratings study by Ionin and colleagues uncovered the same preferences for SVO orders in answers to object questions and for OVS orders in answers to subject questions both in heritage and baseline speakers, in stark contrast to a group of relatively less proficient L2 Russian learners who were also tested in the study and showed no awareness for the effects of information structure on Russian word order (Ionin et al. 2023). While, like in earlier studies, the acceptance rates for inverted structures in the heritage language group were found to correlate with the speakers' level of proficiency, the authors argued against proficiency as the key source of differences between early and late bilinguals, appealing instead to the heritage speakers' advantage of childhood exposure (Montrul 2008) and a general stability of narrow focus marking in heritage languages (Laleko 2021).

In subsequent work, Ionin and Luchkina (2024) reported additional data from Hebrew-dominant heritage Russian bilinguals, comparing them to the results presented in Ionin et al. (2023). Once again, while proficiency affected individual participants' ratings, the heritage speakers were target-like as a

group, prompting the authors to conclude that the relationship between word order and information structure (namely, the association between SVO and object focus and OVS and subject focus) is generally well established in the heritage language.

Taken together with the results of the production studies reviewed earlier, these findings raise some interesting questions about the status of inversion in heritage Russian. While heritage language proficiency appears to impact both the production of inverted structures and their comprehension in context, signs of change are not equally perceptible in these two domains: based on their grammaticality judgments, the heritage and homeland speakers appear to share similar representations of these structures, but in production heritage bilinguals consistently favor basic over inverted orders. On the surface, such patterns seem in line with recent theoretical proposals that draw a fundamental distinction between mechanisms employed for generation and parsing of linguistic structures (see Polinsky, this volume, for an overview), raising the possibility that heritage language change may in principle affect these mechanisms differentially. However, a closer look at the acquisition studies of word order distinctions suggests that comprehension in this domain often lags behind production (e.g., Cannizzaro and Hendriks 2020), a pattern that would predict, contrary to the empirical facts above, a more consistent use of VS orders in heritage language narratives, where speakers are free to apply their internalized knowledge of how to generate and deploy these structures, compared to their evaluation of these structures in strictly controlled, pre-determined linguistic contexts.

Another possibility is that the existing studies diverge on the participants' level of proficiency, and in trying to align their results in a cross-sectional manner we are ultimately drawing generalizations not about production and comprehension, but about distinct, chronologically removed stages in the process of heritage language change. In this account, the grammars of speakers at the lower end of the proficiency continuum would represent outcomes of deeper restructuring characteristic of advanced language shift, while grammars of the more proficient speakers would reflect its incipient manifestations (cf. also Polinsky and Kagan's 2007 application of the creole continuum model to the concept of heritage language proficiency). In the absence of an independently validated and consistently applied proficiency measure, it is not possible to rule out the scenario where the grammar of inversion looks different at different points in the heritage language diachrony.

Yet another dimension to the problem of divergent findings in the literature on inversion in heritage Russian, and possibly in heritage languages more generally (cf. the discussion of the Spanish data in Laleko 2021), concerns the properties of the specific structures examined in the reviewed studies. In the literature on production, the obtained VS counts are invariably generalized from intransitive and transitive contexts and reported jointly, whereas the acceptability studies pointing to target-like representations of focus in heritage speakers are typically restricted to OVS structures.

In a written acceptability judgments study of subject inversion, Laleko (2022) tested both intransitive (VS) and transitive (OVS) structures in two proficiency-based groups of heritage Russian speakers in the United States., compared to homeland speakers and L2 learners. The results were not uniform for the two structures and at the two levels of heritage language proficiency examined in the study. In intransitive contexts, where VS orders in Russian may be triggered by predicate type (unaccusative) or subject focus, the high-proficiency heritage speakers were sensitive to the effects of verb semantics but not to the effects of subject focus; in transitive contexts, however, these speakers displayed the same grammaticality contrasts for subject focus as homeland speakers. The low-proficiency speakers patterned with L2 learners in all conditions, preferring SV(O) orders regardless of subject focus or predicate type. Apart from lending further support for the key role of proficiency as a determinant of what grammatical outcomes one may observe, and hence generalize from, in modeling a heritage language, these results underscore the need for a nuanced approach to the study of broad and multifaceted phenomena in shifting grammars. To wit, the differential effects of focus on the acceptability of subject-final structures observed in intransitive and transitive contexts point to changes in the functions of these structures in the heritage language. With intransitive verbs, where extra-linguistic cues (focus) compete with intralinguistic cues (verbal semantics) in the conditioning of VS orders, heritage speakers seem to place a greater emphasis on the lower level, more immediately accessible cues. In contrast, the mapping between the topic-focus structure of the sentence and its syntactic structure holds more consistently in OVS structures, and it is in these less ambiguous contexts that these higher level contextual cues are employed as word order triggers in the heritage language (Laleko 2022: 13–14).

A serious limitation of the work discussed so far is its categorical bias toward the study of subject inversion. While acknowledging the rich flexibility of Russian word order, all of the above studies draw conclusions about

the directionality of heritage language word order change exclusively on the basis of change (or lack thereof) in one particular domain of its elaborate system. Only one production study to date has examined a general distribution of canonical and non-canonical orders in heritage Russian, with an additional emphasis on the contextual felicity of the attested orders and on rules governing the occurrence of structures involving dislocation, i.e. the movement of non-subject topical material to preverbal positions (Laleko and Dubinina 2018). Despite some quantitative reduction in the use of non-canonical orders by heritage speakers compared to homeland speakers (22% versus 32%), the bilinguals and monolinguals exhibited similar patterns in the distribution of particular word order types; furthermore, the two groups converged in their use of object fronting constructions, whose occurrence in both data samples was strongly predicted by syntactic weight (i.e., "light" objects such as pronouns or single nouns were more likely to undergo fronting than "heavy" phrasal objects, which were more likely to remain in situ). However, heritage and homeland speakers diverged in their placement of verbal modifiers: here, weight continued to dominate in the heritage group but not in the homeland group, where the effects of weight were less categorical (Laleko and Dubinina 2018: 210). Thus, once again, we find evidence of heritage language change manifested in the use of particular word order principles, rather than as a wholesale linearization and leveling of word order distinctions in favor of SVO, with some non-canonical constructions holding their ground in the word order grammars of heritage language speakers despite pressures from the syntactically rigid dominant language.

The complex dynamics of word order change in heritage Russian are further brought to light in a comparative investigation of different types of non-canonical structures, including inversion (OVS) and object fronting (SOV, OSV), in two groups of heritage speakers distinguished by proficiency (Laleko 2024). Drawing on a series of contextualized acceptability judgment experiments, the study identified distinct trajectories in the development of these structures in the heritage language, with interesting effects of speakers' proficiency on the obtained results. While under-rating inverted orders, heritage speakers converged with homeland speakers in their judgments of object fronting, and speakers at the higher level of proficiency in fact showed a trend toward overgeneralization of SOV orders, compared to the lower proficiency speakers and homeland speakers. These results offer strong evidence against the view that heritage language word order change proceeds linearly toward SVO syntax, suggesting instead that some non-canonical orders may

remain resilient or even gain momentum in robustly maintained heritage varieties, mirroring and amplifying quantitative tendencies of their use in the colloquial registers of the homeland variety. However, since the study tested object fronting structures with given objects, an important question left for future work is to examine SOV and OSV constructions derived from focus fronting, as is done here.

4.3.1 The big picture

Taking stock of the findings outlined in this section, while there is strong evidence to believe that heritage Russian varieties tend to exhibit signs of word order change, the scope, directionality, and underlying causes of this change are far from being clear, and several considerations could inform further research on the subject. Broadly, as evident from studies in which heritage language speakers are found to converge with homeland speakers on certain non-canonical word order properties, it is apparent that word order change is not inevitable for all constructions and across all domains of the richly configured Russian word order system, an observation that prompts a more in-depth investigation of contexts and factors that are implicated in the use of particular orders by heritage speakers. Data on the overextension of SOV orders in some heritage Russian varieties are particularly encouraging in this regard, as they attest to a multidimensional nature of heritage language word order change that adds nuance to the more typical story of transfer-induced shedding of complex forms and structures.

At the same time, an overwhelming body of findings do attest to some kind of reductive change taking place in shifting grammars; here, one might recall the extensively researched phenomenon of subject inversion, which all production studies without exception have found to decrease in frequency for heritage Russian, with a proficiency-modulated trend toward the weakening of its link to focus structure observed in experimental work. In light of growing evidence that reductive change in contact situations is often accompanied in the contracting language by related processes in associated domains, one possible avenue of inquiry involves a close examination of other factors involved in the encoding of subject focus; for Russian, this would undoubtedly be prosody. Only one study to date has probed the use of prosodic stress in heritage Russian. Ionin and Luchkina (2024) reported on their experiment in which they tested the link between prosodic prominence and contrastive stress:

sentences bearing felicitous and infelicitous stress were presented as answers to clarification questions, and instances of match and mismatch between what the question is asking and which constituent gets stressed in the answer were analyzed. The results suggested, perhaps to not much surprise considering the pivotal role of prosody in the marking of contrastive focus in English, that English-dominant heritage Russian speakers have no difficulties with the use of contrastive prosody in corrective focus contexts. However, since contrastive focus is semantically and prosodically distinct from presentational (i.e., new information) focus, it is important to examine how heritage Russian speakers use prosody for focus marking outside of corrective contexts, and how their use of prosody measures against their reliance on word order in the expression of new and given information. The present study is the first step toward addressing these questions.

4.4 The study

4.4.1 Research objectives and study design

To address the question about the division of labor between word order and prosody in the encoding of information structural relations in Russian, and to examine ways in which heritage language change happens (or not) in one or both of these domains, scaled acceptability ratings were elicited from homeland and heritage speakers for sentences in three experimental tasks. The first task examined the use of inversion, comparing SVO and OVS structures in prosodically felicitous and infelicitous structures; the remaining two tasks examined object fronting, comparing the distribution of SVO, SOV, and OSV orders in contexts where the object conveys new information focus (task 2) and in contexts where it represents given, contextually salient information (task 3). The target sentences for all structures were embedded into short dialogues, consisting of a question-answer sequence, presented bimodally such that the participants both saw and heard the full dialogue in order to rate the acceptability of the answer relative to the question, using the same five-point scale from "impossible" to "very natural" for all 44 target items, intermixed with fillers. All focused constituents in the target sentences provided information that served as an answer to the question posed in the first part of the dialogue; all given constituents had an overt linguistic antecedent in the question, were definite, and were not focused contrastively or otherwise.

4.4.2 Participant demographics

Thirty-five speakers of Russian took part in the study, including 14 homeland Russian speakers and 21 heritage Russian speakers who grew up in the United States and reported English to be their main language of daily communication. The demographic summary presented in Table 4.2 captures the familiar signature patterns of heritage language acquisition in the bilingual group, where all participants reported early exposure to both Russian (from birth) and English (from birth or, for most participants, in later childhood), with Russian serving the primary role until about the age of schooling and subsequent socialization into the dominant culture. In addition to the sociolinguistic questionnaire by which these demographic data were collected, all speakers including baseline controls have completed a proficiency test consisting of 10 sentences constructed to probe the participants' knowledge of grammatical constraints independently known to be subject to variability in bilingual acquisition and thus serving as expedient indicators of proficiency: case, grammatical gender, verbal aspect, negative concord, and semantically congruent use of indefinite modifiers. To keep the proficiency task comparable to the main experimental task, the same five-point acceptability scale was used across the board, yielding the best possible score of 1 (if the participants rated all target sentences as completely unacceptable) and the lowest possible score of 5 (if the speakers judged them to be perfectly acceptable in Russian). Based on the obtained proficiency scores, the group of heritage bilinguals was divided into two subgroups: speakers with the average result between 1 and 3 were designated as the higher proficiency speakers (HL-h), and speakers whose average score fell within the range between 3 and 5 were designated as the lower proficiency speakers (HL-l). The relevant demographic data, calculated separately for each group, are presented in Table 4.2.

4.4.3 Results and discussion

The results were analyzed statistically using Welch's unequal variances t-tests. First, in order to obtain a general understanding of how prosody and word order are engaged in focus marking in homeland and heritage Russian, comparisons were performed between conditions demonstrating the felicitous application of prosodic (i.e., via stress shift in situ) and syntactic (i.e., via subject-final order) strategies for the encoding of subject focus, labeled as the SVO and OVS conditions in Figure 4.1.

Table 4.2 Participant demographics

	HL(h) ***(N = 11)***	**HL(l)** ***(N = 10)***	**Homeland** ***(N = 14)***
Age	25.5 (18–39)	21.4 (18–26)	40.8 (24–69)
Age of learning English	6.7 (4–12)	3.4 (0–9)	N/A
Age of English becoming primary	9.5 (6–13)	5.6 (0–13)	N/A
Average daily use of Russian (%)	31.4 (5–90)	13.5 (0–50)	100%
Understanding spoken Russian	9.4 (6–10)	8.2 (7–10)	9.9 (9–10)
Speaking Russian	7.5 (4–10)	5.8 (3–10)	9.9 (9–10)
Reading in Russian	6.4 (1–10)	5.3 (1–10)	10.0
Writing in Russian	6.2 (1–10)	4 (1–10)	9.6 (8–10)
Proficiency score	2.7 (1.9–2.8)	4.2 (3.6–4.9)	1.5 (1–2.5)

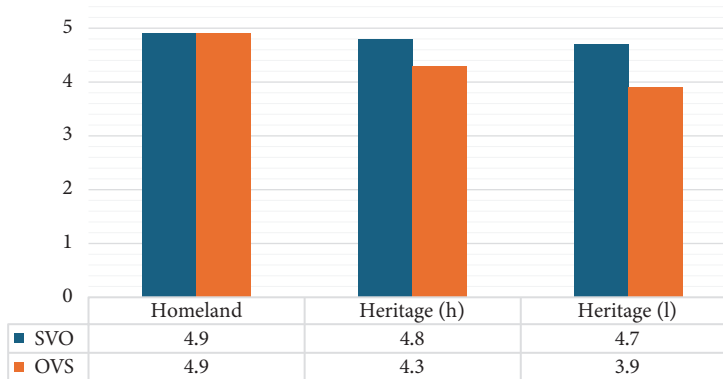


Figure 4.1 Prosodic vs. syntactic encoding of subject focus.

Homeland speakers in the control group provided virtually at-ceiling ratings of sentences in both conditions, showing no statistically significant preference for moving the focused subject to the right-edge position over allowing it to remain in situ: $t(103.3) = 0.2$, $p > 0.01$. These results demonstrate that the marking of subject focus in homeland Russian may be accomplished equally efficiently with or without syntactic movement, with prosody serving as a sufficiently rich resource to identify the focused subject without word order change.

Turning now to the heritage Russian data, the in situ marking of focus through prosody alone emerged as the preferred mechanism for the expression of this information structural category in both bilingual groups, as evidenced by a highly significant acceptability difference observed between the in situ and subject-final conditions: (HL(h): $t(58.5) = 2.5$, $p = 0.01$; HL(l): $t(55.4) = 3$, $p < 0.01$). Across-group comparisons further confirm that the homeland and heritage speakers' judgments of prosodic expression of focus in Russian are in full convergence regardless of the heritage speakers' level of proficiency (homeland vs. HL(h): $t(72.2) = 0.92$, $p > 0.01$, homeland vs. HL(l): $t(54.8) = 1.8$, $p < 0.01$), while the judgments of its syntactic realization diverge significantly, with a statistically robust acceptability dip in both heritage groups compared to the homeland group (homeland vs. HL(h): $t(52.8) = 2.9$, $p < 0.01$, homeland vs. HL(l): $t(44.9) = 4.0$, $p < 0.01$).

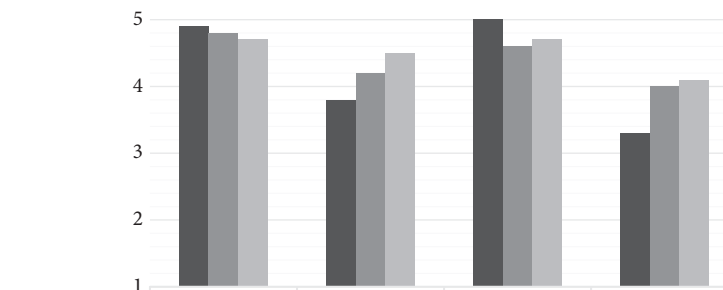
Thus, we see that all varieties of Russian under examination employ prosody in the expression of subject focus, and that movement becomes a less productive focusing strategy in heritage grammars compared to baseline grammars. To put it differently, while the word order principle "place given elements before new elements" bears less weight in heritage Russian compared to homeland Russian, it is notable that violations of this constraint (i.e., the in situ focus structures) do not incur any acceptability penalty in any populations of speakers, thus placing the prosodic strategy of focus marking ("place prosodic stress on the focused constituent") above the word order constraint in all grammars. That this ranking holds without exception across all varieties of Russian tested in the study could be a strong empirical argument in favor of prosodic approaches to information structure marking in Slavic; the question to ask next concerns the degree to which prosodic violations incur an acceptability penalty in the respective varieties. Under the assumption that the stronger constraint will induce a greater penalty when deviated from, comparing the "costs" of such deviations can offer crucial evidence about the ranking of the respective principles in the corresponding grammars.

To assess violations in the prosodic encoding of focus, the word order factor was removed altogether from consideration. Thus, ratings for SVO sentences were examined in four conditions: with a stressed focused subject (felicitous), with an unstressed focused subject (infelicitous), with a stressed focused object (felicitous), and with an unstressed focused object (infelicitous). Figure 4.2 presents the mean ratings in each condition.

Homeland speakers demonstrated the expected prosodic felicity contrasts in both subject and object conditions, as evidenced by a significant difference

between sentences with stressed and unstressed in situ subjects under subject focus ($t(63.6) = 6$, $p < 0.01$) and between sentences with stressed and unstressed in situ objects under object focus ($t(55) = 8.1$, $p < 0.01$). Additionally, homeland speakers were sensitive to context when evaluating sentences of the same prosodic and syntactic structure under distinct discourse conditions: object-stressed SVO structures received significantly higher ratings under object focus than under subject focus, where a stress shift to the subject would have been expected ($t(55) = 7.1$, $p < 0.01$). Similarly, subject-stressed SVO structures received significantly higher ratings under subject focus than under object focus, where a stress shift was contextually unwarranted: $t(60.3) = 7.3$, $p < 0.01$.

Data from the bilingual participants point to a certain degree of neutralization in these prosodic patterns, with an interesting distinction emerging between the subject-focus and object-focus conditions. No significant contrast was obtained in either bilingual group between subject-stressed and object-stressed (i.e., default) SVO sentences under subject focus (HL(h): $t(55.8) = 2.6$, $p > 0.01$; HL(l): $t(76.8) = 0.8$, $p > 0.01$); however, the analogous distinction under object focus did reach significance in both bilingual groups: (HL(h): $t(81.1) = 2.5$, $p = 0.01$, HL(l): $t(60.6) = 2.9$, $p < 0.01$), mirroring the felicity contrast observed for this condition in the data from homeland speakers. In other words, the acceptability penalty for an overextension of the marked prosodic pattern (i.e., unwarranted application of stress shift) remained high in heritage bilinguals, while infelicity stemming from an overuse of a prosodic default (i.e., failure to apply stress shift) became less pronounced. If on the right track, these results suggest that heritage language speakers exhibit a greater reliance than



	Subj Foc	#Subj Foc	Obj Foc	#Obj Foc
Homeland	4.9	3.8	5	3.3
Heritage (h)	4.8	4.2	4.6	4
Heritage (l)	4.7	4.5	4.7	4.1

Figure 4.2 Prosody of subject and object focus in SVO.

homeland speakers on the prosodically and syntactically default pattern – SVO with the right-edge stress – in their acceptability judgments of constructions with in situ focus.

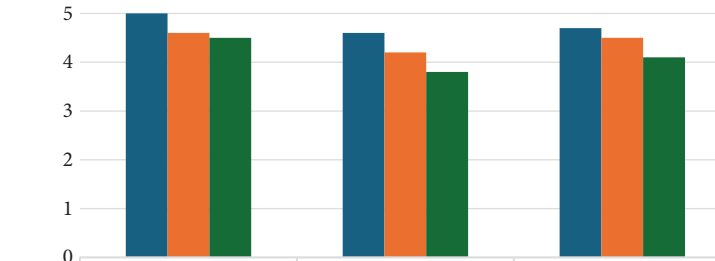
Additional evidence for this generalization comes from comparisons of the heritage bilinguals' judgments of these doubly default structures (i.e., SVO with sentence-final stress) based on the context of occurrence: under object focus, where they are felicitous, and under subject focus, where infelicity is expected to arise. Regardless of proficiency, heritage speakers rated these conditions uniformly: for HL(h), $t(80.3) = 1.4$, $p > 0.01$, for HL(l), $t(72.4) = 1.2$, $p > 0.01$. However, a proficiency-related difference emerged in the bilinguals' ratings of the prosodically marked, subject-stressed SVO orders. Speakers in the higher proficiency group rated these structures differently under subject focus than under object focus ($t(56.3) = 3.9$, $p < 0.01$), as did homeland speakers, while in lower proficiency speakers the same comparison yielded no significant difference ($t(66.6) = 2.6$, $p > 0.01$), suggesting more variability in evaluating stress shift relative to discourse context.

To summarize, both heritage and homeland Russian speakers utilize prosody to mark focus, and in all varieties examined the association of focus with prosody is a stronger requirement for focus marking than the requirement for the given and new material to be linearly arranged in a certain way. However, the findings also reveal signs of change in the heritage language affecting both dimensions of information structure expression: in the realm of syntax, it is manifested as the extension of the word order default; in the realm of prosody, the default right-edge stress pattern is similarly overgeneralized. Viewed together, these two developments may be reflective of a more general mechanism of change, whereby default structures are preferred to and promoted over structures that are more marked – that is to say, more limited in their contextual distribution, more highly specified, and less frequently represented in the input. In itself, this pattern of change is well known from language contact studies; what makes these results particularly interesting here is their co-occurrence across the syntactic and prosodic domains. As discussed earlier, prior accounts have tended to attribute the strengthening of the SVO default in the heritage language to the influence of the dominant language, which in most examined cases involved a more rigidly enforced SVO order. The present findings offer a novel perspective on the role of the dominant language as a mediator of change. In the realm of prosody, the observed over-extension of

the default accentuation pattern is unlikely to be transfer-induced: the dominant language does not create especially favorable conditions for it, with stress shift serving as a robust prosodic marker of subject focus in English.

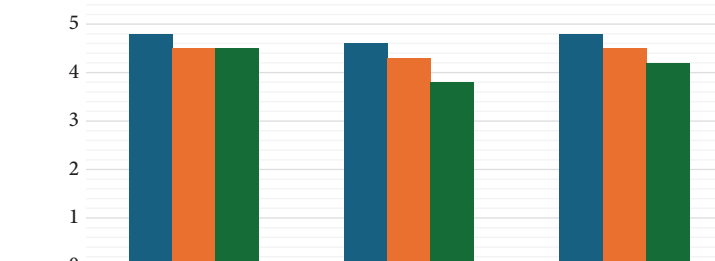
Next, we address the relationship between information structure and word order in the domain of object fronting, looking at the distribution of SVO, SOV, and OSV orders with focused objects (Figure 4.3) and with given objects (Figure 4.4). First, we will examine the acceptability of each structure under distinct discourse conditions (i.e., in object-new and object-given contexts), aiming to obtain the empirical judgments of the two types of Russian object preposing: one associated with focus, and one associated with givenness. Under the approach that attributes word order variation in Slavic to the need to satisfy the "given before new" principle, the fronting of focused objects should bear a greater acceptability penalty than the fronting of given objects: in the former case, both word order and prosodic stress diverge from the default pattern, while in the latter case, the intonation remains neutral and the order of constituents is properly aligned with the expected information structural template. Absence of such penalties would, correspondingly, lead us to consider the notion of communicative dynamism as a much weaker determiner of constituent order in flexible word order systems than most syntactically oriented accounts acknowledge. Second, looking separately at the object-new and object-given conditions, we will assess the acceptability of object movement for each of the two possible scrambled orders relative to the basic order. Our goal here is to twofold: to obtain the empirical data from homeland speakers for object scrambling in Russian and to trace possible change in its information structural functions in the heritage language.

Comparing the object-new and object-given conditions for SVO, SOV, and OSV sentences, no statistically significant differences were obtained for either one of the three order types in either one of the three participant groups (the full statistical data are not presented here for space considerations; however, the nearly identical mean ratings for each order in Figures 4.3 and 4.4, respectively, offer a helpful illustration of the obtained pattern of ratings). These results attest that object scrambling is indeed licensed in both information structural configurations in Russian and is generally stable in the heritage language. We thus see that analyses in which the occurrence of OSV and SOV structures in Russian is discussed narrowly as a manifestation of the "given before new" principle are limited in their capacity to identify the wider range of pragmatic functions associated with these orders.



	Homeland	Heritage (H)	Heritage (L)
SVO	5	4.6	4.7
SOV	4.6	4.2	4.5
OSV	4.5	3.8	4.1

Figure 4.3 Syntactic position of object focus.



	Homeland	Heritage (h)	Heritage (l)
SVO	4.8	4.6	4.8
SOV	4.5	4.3	4.5
OSV	4.5	3.8	4.2

Figure 4.4 Syntactic position of object givenness.

Having established that object scrambling is freely available to homeland and heritage speakers with both focused and given objects, our next step is to measure the acceptability of the scrambled orders relative to their pragmatic context. To achieve that, SOV and OSV structures were each compared with the default SVO pattern in two conditions: object-new and object-given. Data from homeland speakers yielded a nuanced picture of object scrambling in Russian, suggesting that the movement of focused objects is more restricted than the movement of given objects. In the object-new condition, statistically significant contrasts between the basic and scrambled orders were detected with both SOV and OVS structures, revealing an acceptability penalty for the use of focus preposing in monolinguals. In particular, with new objects, SVO sentences were rated as more acceptable than SOV sentences, $t(55) = 5.4$, $p < 0.01$, and as more acceptable than OSV sentences,

$t(55) = 5.3$, $p < 0.01$. With given objects, however, neither type of scrambled structure was associated with a decrease in acceptability, with SVO sentences receiving the same ratings as SOV sentences, $t(93.6) = 2.3$, $p > 0.01$, and the same ratings as OSV sentences, $t(106.5) = 2.5$, $p > 0.01$. In sum, in evaluating object scrambling, homeland Russian speakers are sensitive to the discourse status of the scrambled object; however, they treat middle-field scrambling and left-edge scrambling uniformly within each pragmatic context.

Data from heritage speakers showed a markedly different pattern: in both groups, it was the movement distance, rather than the object's discourse status, that determined acceptability. Comparisons between SVO and SOV sentences yielded no significant differences at either proficiency level in either pragmatic condition: with new objects, HL(h): $t(85.2) = 1.7$, $p > 0.01$, HL(l): $t(75.5) = 1.6$, $p > 0.01$; with given objects, HL(h): $t(81.8) = 1.6$, $p > 0.01$, HL(l): $t(77.9) = 2$, $p > 0.01$. In contrast, comparisons between SVO and OVS sentences produced significant differences at both proficiency levels in both pragmatic contexts, including with new objects, HL(h): $t(77) = 2.9$, $p < 0.01$, HL(l): $t(56.9) = 3.1$, $p < 0.01$, and with given objects, HL(h): $t(70.4) = 3.2$, $p < 0.01$, HL(l): $t(61.8) = 3$, $p < 0.01$. Thus, middle-field object scrambling in heritage Russian was found to apply more freely overall than in the homeland variety, while long-distance scrambling consistently incurred a higher penalty regardless of discourse context. Table 4.3 summarizes these patterns: the plus sign indicates an acceptability penalty relative to the basic, non-scrambled order and the minus sign indicates no such penalty.

In the next section, I review the key findings of the study and discuss their implications. I frame this discussion around the idea that served as an inspiration for this project: that our understanding of heritage language change can build on and benefit from models that seek to study its manifestations from perspectives adopted in studies of linguistic microvariation, comparative typology, and diachronic change.

Table 4.3 Constraints on object scrambling in homeland and heritage Russian

	Homeland Russian		**Heritage Russian**	
	Focused object	Given object	Focused object	Given object
SOV	+	–	–	–
OSV	+	–	+	+

4.5 Conclusions

Broadly, the results of the study confirmed that both prosody and word order are implicated in the expression of information structural distinctions in homeland and heritage Russian. However, the analysis of individual discourse contexts demonstrated that the division of labor between these mechanisms is not generally tilted toward word order in either variety examined, contrary to what one would expect based on accounts advocating for the primacy of word order in the encoding of discourse relations in Russian. In the expression of subject focus, homeland Russian speakers judged movement on par with prosodic marking in situ, pointing to a productive use of both strategies in felicitous contexts. However, comparing contexts targeting non-conformance to the expected prosodic pattern of information structure realization ("focused elements carry stress") and contexts targeting violations of the expected ordering ("given elements linearly precede focused elements") points to an overall much weaker role of word order compared to prosody in all varieties considered. In this light, the weakening of word order as a marker of subject focus in heritage Russian, observed here and in many prior studies, may be viewed as a natural extension of the already existing asymmetry built into the system handling information structural relations in Slavic, rather than a mere result of copying from the syntactically inflexible English.

That this general tendency may hold for Slavic at large is suggested by some recent work on west Slavic languages. In their study of Czech, Polish, and Slovak, Šimík and Wierzba (2017) found a robust correlation between information structure and prosody but a weaker and less pronounced interaction between information structure and word order. The east Slavic data presented here corroborates this overall pattern, with additional effects of microvariation in the relative strengths of the operative constraints in the heritage and homeland varieties. Very similar effects were attested in the west Slavic group: in Polish, for example, scrambling was altogether much more restricted than in the other two languages, and in Czech and Slovak the movement of given expressions proved to be much freer than the movement of focused expressions (Šimík and Wierzba 2017), in a direct parallel to the results from the homeland Russian group obtained here.

Cross-linguistic variation in the relative weights of interacting factors in the expression of information structural distinctions has been conceptualized in terms of plasticity (Vallduví and Engdahl 1995), a notion proposed to capture degrees to which languages may tolerate deviations from the default patterns in word order and intonation to mark discourse relations. The original

insight of the proposal was that a language may display plastic intonation but non-plastic word order (e.g., English) or plastic word order but non-plastic intonation (e.g., Catalan), with the key idea that typologically distinct systems may be parametrized on the basis of their primary mechanism of information structure encoding. However, comparative work on closely related varieties, including work on heritage language change, can productively extend the notion of plasticity to contexts of microvariation, where different processes across distinct subsystems could affect the degrees of plasticity across both dimensions of information structure encoding in either similar or diverging ways.

If that is the case, the observed relaxation of information structural constraints on middle-field object scrambling in heritage Russian, compared to baseline Russian, may be viewed as heritage Russian displaying higher word order plasticity in this domain, considering that the pragmatic unmarking of scrambled orders makes them available in a greater variety of contexts. As noted earlier, a tendency toward overextension of SOV structures to SVO contexts in heritage Russian was independently observed in Laleko (2024); a greater plasticity in OV/OV alternations is also documented in colloquial Russian and child Russian (see, e.g., Laleko and Kisselev 2023 for a recent overview). Conversely, the emergent restriction on long-distance movement, manifested in this study as a greater penalty for OSV and OVS structures relative to the default order in the heritage language group, may be viewed as a development in the direction of decreasing plasticity, whereby the reduced range of landing positions for scrambling could over time limit the set of grammatically allowable structures in the varieties undergoing change. The relative plasticity of prosody is also subject to variation in non-parametrically distinct systems: while both heritage and homeland Russian productively employ stress shift, the relative overextension of the default prosodic pattern observed in heritage Russian suggests a higher degree of rigidity in its use as a focus-marking device. The strength of this approach lies in its potential to capture heritage language change in its multiple manifestations, viewing it as a series of individually motivated developments interacting to produce system-wide outcomes that conspire to create diverse and dynamically developing word order systems in related but gradually diverging language varieties, much in the same way that characterizes diachronic change on a larger scale.

PART III
THE VERB PHRASE

Chapter 5
Non-active voice in heritage Greek

Artemis Alexiadou and Vasiliki Rizou

5.1 Introduction

In this chapter, we examine the production of verbs which bear non-active (NAct) Voice, as observed in narration tasks by heritage speakers (HSs). In particular, we are interested in HSs of Greek in two different majority language contexts, English and German. These speakers are the outcome of immigration primarily due to political and socio-economic reasons and they belong to the Greek diaspora. Immigration to the United States can be traced back to the 19th century, while immigration to Germany is more recent, mainly after World War II (cf. Chasiotis et al. 2006 for an overview of Greek immigration; Smith 2015). Our speakers are thus HSs of Greek, as they are bilinguals and acquire Greek in a context where English and German are the majority languages. Following recent work on the grammar of HSs, we take these speakers to belong to a continuum of native speakers, as they acquire both languages in a naturalistic way. However, the level of proficiency in the heritage language may vary across their lifespan (Benmamoun et al. 2013; Montrul and Polinsky 2011, 2019; Rothman and Treffers-Daller 2014; Scontras et al. 2015; Wiese et al. 2022). In the first years of their life, HSs are dominant in their heritage language, while upon entering in the educational system the opposite pattern may be observed (Montrul 2016; Polinsky 2018). According to various researchers, the linguistic proficiency maintenance in the heritage language depends on the quantity and the quality of input including the instruction of the language via formal education, the age of onset to bilingualism, the socio-economic status, the well-being, and the generation of immigration of their parents/guardians (Daskalaki et al. 2019; De Houwer 2015; Dosi and Papadopoulou 2019; Flores et al. 2017; Kupisch 2019; Meir and Armor-Lotem 2017; Romanowski 2021; Unsworth et al. 2014).

Artemis Alexiadou and Vasiliki Rizou, *Non-active voice in heritage Greek*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro, et al., Oxford University Press.
DOI: 10.1093/9780191987731.003.0006

As heritage languages are usually acquired via oral speech which takes place among the members of the core family and is restricted to conversational informal communication, they are often taken to lack register variation, specifically to lack features that characterize formal communicative situations (Chevalier 2004; Dressler 1991: 101–102). Biber and Conrad (2019: 2) define register as the accumulation of linguistic characteristics that are common in one variety in combination with the situation in which this variety is used. The broad distinction between formal and informal register has been put forth in Halliday (1978: 31–32), who refers to this as the tenor of discourse; Poynton (1985) and Eggins (2004: 100–101) describe each register's characteristics as follows: in formal communicative situations, monolingual speakers tend to use passive forms of verbs, participles, an abundance of adjectives, and lexical verbs instead of periphrastic constructions as well as complex structures such as subordination. It has been observed that HSs exhibit register leveling in various areas; for instance, German heritage speakers in the United States generalize certain clause types, see, e.g., Pashkova et al. (2022). Russian heritage speakers perform poorly on adjectives, as these characterize the formal register, see Polinsky (2005). Greek heritage speakers have been shown to generalize periphrastic constructions across registers in order to express aspect (Alexiadou and Rizou 2023) and to generalize the use of the imprecise and informal determiner *kati* 'some' (Alexiadou et al. 2022).

Our study investigates the production of novel patterns involving NAct Voice in two groups of HSs, asking the question of how these novel patterns can inform our theories of the morpho-syntax and semantic impact of Voice in Greek. These patterns may be very telling about overgeneralization in inflectional morphology in general, as HSs have been reported to overapply regular morphological rules (Polinsky and Scontras 2020). In view of the fact that NAct morphology in Greek may be involved in creating a variety of interpretations, we might in principle expect it to be generalized across contexts. However, as Greek NAct Voice is associated with the formal register, we do not expect HSs to make productive use of it.

The chapter is structured as follows: Section 5.2 presents an overview of the Greek Voice system, while Section 5.3 explores previous literature on L1 and heritage language acquisition regarding verbs that bear NAct Voice such as passives and deponents. In Section 5.4, the research questions based on certain hypotheses alongside our predictions are presented. Section 5.5 presents the methodology of this study, the design and the profile of our participants. Section 5.6 presents our analysis and Section 5.7 concludes.

5.2 The Greek Voice system

Greek has two sets of Voice forms: active and non-active. Voice features have been argued to be next to verbal stem/root, as seen in (1) (Philippaki-Warburton 1998):

(1) verb root- VOICE- ASPECT- TENSE- AGREEMENT

NAct Voice in Greek is linked to the [+learned] feature which characterizes words that either come from Ancient Greek or constitute artificial formations of *katharevousa* (i.e., the artificial form of Greek developed by scholars in the 19th century) (Anastasiadis-Symeonidis and Fliatouras 2003).

NAct Voice is used with passives, reciprocals, reflexives, and certain anti-causatives[1] as well as with deponent verbs, see (2) (Alexiadou and Anagnostopoulou 2004; Embick 1997, 1998). While in (2a–c), NAct is the canonical marking of [-] transitivity, deponents, as in (2d), are taken to exhibit a mismatch between morphology, syntax, and semantics, as NAct Voice appears with [+agent] and [+] transitive verbs (Alexiadou 2024; Embick 1997; Grestenberger 2018; Zombolou and Alexiadou 2014).

(2) a. i supa kege**te**. [Marked anticausative]
the soup-NOM burn-NACT.3SG
'The soup is burning.'

b. to vivlio diavas**tike** ktes. [Passive]
the book-NOM read-NACT.3SG yesterday
'The book was read yesterday.'

c. i Maria htenize**te**. [Reflexive]
the Mary-NOM combs-NACT.3SG
'Mary combs herself.'

d. O Janis ekmetalev**ete** ta kefalea tu. [Deponent]
John exploited-NACT.3SG the funds his
'John exploited his funds.'

[1] Several change of state verbs form anticausatives with active Voice (Alexiadou and Anagnostopoulou 2004):

i To vazo espase
the vase break.ACT
'The vase broke.'

One possible account of the presence of NAct on the verbs in (2a–c) is as follows. Building on Embick (1998), Alexiadou, Anagnostopoulou, and Schäfer (2015) propose that NAct morphology realizes a Voice projection which lacks a specifier. From this perspective, NAct realizes a structure of the type (3a), where [-D], following Schäfer (2008), signals the absence of a specifier. In the absence of Voice or in the presence of a DP in Spec,VoiceP, Greek verbs are assigned default active Voice, (3b–c). (3a) is basically an unaccusative structure and it is associated with passives, reflexives, and NAct-marked anticausatives. The rule that regulates the spell-out of Voice in the absence of a specifier in this system is given in (4).

(3) a. $[_{\text{MiddleVoiceP [-D]}}$ NAct $[_{\text{vP}} [_{\text{ResultP}} \sqrt{}\text{burn}]]]$ [Non-active]
b. $[_{\text{vP}} [_{\text{ResultP}} \sqrt{}\text{open}]]$ [Active]
c. $[_{\text{VoiceP}}$ DP $[_{\text{vP}} [_{\text{ResultP}} \sqrt{}\text{burn}]]]$ [Active]

(4) Voice → NAct/_ (no specifier)

In Oikonomou and Alexiadou (2022), (3a) is a special structure not only morphologically, but also semantically: in (3a), Voice forms a single interpretation domain with the vP with which it combines. Thus, Voice may receive a contextual interpretation, but also idiosyncratic and novel interpretations may arise. As we will see, (3a) is also employed in the context of transitive deponent verbs, whereby the external argument in this case is a low agent, introduced below VoiceP, see Grestenberger (2018) and Oikonomou and Alexiadou (2022) for details.

5.3 Voice in L1/2L1 acquisition

The acquisition of NAct forms is considered to be a late-acquired phenomenon in L1 Greek. The approximate age of L1 NAct acquisition beginning with short passives is around 5 years old (Fotiadou and Tsimpli 2010; Grey 2020; Zombolou et al. 2010). Previous studies on bilingual children show that these do not seem to behave radically differently from monolingual children with respect to reflexives, while they seem to be more vulnerable with passives (Grey 2020; Tsimpli 2006). Furthermore, as observed in Paspali et al. (2022) and Unsworth et al. (2011), Greek bilingual children exhibit lower accuracy with anticausatives bearing NAct Voice.

It has been pointed out that heritage speakers appear to face difficulties especially with inflection in the verbal domain. Several studies have shown

that these bilingual speakers are prone to loss of morphological features and thus they simplify the inflectional features, deviating from canonical verbal paradigms (cf. Putnam et al. 2021 for an overview). For instance, Polinsky (2000) observes an incomplete verbal paradigm in American Russian, while Fernández-Dobao and Herschensohn (2021) mention that HSs of Spanish in contact with English overregularize irregular verb stems.

Voice alternations in heritage Greek have not been the focus of extensive investigation. One of the first researchers who noticed some inconsistencies regarding the production of verbs with NAct by HSs is Seaman (1972: 165–166). He reports that Greek HSs in the United States avoid NAct forms and hesitate to use them. Despite the low frequency of verbs produced in NAct Voice identified in his research, one of the examples he includes is that of the deponent verb *diigume* 'narrate.NAct' which HSs changed to active Voice.

Zombolou and Alexiadou (2012, 2014) discuss Greek HSs in Argentina and Australia and show that these speakers generalize active on deponents and so-called marked anticausatives. This observation alongside Paspali et al.'s (2022) and Unsworth et al.'s (2011) findings can be explained on the basis of the *Syntactic Complexity Hypothesis* (SCH) (Zombolou et al. 2010), according to which, NAct forms are more complex and thus avoided by learners and HSs. This is readily accounted for under the representations in (3), whereby (3a) is more complex than (3b).

Specifically, the majority of the participants in Zombolou and Alexiadou's study (63%, $p < 0.01$) changed the morphology of transitive deponents by using active in the place of NAct, (5):

(5) a. ≠Onirevo (= onirevome) ta pragmata tu
dream.ACT.PRES.1SG the thing.PL.ACC the
panepistimiu
university.SG.GEN
'I am dreaming of the university things.' (HS Argentina, 22 years)

b. ≠Emis ekmetalevome (= ekmetalevomaste) tis mikres
we exploit.ACT.PRES.1SG the small
hores
countries.PL.ACC
'We exploit the small countries.' (HS Australia, 35 years)

Similar loss of verbal deponency is already observed within the geographical bounds of Greece.

Zombolou and Alexiadou thus concluded that since their participants changed transitive deponents only, this happens due to overregularization.

Transitive verbs in Greek canonically bear active and not NAct; NAct, as shown in the previous section, is reserved for intransitive members of an alternation. HSs thus replace NAct with active. As Greeks in Australia changed the morphology of deponent verbs less than Greeks in Argentina, 59% vs. 67% respectively, language interference from English can be excluded. Thus, the authors proposed that we have reanalysis resulting in simplification of the system and loss of morphological variation.

While this study observed the generalization of the active default on marked anticausatives and transitive deponents, other work has observed the opposite generalization: it is the marked, i.e. reflexive morphology that is generalized in, e.g. heritage Spanish (Polinsky 2018). According to Polinsky, this is correlated with low proficiency; (6) is described as odd, as it is not attested in the input, but is produced by low-proficiency heritage speakers of Spanish:

(6) El lobo se perseguió el [Heritage Spanish]
the wolf se chased the
conejo.
rabbit
'The wolf chased the rabbit.' (Polinsky 2018: 177)

5.4 Research questions

In this study, we aim to explore the following research questions (RQs). The first one (RQ1) is: which verb categories bearing NAct Voice are produced by two groups of HSs of Greek in Germany and in the United States? On the basis of our results, we will further explore whether these categories diverge from monolingual productions and whether we can attribute this to language contact (RQ2). Based on previous findings (Polinsky 2018; Seaman 1972; Zombolou and Alexiadou 2012, 2014), we expect two possible scenaria (RQ3): either to find overgeneralization of the active default or generalization of NAct correlated with low proficiency. RQ4: Given the methodology that we implemented in the present study, the addition of an extra-linguistic factor such as register variation is expected to trigger monolinguals to associate formal registers with NAct Voice verbs and differentiate their narration in informal communication settings by producing less verbs in NAct Voice, unlike HSs who are expected to exhibit register leveling according to previous studies on different phenomena (Alexiadou et al. 2022).

5.5 Methodology

Following Wiese's (2017) "language situation" setting, we conducted our research by collecting naturalistic data through narration tasks. By implementing this methodology, we are provided with a variety of phenomena allowing us to conduct exploratory research, unlike the data collected from controlled experiments. This methodology is suitable for detecting patterns or even slight changes that are happening in a linguistic system and is appropriate for eliciting both monolingual and bilingual L2 learners or HSs, regardless of the age group they belong to. In addition, the task set-up provides comparable oral and written data in different levels of formality (data sets 2×2). From this perspective, we can explore to what extent participants can differentiate their narrations and which linguistic features are present in the different communicative situations. HSs narrated the event in both their heritage and their majority language and the elicitation sessions in the different languages were scheduled from three to five days apart while monolingual participants took part only in their mother tongue. Different elicitation combinations were created to avoid biasing our participants starting each time with the same setting (see Section 5.9: Appendix). The balanced elicitation process consists of 16 elicitation orders, half starting with the heritage language and half with the majority language.

5.5.1 Design

The task consisted of a short video (00:42 minutes) simulating a fictional event. A minor car accident takes place in a parking lot and participants were asked to narrate to different people what happened supposing that they were present at the time the accident occurred. The video presented is inspired by an everyday event which happens in urban cities and, as the pilot study revealed, participants are familiar with this scene and they feel comfortable to talk about such an instance.

For the elicitation of the different communication situations, we simulated two distinct communication settings: the formal and the informal. The elicitation of the formal setting took place in an office with an elicitor formally dressed. Participants were sitting opposite the elicitor at an adequate distance. The elicitor of the formal register used standardized language with honorifics asking participants to testify in detail how the accident happened to the police. The oral narration of the formal register involved leaving a voice mail on the

answering machine of the police department, while the written one was to type a testimony on the police's laptop as they witnessed the incident.

The elicitation of the informal part took place in a different office, where the two interlocutors were sitting close to each other. A different elicitor was responsible for the informal setting who was casually dressed, and very talkative without using any honorifics. A 20-minute chitchat preceded the elicitation of the data to involve participants in a more relaxed and friendly environment so participants could simulate more easily the targeted communication setting. While they were chatting, the elicitor offered treats to the participants in order to make them feel comfortable. After this conversation, the main part of the study took place. In this communication situation, participants had to narrate in a voice message on WhatsApp what happened to a close friend. The written mode was to send a text message to the same close friend and describe the accident again on WhatsApp. The informal elicitor always reminded participants of their appointment one day before via text message in a friendly and familiar way. The whole process was recorded.

5.5.2 Participants

For our study, two groups of HSs were recruited, one in the United States (NY and Chicago) and another one in Berlin, Germany. Moreover, one control group was recruited in Athens, Greece, in order to compare the monolingual variety with the heritage ones. Table 5.1 presents the metalinguistic data that are collected in the form of a questionnaire at the end of the elicitation task. Beginning from the upper rows, the table exhibits the number of speakers in every group, their gender, and their mean chronological age by the time of testing. The following three variables have been measured only for HSs, as the control group has been raised monolingually: mean age of onset to bilingualism, and current input in the heritage and in the majority language. The age of onset ranges from 0 to 8 years old meaning that within the groups there are both simultaneous and sequential speakers, while the current input is measured bidirectionally in two scales from 1 to 3, the first one counting how often participants speak Greek to different members of their family and friends and the other one counting how often each of these people speaks Greek to them. The mean score of the self-ratings has been computed out of four questions on reading, writing, comprehending the spoken language, and speaking the language, first calculated for Greek and then for the majority language on a scale from very easy to very difficult (1 to 5) for each question. The input received

Table 5.1 Metalinguistic data across groups

		HSs in the United States	HSs in Germany	Monolingual controls
N		63 (34 females)	48 (24 females)	64 (32 females)
Chronological age		22;9 (min 14 – max 35) SD 7.335	23;2 (min 14 – max 36) SD 6.617	21;5 (min 13 – max 35) SD 6.638
Age of onset		1;41 (min 0 – max 6) SD 1.956	1;9 (min 0 – max 8) SD 2.162	–
Current input in the heritage language		0.789 (min 0.1 – max 1.5) SD 0.3482	1.1 (min 0.0 – max 1.8) SD 0.423	–
Current input in the majority language		1.16 (min 0.5 – max 2.0) SD 0.3458	0.82 (min 0.1 – max 1.6) SD 0.342	–
Self-ratings in the heritage language		3.59 (min 1.75 – max 5) SD 0.9369	4.08 (min 2.25 – max 5) SD 0.838	4.87 (min 3.25 – max 5) SD 0.3118
Self-ratings in the majority language		4.91 (min 4 – max 5) SD 0.2153	4.83 (min 3.5 – max 5) SD 0.362	–
Literacy practices in the heritage language		1.01 (min 0 – max 2) SD 0.5252	1.2 (min 0 – max 2) SD 0.551	1.54 (min 0.33 – max 2) SD 0.4614
Literacy practices in the majority language		1.81 (min 0.3 – max 2) SD 0.3601	1.52 (min 0 – max 2) SD 0.500	–
Parent's generation	Both 1st	25 prt	25 prt	–
	One* 1st	3 prt	–	
	One 1st one 2nd	18 prt	3 prt	
	One 1st one foreign	2 prt	16 prt	
	Both 2nd	8 prt	–	
	One 2nd	1 prt	–	
	One 2nd one foreign	1 prt	2 prt	

Continued

Table 5.1 *Continued*

	HSs in the United States	HSs in Germany	Monolingual controls
Years of education in the heritage language	9;13 (min 0 – max 12) SD 3.401	7;2 (min 0 – max 12) SD 3.511	–
Hours of education in the heritage language	2166.19 (min 0 – max 3120) SD 890.692	6155 (min 0 – max 12,480) SD 4605.401	–

*This category refers to single-parent families that are of first- or second-generation immigrants.

from media (TV, radio, blogs in Greek/majority language) is measured on a three-point scale (0–2) measuring the frequency from never to often and is listed under the term literacy practices. Another variable that is included in the metalinguistic data is the generation to which our participants' parents belong. Thus, we can have a clear picture of their baseline information. Additionally, the years and the hours of formal bilingual education have been calculated for the heritage group based on their reports and the equivalent curricula in each country.

Turning to between group differences we ran several non-parametric Mann–Whitney tests: we indeed find some significant differences between the communities. Comparing the two HS groups, we find that there is a significant difference regarding the literacy practices in the heritage language ($U = 953.500$, $Z = -2.673$, $p < 0.05$) and in the majority language ($U = 867.000$, $Z = -3.525$, $p < 0.001$) and regarding the current input in the heritage language ($U = 680.000$, $Z = -4.135$, $p < 0.001$) and in the majority language ($U = 649.500$, $Z = -4.348$, $p < 0.001$). Further significant differences are observed for the visits to Greece ($U = 678.000$, $Z = -5.492$, $p < 0.001$). No other significant differences were found for the remaining metalinguistic factors.

5.6 Results

In this section, we present an overview of our results. The total tokens per speaker group from the latest version of the RUEG corpus (Wiese et al. 2021) and the tokens of verbs in absolute numbers and frequencies are presented in Table 5.2. Table 5.2 exhibits also the verbs found in ACT and NAct Voice per group and their frequency based on the number of verbs. Among the matches

Table 5.2 Number of words, verbs, verbs in ACT, and NAct Voice in the Greek subcorpus per group in tokens and frequencies

	HSs in the United States	**HSs in Germany**	**Monolingual**
No. of tokens	22,577	22,942	30,777
Tokens of verbs	3671 (0.16)	3661 (0.15)	4957 (0.16)
Tokens of verbs in ACT	3044 (0.82)	2946 (0.80)	3723 (0.75)
Tokens of verbs in NAct	614 (0.16)	708 (0.19)	1219 (0.24)
Tokens of verbs in NAct excluding *ime* and *ginome*	158 (0.04)	310 (0.08)	667 (0.13)

of verbs in NAct Voice are copulas like *ime* 'be', passives *prokalume* 'caused', reflexives like *travmatizome* 'get hurt', reciprocals like *sigkruome* 'collide', and intransitive deponents like *erhome* 'come'. The third line of Table 5.2 exhibits the verbs in NAct Voice having excluded the copula *ime* 'be' and the deponent *ginome* 'become'. All queries can be found in the Appendix (Section 5.9).

A manual categorization of the different verb categories is presented in Table 5.3, which includes row instances and frequencies derived from normalization per 100 tokens based on the tokens of verbs found in the corpus. The data upon which the manual categorization has been conducted are stored in an OSF repository.[2]

What we can notice in Table 5.3 is that, as expected, monolinguals produce the most deponent verbs bearing NAct forms followed by the group of heritage speakers in Germany (although in this group less participants were recruited) and finally the group that produced the least deponent verbs is HSs in the United States.

We observed 10 non-target forms in our two heritage speakers' groups (see examples (7)–(9)), 7 of which were found in the US data. Examples (7)–(9) are cases where NAct Voice appears on anticausative verbs that do not combine with this form. We found six such instances.

(7) Ena zevgari *__pernotan__ to dromo (USbi97FG_fs)
a couple cross.NAct.PST.3SG the street
'A couple was crossing the street.'

[2] By exporting the data that have the tag "nonact" there are some past participles tagged in the Voice tier as 'nonact', and this is the reason why the number of verbs appears to be smaller in Table 5.2. For comparison see the data stored in OSF: https://osf.io/hzqa3/.

Table 5.3 Categories of verbs found in the data presented in raw instances and frequencies

VERBS	HSs in the United States	HSs in Germany	Monolinguals
Deponent (excluding *ime* and *ginome*)	126 [-trans.] 1 [+trans.] (3.4)	230 [-trans.] 9 [+trans.] (6.5)	453 [-trans.] 17 [+trans.] (9.4)
Reflexive	2 (0.05)	21 (0.5)	55 (1.1)
Reciprocal	17 (0.4)	5 (0.1)	60 (1.2)
Passive	9 (0.2)	26 (0.7)	46 (0.9)
Non-target	7 [6 anticausatives and 1 + trans. deponent] (0.1)	3 [1 anticausative, 1 unaccusative and 1 +trans. deponent] (0.08)	0
N of tokens (Verbs)	3671	3661	4957

(8) Ena aspro aftokinito pu den ihe ora na
a white car that not had time SBJV
*__stamatithi__ (DEbi12FG_fw)
stop.NACt.3SG
'A white car that didn't have time to stop.'

(9) To aftokinito aspro *__spastike__ to (USbi07FG_is)
the car white break.NACt.PST.3PL the
aftokinito ble
car blue
'The white car crashed the blue car.'

Example (10) exhibits an unaccusative verb which remains [-] transitive but is combined with NAct Voice instead of canonical active.

(10) ihe figi ap ta heria tu ke (DEbi05FG_fs)
has gone from the hands his and
*__kilithike__ ap tin apenanti plevra
roll.NACt.3SG from the opposite side
'[The ball] has gone from his hands and rolled in the opposite side of the street.'

Examples (11)–(12) exhibit the two novel transitive deponent verbs bearing NAct:

(11) To ble aftokinito stamatise gia na min (DEbi12FG_fw)
the blue car stopped in order to not

*patithi tus anthropus
run over.NACT.3SG the people.

'The blue car stopped not to run over people.'

(12) Meta tis andres mesa stis aftokinites (USbi07FG_fw)
then the men in the cars

*eparthine to ennea mia mia
call.NACTPST.3PL the nine one one

'Then the men in the cars called 911.'

Table 5.4 Distribution of verbs in NAct in the different formality levels having excluded *ime* and *ginome*

Tokens of verbs in NAct	HSs in the United States	HSs in Germany	Monolingual
Formal setting	245	198	240
Informal setting	201	195	303

Regarding register variation, Table 5.4 presents the distribution of the verbs in NAct in the different formality settings. HSs produce more NAct verbs in the formal register while the opposite and unexpected pattern is found in monolinguals. The behavior of HSs does not align with our expectations in RQ4 as there is a slight differentiation, although it is not statistically significant in either of our HS groups; the exception here is participant USbi07FG who repeated the same non-target form in all four settings, something that was also predicted in RQ3 (second scenario).

Coming back to the non-target forms in Table 5.3, as these were produced by certain speakers only, namely two in the United States and one in Germany, we looked at the metalinguistic data of these speakers in some detail. With respect to the two HSs in the United States, when comparing their data to those of the whole group (Table 5.5), we observe that, at least for participant USbi07FG, her current input, the self-ratings in Greek, the literacy and the education in Greek both measured in years and in hours that she received are comparatively lower than the majority's average. This supports Polinsky's correlation of non-canonical marking and low proficiency. Regarding both USbi97FG and USbi07FG, we can observe that one of their parents is a second-generation immigrant, leading to interesting issues concerning the baseline these speakers were exposed to, which, however, cannot be answered in the context of our methodology.

Table 5.5 Metalinguistic data of the US group and the speakers who produced the non-target forms

		HSs in the United States	USbi07FG	USbi97FG
Age of onset		1;41 (min 0 – max 6) SD 1.956	0 years (simultaneous)	0 years (simultaneous)
Current input in Gr		0.789 (min 0.1 – max 1.5) SD 0.3482	0.25	0.75
Current input in AE		1.16 (min 0.5 – max 2.0) SD 0.3458	2	2
Self-ratings in Gr		3.59 (min 1.75 – max 5) SD 0.9369	1.75	4
Self-ratings in AE		4.91 (min 4 – max 5) SD 0.2153	4	4
Literacy practices in Gr		1.01 (min 0 – max 2) SD 0.5252	0.6	1
Literacy practices in AE		1.81 (min 0.3 – max 2) SD 0.3601	2	2
Parent's generation	Both 1st	30 prt		
	One 1st	3 prt		
	One 1st one 2nd	18 prt	☑	☑
	One 1st one American	2 prt		
	Both 2nd	8 prt		
	One 2nd	1 prt		
	One 2nd one American	1 prt		
Years of education in Gr		9;13 (min 0 – max 12) SD 3.401	1 year	10 years
Hours of education in Gr		2166.19 (min 0 – max 3120) SD 890.692	208 hours	2600 hours
Visits to the country of heritage		1.03 (min 0 – max 2) SD 0.314	1	1

Table 5.6 Metalinguistic data of HSs in Germany compared to the speaker who produced the non-target form

		HSs in Germany	DEbi12FG
Age of onset		1;9 (min 0 – max 8) SD 2.162	0 year (simultaneous)
Current input in Gr		1.1 (min 0.0 – max 1.8) SD 0.423	0.8
Current input in DE		0.82 (min 0.1 – max 1.6) SD 0.342	1.2
Self-ratings in Gr		4.08 (min 2.25 – max 5) SD 0.838	2.75
Self-ratings in DE		4.83 (min 3.5 – max 5) SD 0.362	5
Literacy practices in Gr		1.2 (min 0 – max 2) SD 0.551	1.6
Literacy practices in DE		1.52 (min 0 – max 2) SD 0.500	1
Parent's generation*	Both 1st	25 prt	☑
	One 1st one 2nd	3 prt	
	One 1st one German	16 prt	
	One 2nd one German	2 prt	
Years of education in Gr		7;2 (min 0 – max 12) SD 3.511	4
Hours of education in Gr		6155 (min 0 – max 12,480) SD 4605.401	4160

*One participant's data concerning the generation in which his/her parents belong are missing due to technical problems in retrieving the data.

Running the same comparison between the group of heritage speakers recruited in Germany and the participant that produced non-target forms (Table 5.6), we observe almost the same pattern as with participant USbi07FG,

namely lower self-rating score and reduced input and education received in Greek compared to the majority of the speakers in this group.[3]

5.7 Discussion

Turning now to a more detailed discussion of our results, what we found is actually the opposite of what has been observed in previous literature. Specifically, in our case, NAct appears on anticausative verbs that do not combine with NAct, and certain HSs form novel transitive deponent verbs bearing NAct, although this pattern is very restricted. These findings cannot be explained by language interference, as the two languages with which Greek is in contact, (a) lack Voice marking altogether and deponents (English), (b) do not mark the same set of anticausatives (German) with Voice morphology, and (c) while German, like Spanish, has inherent reflexive verbs which could be viewed as deponents (Alexiadou 2019 for discussion and references), the one verb that German HSs recategorize as deponent is not an inherent reflexive in German. We cautiously suggest that the production of non-target forms might be an outcome of low input in the heritage language, as the metalinguistic data show, and further exploration is needed. Regarding the slight tendency of HSs to produce more verbs bearing the NAct Voice in formal settings, we cannot make strong claims, as it is not statistically significant. However, this striking pattern needs further exploration in future studies. Register leveling is not at stake here unlike our predictions and previous literature introduced in Section 5.1.

Nevertheless, the question arises as to what enables such formation of these non-target forms to begin with. Recall that, building on Embick (1997), Alexiadou, Anagnostopoulou, and Schäfer (2015) proposed that NAct morphology realizes a Voice projection which lacks a specifier. In their analysis, NAct realizes a structure where Voice is marked [-D], signaling the absence of a specifier. In the absence of the Voice projection altogether or in the presence of a DP in Spec,VoiceP, Greek verbs are assigned default ACT Voice. The fact that certain HSs prefer the marked construction for anticausatives might indicate a movement toward a more canonical system, where all intransitive variants bear NAct. Basically, the structure in (13a) is an unaccusative structure and it is associated with passives, reflexives, and NAct-marked anticausatives. As

[3] The metalinguistic data from DEbi05FG, which produced a non-target form, could not be retrieved due to a technical problem.

this corresponds to a greater number of non-target forms, we suspect that such overgeneralization is a viable explanation:

(13) a. $[_{\text{MiddleVoiceP [-D]}}$ NAct $[_{\text{vP}} [_{\text{ResultP}} \sqrt{}\text{burn}]]]$ [Non-active]
b. $[_{\text{vP}} [_{\text{ResultP}} \sqrt{}\text{open}]]$ [Active]
c. $[_{\text{VoiceP}}$ DP $[_{\text{vP}} [_{\text{ResultP}} \sqrt{}\text{burn}]]]$ [Active]

The rule that regulates the spell-out of Voice in the absence of a specifier in this system is repeated in (14) below:

(14) Voice → NAct/_ (no specifier)

Unlike Zombolou and Alexiadou (2012), we did not observe a generalization of actives and we identified a correlation with low input and proficiency in Greek. We may speculate that this is so, as the non-active paradigm shows less allomorphy, and thus appears more canonical and preferred by HSs.[4] Being more canonical, it is especially preferred by speakers with low proficiency. It could very well be that the speakers in Zombolou and Alexiadou's study were not of low proficiency, but we do not know that. Nevertheless, which verbs will form anticausatives with non-active is highly idiosyncratic, making the overall distribution of non-active in Greek rather non-canonical. Thus, it might very well be that these two pressures lead to variation among participants: certain participants generalize active in order to avoid non-canonicity at the level of verb form choice. In other words, since some anticausatives bear active, all should bear active eliminating the need to choose the idiosyncratic forms. Others, however, identify the non-active form as the marker of intransitivity, thus generalizing this form to all verbs and eliminating allomorphy at the level of the verbal paradigm.

Matters are somewhat more complex when it comes to the creation of new transitive deponent verbs. Although we only found two such instances in our data, their formation is indeed surprising. It has been argued that transitive deponents belong to at least two different classes (Embick 1997): on the one hand, there are transitive deponents which are psych predicates, and specifically subject experiencer verbs; on the other hand, there are transitive deponents which are not psych verbs and in fact seem to mostly behave like canonical transitive verbs. Our speakers create the latter type of deponents.

[4] We are grateful to an anonymous reviewer for this suggestion.

According to Grestenberger (2018), the external argument in this case is introduced lower, i.e. it is a low agent, and Embick's NAct rule applies, see also Oikonomou and Alexiadou (2022) and Alexiadou (2024):

(15) [VoiceP -external argument [XP_{AGENT} [vP DP]]]

We would like to apply this analysis to the new deponents created by HSs. Since (15) is a possible structure in Greek, certain HSs may also generalize it: our novel verbs include an agentive Voice and NAct is the result of an agent being introduced lower than Voice. Oikonomou and Alexiadou (2022) suggest that perhaps Greek grammar productively uses this rule to derive novel interpretations of verbs. Voice forms a single interpretation domain with the vP with which it combines. These particular structures guide speakers to derive novel interpretations.

The final question here is why this option too would be chosen by low-proficiency HSs. We speculate that low proficiency correlates with what Scontras et al. (2018) call restructuring of heritage grammars. In our case, HSs restructure their grammar in making extensive use of a productive structure of Greek: it is the minimal domain that yields novel verbal interpretations.

5.8 Conclusions

We discussed two changes in the Voice system of certain HSs of Greek. Unlike what has been reported in previous literature, certain HSs generalize NAct on anticausatives and to a certain extent create novel deponent verbs generalizing a productive rule of Greek grammar. Their behavior supports views of NAct Voice as being productively used by speakers to derive novel interpretations of verbs. This non-default overregularization correlates with proficiency.

5.9 Appendix

Corpus queries:

Tokens in monolinguals: https://hu.berlin/tok_monos
Tokens in heritage speakers in Germany: https://hu.berlin/tok_DE_HSs
Tokens in heritage speakers in the United States: https://hu.berlin/tok_USA_HSs

Table 5.7 Elicitation orders

	Order 1(if/swsw)	**Order 2(if/wssw)**	**Order 3(if/swws)**	**Order 4(if/wsws)**
informal	spoken	written	spoken	written
	written	spoken	written	spoken
formal	spoken	spoken	written	written
	written	written	spoken	spoken
	Order 5(fi/swsw)	**Order 6(fi/wssw)**	**Order 7(fi/swws)**	**Order 8(fi/wsws)**
formal	spoken	written	spoken	written
	written	spoken	written	spoken
informal	spoken	spoken	written	written
	written	written	spoken	spoken

Total number of verbs in monolinguals: https://hu.berlin/verbs_in_monos
Total number of verbs in heritage speakers in Germany: https://hu.berlin/verbs_in_DE_HSs
Total number of verbs in heritage speakers in the United States: https://hu.berlin/verbs_in_USA_HSs
Verbs in active Voice in monolinguals: https://hu.berlin/tok_monos_active
Verbs in active Voice in heritage speakers in Germany: https://hu.berlin/tok_DE_HSs_active
Verbs in active Voice in heritage speakers in the United States: https://hu.berlin/tok_USA_HSs_active
Verbs in non-active Voice in monolinguals: https://hu.berlin/tok_monos_nact
Verbs in non-active Voice in heritage speakers in Germany: https://hu.berlin/tok_DE_HSs_nact
Verbs in non-active Voice in heritage speakers in the United States: https://hu.berlin/tok_USA_HSs_nact
Matches of *ime* in monolinguals: https://hu.berlin/tok_monos_ime
Matches of *ginome* in monolinguals: https://hu.berlin/tok_monos_ginome
Matches of *ime* in heritage speakers in Germany: https://hu.berlin/tok_DE_HSs_ime
Matches of *ginome* in heritage speakers in Germany: https://hu.berlin/tok_DE_HSs_ginome
Matches of *ime* in heritage speakers in the United States: https://hu.berlin/tok_USA_HSs_ime
Matches of *ginome* in heritage speakers in the United States: https://hu.berlin/tok_USA_HSs_ginome

Matches of verbs in NAct in formal setting in monolinguals: https://hu.berlin/nact_verbs_in_monos_formal

Matches of verbs in NAct in informal setting in monolinguals: https://hu.berlin/nact_verbs_in_monos_informal

Matches of *ime* and *ginome* in formal setting in monolinguals: https://hu.berlin/ime-ginome_monos-formal

Matches of *ime* and *ginome* in informal setting in monolinguals: https://hu.berlin/ime-ginome-monos-informal

Matches of verbs in NAct in formal setting in heritage speakers in Germany: https://hu.berlin/nact_verbs_in_DE_HSs_formal

Matches of verbs in NAct in informal setting in heritage speakers in Germany: https://hu.berlin/nact_verbs_in_DE_HSs_informal

Matches of *ime* and *ginome* in formal setting in heritage speakers in Germany: https://hu.berlin/ime-ginome_DE_HSs-formal

Matches of *ime* and *ginome* in informal setting in heritage speakers in Germany: https://hu.berlin/ime-ginome_DE_HSs-informal

Matches of verbs in NAct in formal setting in heritage speakers in the United States: https://hu.berlin/nact_verbs_in_US_HSs_formal

Matches of verbs in NAct in informal setting in heritage speakers in the United States: https://hu.berlin/nact_verbs_in_US_HSs_informal

Matches of *ime* and *ginome* in formal setting in heritage speakers in the United States: https://hu.berlin/ime-ginome_US_HSs-formal

Matches of *ime* and *ginome* in informal setting in heritage speakers in the United States: https://hu.berlin/ime-ginome_US_HSs-informal

Acknowledgments

We are indebted to two anonymous reviewers and the editors of this volume for their comments. Our research was funded by the Deutsche Forschungsgemeinschaft grants AL 554/13-1, 394,836,232 and AL 554/15-1, 313,607,803 to Artemis Alexiadou. The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the DGfS Ethics Committee of the Deutsche Gesellschaft für Sprachwissenschaft (German Society for Linguistics). Dates of approval: January 12, 2017, and February 1, 2019.

6

The shape and size of defective domains

Non-finite clauses in Pennsylvania Dutch

Michael T. Putnam

6.1 Introduction

A guiding intuition concerning the syntax of heritage languages (HLs) is that this module of grammar remains ostensibly constant and unaffected over the course of the lifespan (Lohndal 2021). The syntactic properties of HLs thus show a greater deal of resistance to change and attrition when compared with other domains of grammar, such as morphology (Lohndal 2021; Polinsky 2018; Putnam, Schwarz, and Hoffnam 2021) and phonology (Chang 2021). Although "core" operations of syntax, such as Merge, remain intact, a number of proposals suggest that computational domains and the determination of locality more generally can be reduced. In their review of the general properties and developmental trends of HLs, Polinsky and Scontras (2020) highlight a number of factors that may influence the development of HLs across generations. Two of these, namely, **economy of online resources** and **shrinking of structure**, are of central importance for this chapter. The general tendency in HL syntax research to date has been to interpret the development of syntax in these grammars as the "shrinking" of computational domains, due to Representational Economy factors (Putnam 2020; Scontras, Polinsky, and Fuchs 2018) or an overall reduction in systemic economy (Lohndal and Putnam 2021). Studies focusing on long-distance *wh*-movement (Hopp, Putnam, and Vosburg 2019) and long-distancing anaphoric binding (Gürel 2004; Kim, Montrul, and Yoon 2009; Putnam and Arnbjörnsdóttir 2015) confirm the reduction of computational domains in HL syntax.

In this chapter, I investigate a series of related phenomena that pose somewhat of a challenge to the notion that structure always *shrinks*. In fact, I advance the claim that not only do we observe instances of domain-*shrinking*, but also domain-*stretching* in HL syntax. The empirical focus of this chapter centers on non-finite clauses, i.e. *defective domains* in Pennsylvania Dutch

Michael T. Putnam, *The shape and size of defective domains*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro et al., Oxford University Press.
DOI: 10.1093/9780191987731.003.0007

(PD).[1] For the remainder of this chapter, the *defective domains* we primarily concern ourselves with are infinitives and gerunds. In spite of sharing many similarities with the clausal syntax of standard German and its continental dialects, two elements of defective domains in PD that distinguish it from German require further attention (Börjars and Burridge 2011; Huffines 1986, 1990; Louden 1988, 1997, 2005, 2016b, 2019; van Ness 1994). The first of these concerns the loss of the infinitival marker *zu* 'to' in PD, which has been replaced by the preposition/complementizer *fer* 'for', which appears at the leftmost edge of the non-finite clause, i.e. in C, see (1).

(1) Es iss hatt [$_{CP}$ fer's Buch uffpicke].
it is hard INF=the.NEUT book up-pick
'It is difficult/hard to pick up the book.'

Second, although clausal gerunds are not permitted in PD (2a), defective clausal gerunds frequently occur (2b) (see Louden 2019 for supporting arguments).

(2) a. *[$_{CP}$ Rose zu die Hochzich komme] war n
Rose to the wedding come.NF was a
Surprise zu alliebber.
surprise to everyone
Intended: 'Rose showing up at the wedding was a surprise to everyone.' [Clausal gerund]

b. Ich bin n browiere [$_{VP}$ die Daer uffmache].
I am PROG try the door open.NF
'I am trying to open the door.' [Defective clausal gerund]

The two aspects of defective domains in PD that are the focus of this chapter, namely, (i) the "loss" of the infinitive marker *zu* and (ii) the licensing of defective clausal domains, require an appeal to both the reduction *and* expansion of syntactic domains. Turning first to the need for clausal-expansion, similar

[1] PD is "a language that has outgrown its name" (Keiser 2012: 1), to the extent that it has grown exponentially beyond the borders of southeastern Pennsylvania over the past 325 years. Although it is the native language (L1) of over 400,000 individuals today throughout the Northern Hemisphere, PD should still be considered an HL in the broader sense, based on two fundamental criteria: First, it is not the sociolinguistically dominant language of society, hence, there are a restricted number of domains in which the language is used. Second, and related to the previous point, PD-children receive truncated L1-input both from a qualitative and quantitative perspective, since they begin acquiring English when they begin their formal education, usually at the age of 6, and often sooner if they have older siblings, i.e. sometimes as early as 3–4 years of age.

phenomena related to the *stretching* of defective domains include the lexicalization of an overt element in C, which has been observed in other extra-territorial varieties of Germanic such as American Norwegian (Putnam and Søfteland 2022) and Brazilian Pomeranian (Postma 2014), with the latter case also resulting in the loss of an infinitival marker. This strategy of domain expansion reflects a desire to establish structural salience (Polinsky 2018: 174–175) in HL syntax, and as we will discuss in this chapter, avoid substantial conflicts between dueling representational preferences in the minds of bilingual speakers. In the second case, the emergence of defective clausal gerunds in PD is analogous to the historical development of gerunds in English from locative PPs (Alexiadou 2013; Fischer, De Smet, and van der Wurff 2017; Higginbotham 2009). If this comparison holds, we observe the *shrinking* of the structural workspace of PD (D'Alessandro, Natvig, and Putnam 2021; Polinsky and Scontras 2020), with respect to the development of verbal defective clausal gerunds in PD. Our analysis of these nuances in defective domains of contemporary PD-syntax stands to have a broader impact on theoretical approaches to these grammars, as laid out in the following research question (RQ) below:

- RQ: How does the occasional expansion of clausal structure situate with the general tendencies in HL syntax to shrink computational domains?

In line with recent proposals by Biberauer (2018) and Putnam and Søfteland (2022, 2024), we highlight two motivations for clause-*stretching* in HL syntax. The first concerns a tendency to provide lexical material in the position of the leftmost functional head in extended projections. Second, and more specific to Germanic syntax, instances of clause-*stretching* are carried out in order to avoid TP-projections to house subjects, which German and related languages use sparsely when compared with contemporary English (Haider 2010: Chapter 7). Here I discuss how this tendency can be united with the conceptualization of locality and selective opacity in current minimalist theorizing as well as the existing parallels between diachronic and HL syntactic development and change.

The structure of this chapter is as follows: In the next section (6.2), we review the general syntactic properties of defective domains, based largely on recent proposals by Pires (2007), Wurmbrand (2001, 2014), and Wurmbrand and Lohninger (2023). We illustrate the shape and size of various defective domains in English, making the case that defectives domain project a wide gamut of structures (e.g., VP, *v*P, TP, and CP). In Section 6.3, we review

the general properties of defective domains in PD, contrasting the general properties of PD with English and (standard) German. We then transition into questions regarding the historical development of verbal gerunds in PD in Section 6.4, discussing and establishing interesting parallels that exist between the emergence of verbal gerund in both English and PD. Following this brief historical treatment of verbal gerunds, we consider how this account of non-finite structures in PD contributes to larger questions concerning the structural development of HLs in Section 6.5, paying particular attention to the central role of representations in theoretical treatments of HLs.

6.2 The syntax of *defective domains*

With respect to the syntax of *defective domains*, in this chapter we concentrate on the following phenomena:

(3) Defective domains in English

a. *It began to rain.*	[Infinitive]
b. *I heard her singing in the shower.*	[AcI]
c. *Mike delivering a decent presentation was a surprise.*	[Clausal gerund]
d. *I tried opening the door.*	[Defective clausal gerund]
e. *I expect him to order a whiskey sour.*	[ECM]
f. *He seems to be upset that the Steelers lost.*	[Raising]
g. *He tried to quit yelling at his kids in public.*	[Control]

Here we undertake a brief review of some of the general properties of defective domains in English, i.e. the examples in (3). The primary point of emphasis here is that the size of the clausal structure for these various defective domains differs substantially, both in English and cross-linguistically (Pires 2007; Wurmbrand 2001, 2014; Wurmbrand and Lohninger 2023). Here I make a point of comparison with English in this section and an occasional reference to (standard) German non-finite structures. Admittedly, German is not a suitable baseline for PD[2], however, as we shall see through this comparison in this and the subsequent section, PD (still) shares a number of syntactic characteristics with German.

Non-finite/defective domains exhibit different underlying structural properties. Abundant literature exists debating the position(s) of the infinitival

[2] See Polinsky (2018: Section 1.4) and D'Alessandro et al. (2021) for discussion of the difficulties of establishing a suitable baseline in HL research.

marker *to* in English (see, e.g., Levine 2012). Minimally, the position of English *to* as been argued to occupy the following positions:

- Adjoined to *to*P immediately dominating *v*P,
- As T, or
- As C

Turning to defective domains that do not require the presence of *to*, AcI-predicates (*accusativus cum infinitivo*) lack of both *to* and an overt subject, which suggests that AcIs are best understood as *v*Ps, see, e.g., (4).[3]

(4) a. I heard Marsha [$_{vP}$ singing in the shower.
b. I saw Peter [$_{vP}$ throwing the football in his backyard.

German, which is not an EPP-language (in that nominative subjects do not occupy Spec,TP) also licenses AcIs, as shown in (5).[4]

(5) Ich habe ihn sprechen hören/gehört.
I have him speak hear/heard
'I heard him speaking/talking.'

Gerunds, which also do not license the presence of infinitival-*to*, can be split into two distinct sub-classes, namely, (i) *clausal gerunds* and (ii) *defective clausal gerunds*. Consider the contrast between the control predicate licensing PRO in (6a) and the clausal gerund in (6b).

(6) a. Carol worried about [PRO being late for dinner.
b. Carol worried about [Jim being late for dinner.

Pires (2007: 16) provides support for treating clausal gerunds as units that project a TP. As illustrated in (7) (again, adopted from Pires 2007: 16), clausal gerunds (in English) can appear as (i) complements to verbs (7a), (ii) as complements to prepositions (7b) and (7c), and (iii) as phrases in "subject position" (7d). More directly, clausal gerunds appear in what are traditionally classified as "Case positions," i.e. those in which the arguments that reside within them receive structural Case (either nominative or accusative).

[3] Lee-Schoenfeld (2007) discusses binding evidence that suggests a further distinction between *v*P and VP projections with respect to AcIs in German.
[4] See Haider (2010: Chapter 7) for evidence in support of this claim.

(7) a. Mary favored [Bill taking care of her land].
b. Susan worried about [Mark being late for dinner].
c. Sylvia wants to find a new house without [Anna helping her].
d. [Sue showing up at the game] was a surprise to everybody.

The "Case position" argument advanced by Pires (2007) also explains why clausal gerunds can occur as complements to prepositions (8b), unlike finite and infinitive clauses (8a). Clausal gerunds thus behave similar to possessive *-ing* and DPs, as shown in (8c) (data from Pires 2007: 21).

(8) a. *Mary talked about [(that) John moved out/ John to move out].
b. Mary talked about [John moving out].
c. Mary talked about [John's moving out / John's move].

In contrast, *defective clausal gerunds* appear to have a more reduced syntactic structure. They can appear as (gerund) complements of aspectualizers (e.g., *start, finish*, and *keep*) and verbs such as *try* and *avoid* (Pires 2007: 70).

(9) a. Mary started/finished/continued [reading the newspaper].
b. Bill_j tried [e_j talking to his boss].
c. Philip_j avoids [e_j driving on the freeway].

Due to (i) their lack of independence regarding tense and aspect in relation to the matrix clause and (ii) the questionable status of PRO, these *defective clausal gerunds* are generally regarded to be relatively small, i.e. VPs, and as we shall argue, DPs.[5] Wurmbrand (2001, 2014) and Fukuda (2012) provide cross-linguistic evidence for the reduced syntactic structure of *defective clausal gerunds*.

Finally, we conclude our brief overview of defective domains in English by taking a closer look at ECM (Exceptionally Case Marked) (10), Raising (11), and Control predicates (12).

(10) Mike expected [him to win the game]. [ECM]

[5] One could plausibly make the argument that *defective clausal gerunds* consist of a TP-projection with a "defective" head (with "null" Tense), which Pires (2007: Chapter 3) suggests. He supports this analysis based on fact that the absence of a TP-projection altogether in *defective clausal gerunds* falls to account for the assignment of null Case (which commonly occurs in Spec,TP). We do not pursue this proposal here, opting rather for recent proposals put forward by Alexiadou (2005), Alexiadou, Iordăchioaia, and Soare (2010), and Alexiadou (2013), which interpret verbal gerunds as DPs that contain an internal AspP. We discuss the relevance of this in Sections 6.4 and 6.5.

Table 6.1 Defective domains in English

Defective domain	Projection
AcI	*v*P
Clausal Gerund (CG)	TP
Defective CG	VP (or "defective" TP)
ECM	TP
Raising	TP
Control	CP

(11) Cindy seems [to be sick]. [Raising]

(12) a. Jan convinced Cindy$_j$ [PRO$_j$ to taddle on Marsha]. [Object control]

b. Bobby$_j$ tried [PRO$_j$ to eat more ice cream than his brothers]. [Subject control]

Based on long-standing assumptions, we assume that nominal elements that receive accusative case in ECM-structures are in Spec,TP (i.e., *him* in (10)). In similar fashion, raising predicates are also held to TPs, while control predicates require a CP-complement (in order to license PRO).

The proposed underlying projections of defective domains in English are summarized in Table 6.1. Following the proposals of Wurmbrand (2001), Pires (2007), and Fukuda (2012), we see that, at least for English, the projection of TP is suggested for (at least) clausal gerunds, ECM-predicates, and Raising-predicates, and possible defective clausal gerunds.

In the following section, we discuss the syntax of defective domains in Pennsylvania Dutch, showing that in many respects, PD has overwhelmingly retained its German(ic) syntactic heritage.

6.3 Defective domains in Pennsylvania Dutch: an overview

The syntax of defective domains in Pennsylvania Dutch is the primary focus of this section. Once again, for the sake of typological comparison, English and (standard) German are called upon to establish to what extent PD has retained "German-like" syntax with respect to these structures. In Section 6.3.1, we discuss a historical development in PD, namely, the loss of *zu/ze* as a marker of non-finiteness. We turn to more synchronic data in Section 6.3.2, where we show, that although PD has largely retained a more German-like syntax with

respect to its licensing of defective domains, it has developed a *defective clausal gerund*, which is present in modern English yet absent in German. Before providing more detailed comments on the PD data, it is worth pointing out that the majority of these examples were elicited through informal interviews with L1 speakers of PD.

6.3.1 The loss of *zu* as a marker of non-finiteness

In their work on Pennsylvania Dutch spoken in Waterloo County, Canada (WCPD), Börjars and Burridge (2011) demonstrate that the grammaticalization of non-finite clauses headed by *fer* originates from purposive clauses similar to those found in modern Standard German as shown in (13).

(13) Wir haben Milch gekauft [$_{CP}$ **um** einen Kuchen **zu** backen].
we have milk bought C a cake INF bake
'We bought milk in order to bake a cake.' [Standard German]

Many dialects and non-standard varieties of German make use of the complementizer *fer* in these purposive constructions rather than *um*, which is found in Standard German. Louden (1988, 1997) and van Ness (1994) classify *fer* in Pennsylvania Dutch as a generalized complementizer, and we will see that this assertion holds in most structures found in our elicited data. Börjars and Burridge (2011) analyze a path toward grammaticalization in which this initial element in the historical purposive clause, namely *fer*, becomes the head of a clause that signifies non-finiteness, allowing for the traditional infinitival marker *zu* to be eventually reduced and eliminated from the grammar. Roughly at the turn of the previous century, PD "lost" *zu* as an infinitival marker. Consider the historical examples in (14) (cited by Börjars and Burridge 2011: 390–391).

(14) a. **Fer** Sauder **zu** haysa is doch gar ke Shand.
for Sauder INF be-called is but absolutely no shame
'To be called Sauder is no shame at all.'
(Sauder 1955)

b. Se wore ols so shlim **fer** danse.
they were always so eager INF dance.NF
'They were always so eager to dance.'
(Horne 1905)

Huffines (1986, 1990) reports that the usage of both *fer* and *zu* simultaneously to mark a non-finite/defective domain is extremely rare today, restricted mostly to the speech of remaining speakers of non-sectarian PD, which is now moribund. As we shall discuss later in this chapter (Section 6.5.1), the "loss" of the infinitival marker *zu* is not restricted to PD; in fact, the proposed stages of this syntactic change, outlined in (15), receives a straightforward explanation.

(15) Three stages of development of *fer* in PD (Börjars and Burridge 2011: 395)

a. *fer* + NP benefactive
b. *fer* + clause purposive
c. *fer* + clause general infinitival

This historical evidence shows that the *zu*-infinitival marker did previously exist in Pennsylvania Dutch, which is unlike other closely related German dialects such as Alemannic which have never realized this infinitival marker (see Brander 2020 for a discussion). Additionally, Huffines' (1990) study of purposive clauses and the continued usage of *zu*-infinitive in Pennsylvania Dutch spoken in Pennsylvania concluded that the relatively sporadic use of either of these options was produced predominantly by elderly non-sectarian speakers of the language. For the native speakers who provided us with preliminary data for this study, the use of *zu* as an infinitival marker is no longer a viable option in their grammar. This is best illustrated with the elicited translation of the *tough*-construction *It is difficult to pick up the book*:

(16) a. Es iss hatt [$_{CP}$ fer's Buch uffpicke.
it is hard INF=the.NEUT book up-pick
'It is difficult/hard to pick up the book.'

b. *Es iss hatt ['s Buch uff**fer**picke.

c. *Er iss hatt ['s Buch uff**zu**picke.

Louden (1988: 210–212) remarks that the bare infinitive in PD correlates closely to gerund complements in English, while the non-finite clauses headed by *fer* approximate to *to*-infinitives in English. Börjars and Burridge (2011) confirm that this also holds for the variety of Pennsylvania Dutch spoken in Waterloo County, Canada.

It is also worth pointing out that the infinitival markers *to* and *zu* in English and German respectively do not occupy the same structural positions, cf. (17).[6] Assuming that the syntax of PD (strongly) resembles German in this respect, the infinitival marker *zu* can/could occupy multiple positions in the syntax (Wilder 1989; Wurmbrand 2001, 2014; Wurmbrand and Lohninger 2023).

(17) a. Maria decided to carefully remove the bigger splinter. [English]

b. Maria beschloss den größeren Splitter vorsichtig zu
Maria decided the bigger splinter carefully to
entfernen / *zu vorsichtig entfernen.
remove to carefully remove
'Maria decided to remove the bigger splinter.' [German]

6.3.2 Non-finiteness in modern Pennsylvania Dutch

In many respects, PD has maintained a significant number of core traits of German(ic) syntax. For example, as illustrated in (18), PD permits clausal union between two matrix clauses (in which the finite verb appears in the second position in both clauses) (18a), and finite verbs appear in the final position of the clause (18b).[7]

(18) a. Die Sarah meent [$_{CP}$ die Rose schwetzt Deitsch gut.
the Sarah thinks the Rose speaks Dutch well
'Sarah thinks Rose speaks Dutch well.'

b. Die Sarah denkt, [$_{CP}$ ass die Rose gut Deitsch
the Sarah thinks that the Rose good Dutch
schwetze kann.
speak can
'Sarah thinks that Rose can speak Dutch well.'

In light of our brief overview of the underlying syntactic properties of defective domains in Section 6.2 above, PD makes use of multiple syntactic projections

[6] Given that the subject is expressed in German, English, & PD but PRO in control predicates, this supports the hypothesis that AcI-constructions are *v*Ps underlyingly.

[7] In connection with discussions pertaining to the leftmost element of clausal structure, see Fitch (2011) for evidence that PD allows extraposed temporal PPs more frequently than other continental German vernaculars.

to represent non-finiteness much like English and (standard) German. Louden (2016a, 2019) has proposed the following distribution regarding the marking of defective domains in Pennsylvania Dutch:

(19) Louden's (2016a, 2019) generalizations
   a. Infinitival constructions are introduced with *fer* 'for' or are phonologically null Ø
   b. The distribution of these two options is dependent on semantically equivalent English expressions
   c. If the infinitival *to* is required in English, *fer* must occur in PD
   d. If English requires or permits a gerund or bare infinitive, *fer* is omitted

Although these generalizations hold in the data reviewed below, in this section I (i) highlight here that the variation found in contemporary PD non-finite clauses is primarily *syntactic* in nature and, in turn, (ii) I provide an overview of the syntactic properties of defective domains in PD with an eye toward how this impacts HL syntactic development and change more generally.

6.3.2.1 *Fer* as a complementizer in PD

As a starting point, PD licenses AcI-constructions, which is to be expected given that both English and (varieties of) German also license them. Syntactically speaking, this means that PD also makes use of the *v*P-projection in expressing defective domains without an infinitival marker, as illustrated in (20).

(20) Die Sarah hot die Rose gheert [$_{vP}$ gut Deitsch schwetze.
the Sarah has the Rose heard well Dutch speak.NF
'Sarah hear Rose speak(ing) Dutch well.'

As discussed above, the complementizer *fer* – which appears in C – has assumed the duties of the infinitival marker *zu* in contemporary (sectarian) PD, since the latter marker is no longer realized in modern variants of PD. We take this as evidence that *fer* is the lexicalization of C in (21), especially since PD still strongly adheres to the V2-preference, which means that subjects appear in the CP-layer (Spec,FinP) rather than Spec,TP.

(21) a. Der Tim hat gemeindt [$_{CP}$ fer die Bicher wegduh.
the Tim has remembered INF die books away-make
'Tim remembered to put the books away.'

b. Die Kinner hen admit [CP fer's Fenschder
the children have admitted INF=the.NEUT window
verbroche hawwe mit Schtee.
broken have with stone
'The children admitted to break the window with (a) stone.'

It is interesting to observe that *wh*-infinitives, which are ungrammatical in German and most of its variants, are perfectly acceptable in PD; see, e.g., (22). Putnam and Søfteland (2024) found similar behavior in North American Norwegian, which licenses *wh*-infinitives, although these are ungrammatical in European Norwegian varieties. This once again can be interpreted as a general trend to lexicalize infinitival markers at the edge of clauses (in this case again, C) in heritage Germanic syntax (see also Section 6.5).

(22) Ich wees nett [CP was fer tun.
I know not what INF do
'I don't know what to do.'

Fer-clauses can also occupy the subject position as in (23); however, there are environments where *fer* is not required (24).[8]

(23) [CP Fer happy sei] misse die Kinner gut schloofe.
INF happy be.NF must the children well sleep
'To be happy the children must sleep well.'

(24) Die Lisa hat vergesse [CP (*fer) der Allen saage wege
the lisa has forgotten INF the Allen say.NF about
der Gaul.
the horse
'Lisa forgot to tell Allen about the horse.'

Turning to control, raising, and ECM-predicates, we see that PD licenses control predicates with defective domains headed by *fer*. This is demonstrated with the example of object control in (25).

[8] An anonymous reviewer poses an interesting question as to what condition(s) the ban the occurrence of *fer* in (24). At this point, anything I would offer would be mere speculation, since, according to Louden's (2016b, 2019) observation, this *should* be possible. I leave this interesting question for future research.

(25) Ich haawe ihn verschwetzt [$_{CP}$ fer uffheere schmoke.
I have him convinced INF stop smoke.NF
'I convinced him to stop smoking.'

There does not seem to be an equivalent to subject-to-subject raising in Pennsylvania Dutch. As we see in (26), in spite of the appearance of quasi-raising predicates such as *seemt* 'seems' and *act* 'acts', and *guckt* 'looks', a CP-clause headed by a *wh*-element containing a finite verb is still required.

(26) a. Der John seemt / guckt [$_{CP}$ wie er grank iss.
the John seems look like he sick is
'John seems to be sick / looks like he is sick.'

b. Der John act / guckt [$_{CP}$ wie er zu der Schtoor
the John acts looks like he to the store
geh will.
go wants
'John acts like he wants to go to the store.'

This point notwithstanding, we do find some examples, such as those in (27), in which the subject of the embedded clause "raises" from the non-finite clause into the matrix clause (27a), and those that include what appears to be the combination of an ECM-predicate structure (with accusative cause being assigned to *ihn* 'him') in addition to the CP-phrase complement (27b).[9]

(27) a. Niemand expect teachers [$_{CP}$ fer perfect sei.
no one expects teachers INF perfect be
'No one expects teachers to be perfect.'

b. Er weest, ass ich ihn expect haawe [$_{CP}$ fer die
he knows that I him expect have INF the
Daer schliesse.
door shut.NF
'He knows that I expected him to lock/shut the door.'

Generally speaking, English-style ECM is avoided in contemporary spoken Pennsylvania Dutch. These structures are usually realized as either a subordinate clause or a conjoined main clause, as in (28). Although much more

[9] For a more detailed and nuanced discussion of ECM-predicates across Germanic, see Wurmbrand and Christopoulos (2020).

research needs to be conducted on these structures in PD, a noticeable trend is clear even at this juncture; that is, while non-finite clauses that license control PRO (which are CPs) are possible in PD, raising and ECM-predicates, which project TPs, are ungrammatical, or circumvented to avoid projecting a TP.

(28) Die Bawwer meene [$_{CP}$ d/seller Gaul iss grank.
the farmers think/consider this horse is sick
Intended: 'The farmers consider this horse to be sick.'

6.3.2.2 Gerunds in PD

We now turn to gerunds, which, as indicated in (19) above, PD tends to realize as the absence of the complementizer/infinitival marker *fer*. The sentences in (29), which were all provided and evaluated by L1 speakers of sectarian PD, contain the phrasal verb *meinde vun* 'remember' in the matrix clause.[10] Note that example (29c) is ill-formed, due to the assignment of accusative Case on the pronoun (which would resemble an ECM-structure).

(29) a. Der Tim hot gemeindt vun [$_{XP}$ die Bicher wegduh.
the Tim has remembered P the books away-make
'Tim remembered putting the books away.'

b. Die Sarah meindt vun [$_{XP}$ die Rose gut Deitsch
the Sarah remembers P the Rose well Dutch
schwetze.
speak.NF
'Sarah remembers Rose speaking Dutch well.'

c. Ich meind(e) noch vun [$_{XP}$??ihn mit uns zu die
I remember still P him mit us to the
Gmee geh.
church go.NF
'I remember him going to church with us.'

The XP-complement of *meinde vun* is a gerund-phrase and not a conjoined matrix clause, which we can confirm due to the non-finite (NF) at the right edge of the gerund phrase. In fact, as we see in example (30), combining a matrix clause with a finite predicate in non-final position is ungrammatical.

[10] Although my informants reported that the preposition is obligatory for them, Mark Louden (p.c.) reports that some speakers, and some dialects of PD, allow the preposition *vun* to be elided.

(30) *Die Sarah meindt vun [XP die Rose **schwetze** gut
the Sarah thinks P the Rose speak.NF well
Deitsch.
Dutch
'Sarah remembers Rose speaking Dutch well.'

The XP-projection that houses these gerunds in PD is quite small, likely consisting of nothing more than a bare VP. We base this claim on data from Putnam and Rocker (2019), which focused on the syntactic and semantic properties of eventive aspectualizers – such as *schtaerte* 'start' – in PD. Similar to what is found in vernacular German (Van Pottelberge 2004, 2007), PD expresses progressive and imperfective aspect with the morpheme *an/m* which appears to the left of the verb phrase as shown in (31) below. Of particular relevance, examples (31c) and (31d) confirm that the complement of the progressive marker and eventive aspectualizers is likely nothing more than a VP (due to the lack of *fer*).

(31) a. Die Ime schtaerte ihn nochgehe.
the bees start him after-go.NF
'The bees start going / to go after him.'

b. und er is an schtaerte fer gehe.
and he is PROG start INF go.NF
'and he is starting to go (*going).'

c. 'S is an schtaerte reggere.
it is PROG start rain.NF
'It is starting (*to start) to rain (*raining).'

d. Ich haawe welle uffheere [XP schmoke.
I have want stop smoke.NF
'I wanted to stop smoking (*to smoke).'

e. Er hot aaghalde [XP in die Scheier schaffe der gans
he has kept in the barn work.NF the whole
Daag.
day
'He kept working (*to work) in the barn all day long.'

A critical point that cannot be overstated is that the previous examples of gerunds in PD all meet the requirements of being classified as *defective clausal gerunds*. The data in (32) highlight this contrast when we compare an example with progressive aspect (32a) and one in present perfect (32b). The

example in (32a) is usually strongly marked, if not fully ungrammatical, in most German-based vernaculars, but is quite common in English.

(32) a. Ich bin n browiere [$_{DP}$ die Daer uffmache.
I am PROG try the door open.NF
'I am trying to open the door.'

b. Ich haawe browiert [$_{DP}$ die Daer uffmache.
I have tried the door open.NF
'I tried to open the door.'

The situation concerning *clausal gerunds* in PD, however, is quite different. In fact, *clausal gerunds* are ungrammatical in contemporary sectarian PD. The data in (33) confirm this assumption. Compare these PD examples with those previously cited from Pires (2007: 16) in (7), repeated below as (34) for the sake of the reader. Recall that clausal gerunds in English can occur as complements to verbs (34a), complements of prepositions (34b,c), and in subject position (34d) in English. None of these conditions hold in PD, cf. (33d) and (34a) as complements of verbs, (33e) and (34b,c) as complements of prepositions, and (33f) and (34d) in subject position.[11]

(33) a. *Ich meind vun [$_{CP}$ Sally zu die Gmee geh mit
I remember P Sally to the church go.NF with
uns.
us
'I remember Sally going to church with us.'

b. Ich meind (vun) [$_{CP}$ wann die Sally in die Gmee
I remember P when the Sally in the church
gange is mit uns.
gone is with us
'I remember when Sally went to church with us.'

[11] An anonymous reviewer makes the observation that although PD appears to have little problem with structures containing obligatory control, it is questionable based on these initial data whether this is also the case with constructions with arbitrary control (pro *to watch TV is not allowed*) or those that have an overt subject (*For Sally to come to the wedding was unexpected*). This reviewer suggests that we could perhaps advance the claim that different types of subjects in the absence of finiteness are relativized for HL speakers. Along these lines, PRO is fine because it is easy to identify and bind, with little-*pro* being somewhat difficult, and object subjects being particularly challenging. This proposal is indeed quite interesting and should be the focus of future research in PD-syntax.

c. Ich meind noch vun [$_{CP}$ zu die Gmee laafe.
I remember still P to the church run.NF
'I still remember running to (the) church.'

d. *Sarah wett n neier Haas finne [$_{CP}$ ohni Rose
Sarah wants a new house find without Rose
sie helfe.
her help.NF
Intended: 'Sarah wants to find a new house without Rose helping her.'

e. *Sarah worry wege [$_{CP}$ Sally spät zu Owetesse
Sarry worries about Sally late to dinner
komme.
come.NF
Intended: 'Sarah worries about Sally coming/being late for dinner.'

f. *[$_{CP}$ Rose zu die Hochzich komme war n Surprise
Rose to the wedding come.NF was a surprise
zu alliebber.
to everyone
Intended: 'Rose showing up at the wedding was a surprise to everyone.'

(34) a. Mary favored [Bill taking care of her land].
b. Susan worried about [Mark being late for dinner].
c. Sylvia wants to find a new house without [Anna helping her].
d. [Sue showing up at the game] was a surprise to everybody.

In summary, PD-speakers have integrated *defective clausal gerunds* into their syntax, while not developing structures that license *clausal gerunds* as they appear in English. Table 6.2 summarizes the non-finite clauses licensed by English, German, and PD.

A quick comparison of these three languages reveals once again that PD has retained many German(ic)-syntactic attributes. Most relevant for our treatment of defective domains in PD, this language seems to lack structures that are headed by a TP-projection (i.e., ECM, Raising, and clausal gerunds). Bearing these generalizations in mind, in the following section we turn to the formal analysis of these defective domains in PD and investigate to what extent the emergence of defective gerunds echos the diachronic development of similar constructions in the history of English.

Table 6.2 English-PD-German non-finite clauses

	English	Penn Dutch	German
AcI	✔	✔	✔
Clausal Gerund (CG)	✔	✗	✗
Defective CG	✔	✔	✗
ECM	✔	✗	✗
Raising	✔	✗	✗
Control	✔	✔	✔

6.4 The emergence of the verbal gerund in Pennsylvania Dutch

As seen in the previous section, PD continues to strongly reflect its German(ic) heritage in its syntactic representations (see, e.g., Table 6.2 immediately above). Contemporary PD syntax does not project Spec,TP as a licit landing site for subjects, which provides a straightforward account as to why *defective clausal gerunds* have emerged in this language, while *clausal gerunds* have not. In fact, the size of defective domains in PD varies much like English and German, however like the latter, TPs are not projected and associated with these defective domains. PD licenses (i) AcIs (*v*P), (ii) *defective clausal gerunds* (VP), and (iii) *fer*-clauses (CP) (for *to/zu*-infinitives and control predicates).

Avoiding projecting (Spec,)TP in modern PD delivers a rational explanation for **why** *defective clausal gerunds* developed in PD, but it doesn't explain **how** this may have occurred. In addition to concerns regarding the shrinking of computational domains in HL grammars, another interesting factor which has regained traction in recent research is the potential symmetry between diachronic developments and what we observe in HL settings (Kupisch and Polinsky 2022).[12] A foundational question here concerns not only whether similarities exist in the trajectory of observed changes in diachronic and HL settings, but also whether or not we witness these changes at an exacerbated pace in the latter.

In this section, we review the historical development of gerunds in modern English, showing certain parallels that exist between the development of gerunds in English and PD. More specifically, we focus on the emergence of *verbal gerunds* in the history of English, which are not present in

[12] See previous work by Lipski (1985, 1999) for seminal proposals concerning the potential existing parallels between diachronic and HL developments.

modern German, which only licenses *nominal gerunds* (Van Pottelberge 2004, 2007). We demonstrate in this section that PD has developed a verbal gerund distinguishing itself from other German vernaculars.

The origins of the gerund in English find its roots in deverbal nouns ending in *-ing/-ung* (Fischer et al. 2017: Section 8.2.2). Jack's (1988) comprehensive research on the emergence of verbal gerunds from nominal forms homes in on four primary reasons as to how this development may have occurred. Two are particularly relevant for our treatment of gerunds in PD undertaken here. First, the production of the *-ing*-derivation increased substantially during the Old English and early Middle English periods. Second, unlike infinitives, verbal gerunds could follow prepositions. If this second hypothesis was a determining factor of the development of verbal gerunds in English, it explains why they continued on in further developmental stages of the language through the Middle and Modern periods. As pointed out by Brown and Putnam (2015), whereas (colloquial varieties of) German does not license an aspectual marker *an* immediately preceding a PP, this is acceptable in PD, see, e.g., (35).

(35) a. *Die Kinder sind (*an) mit einem Ball (*an)
the kids are PROG with a ball PROG
Spielen.
play
Intended: 'The kids are playing with the ball.' [Coll. German]

b. Die Kinner sin (a)n mit 'em Ball spiele.
the kids are PROG with a ball play.NF
'The kids are playing with the ball.' [Penn Dutch]

The development of progressive aspect in English must also be mentioned at this juncture in tandem with the emergence of the verbal gerund. The example in (36), taken from Higginbotham (2009: 154), illustrates that progressive aspect in English initially came into existence as a nominal construction with a gerundive object.

(36) John is at [PRO crossing the street].

The prevalence of progressive aspect in spoken colloquial American English and the vestige of the progressive aspect morpheme *an* presented a fertile

environment for the growth and expansion of purely nominal gerunds to also develop a verbal gerund correlate in Penn Dutch. Alexiadou et al. (2010) provide the following decomposition of nominal and verbal gerunds in (37):[13]

(37) a. **Nominal gerunds**: DP ⟩ (NumP ⟩ ClassP ⟩ *n*P) ⟩ (AspP) ⟩ VoiceP
b. **Verbal gerunds**: DP ⟩ AspP ⟩ VoiceP ⟩ *v*P ⟩ Root

Throughout the development of PD, we observe an expansion of the different types of complements – both regard to the shape and size of these syntactic objects – that can co-occur with progressive aspect, such as (38) (example from Huffines 1986), in which an accusative object *Gleeder* 'clothes' and a directional-PP *ins Klasset* 'into the closet' appear in the verbal gerund:

(38) Er is an [$_{XP}$ Gleeder ins Klasset henke.
he is PROG clothes into-the closet hang.NF
'He is hanging clothes in the closet.'

Although a detailed treatment of the historical development of the verbal gerund in Pennsylvania Dutch is beyond the scope of this chapter, we can still establish some clear parallels between its emergence here and what has been observed in the history of related English forms. An anonymous reviewer raised the question as to whether or not some notion of transfer may be responsible for the emergence of *Defective CGs* in contemporary PD. As we discuss in the subsequent section, some notion of transfer, especially as it is conceptualized by Aboh (2015), is present in our reprisal of *Representational Economy* that we introduce below.

6.5 Representations matter

Our analysis of the structural projections associated with various defective domains in PD reveals the importance of mental representations in predicting which forms may and may not emerge due to sustained language contact in an HL setting. The avoidance of projecting TPs in PD in connection with

[13] The proposed underlying structures in (37) are both extended DP-projections. The connection between these complex DP-nominalizations and the projection of TPs and VP-projections are important topics for future research.

non-finiteness provides a straightforward explanation as to why PD has developed *defective clausal gerunds* while (thus far) it has not included *clausal gerunds* in its syntax. I demonstrate that this state of affairs is not unique to PD, but is attested in other heritage varieties of extra-territorial German(ic) in Section 6.5.1. Bearing this evidence in mind, I then comment in Section 6.5.2 on how these data impact our understanding of *Representational Economy*.

6.5.1 The status of (Spec,)TP across (Heritage) Germanic syntax

The status of the TP-projection from a pan-Germanic perspective – both synchronically and diachronically – is a key fulcrum of much parametric variation among Germanic languages.[14] A key factor that distinguishes most Germanic languages (sans English) is a general ban on the projection of Spec,TP as a final landing site for subjects. This representational difference accounts for a number of syntactic properties that are found in Germanic languages; for example, it explains why a language like German does not license ECM and raising predicates while English does. PD, which continues to show strong German(ic) syntactic tendencies, also appears to follow suit here and not project Spec,TP has a licit landing site for subjects. I show below that this is not an isolated phenomena in Heritage Germanic, since North American Norwegian (NAmNo) (Section 6.5.1.1) and Brazilian Pomeranian (Section 6.5.1.2) exhibit similar strategies to avoid projecting Spec,TP.

6.5.1.1 North American Norwegian (NAmNo)

North American Norwegian (NAmNo) is the language of Norwegian immigrants spoken predominantly in the American upper midwest since the 1850s. NAmNo is now moribund, with the eldest generation representing the remaining speakers who possess significant proficiency in the language. Taking a closer look at non-finite clauses in this HL, Putnam and Søfteland (2022) noted the following trends (data taken from the Corpus of American Nordic Speech (CANS); Johannessen 2015a). The examples in (39) are representative samples of defective domains found in NAmNo. Examples (39a) and (39b) follow the "expected" pattern when we compare NAmNo with Norwegian Bokmål;

[14] See Biberauer (2003) for arguments that historically Germanic languages did not originally license *pro* in Spec,TP in the absence of overt subjects.

however, examples (39c) and (39d) are not found in standard Norwegian or any of its dialectal variants.

(39) a. je har inngentinng **å** jøra hær i (Bokmål: **å**)
I have nothing INF do here in
biin
town.DEF
'I have nothing **to** do here in town.' (Outlook-06gm)

b. i e fø gammall **ti** **å** reise (Bokmål: **til å**)
I am too old PRP INF travel
'I'm too old **to** go.' (Hatton-02gm)

c. je har leRRt **te** **å** lœsa no (Bokmål: only **å**)
I have learnt PRP INF read some
nåRsk
Norwegian
'I have learned **to** read some Norwegian.' (Fargo-02gm)

d. nårr je var stor nåkk **å** tjøre (Bokmål: **til å**)
when I was big enough INF drive
bannd'lvægen
bundle-wagon
'When I was big enough **to** drive bundle wagon.' (Decorah-02gm)

Although examples such as (39c) and (39d) do not represent the dominant pattern in CANS, they occur as an alternative whenever the "expected" projections do not occur. Putnam and Søfteland (2022) analyze (39c) as projecting a CP-layer and (39d) consisting of only a *v*P. Identical to the treatment of PD *defective clausal gerunds* advanced here, NAmNo speakers tacitly invoke a strategy to avoid projecting Spec,TP. This strategy, often referred to as *overextension* or *hypercorrection*, during which HL speakers "hypercorrect" their HL by producing forms that disassimilate representations from traits found in the sociolinguistically dominant L2, is commonly attested in HL grammars (Kupisch 2014; Putnam and Hoffman 2021; Rinke and Flores 2014; Rinke, Flores, and Barbosa 2018).

6.5.1.2 Brazilian Pomeranian

Next, we turn to Brazilian Pomeranian (BP), a heritage variety of Low German spoken in Brazil. Postma (2014, 2018) documents a parallel shift in BP that

resembles the "loss" of the infinitival marker *zu* in the historical development of Pennsylvania Dutch (see, e.g., Section 6.3.1 above). In (40), the infinitival marker now occupies the complementizer position in C, however, unlike the situation in Pennsylvania Dutch, in BP lexical element *taum* in C is a morphological portmanteau consisting of T+C (=*tau*+*um*). Postma 2014: 636 reports that this strategy is unattested in European Pomeranian, and that the elimination of *tau* 'to' while only producing the complementizer *um* is extremely rare.

(40) Wij arbeira upm laand [$_{CP}$ **taum** da arme lüür **ø**
we work on.the land INF the poor people
helpen.
help
'We work on the land to help the poor people.' [Brazilian Pomeranian]
(Postma 2018: 132)

In summary, although BP and PD have carried out different operations that have led to a "loss" of their infinitival markers in defective domains, both changes appear to be motivated by similar syntactic factors, leading to instances of clausal-*stretching*.

6.5.2 *Representational Economy* reprised

What emerges from this discussion of the creation of *defective clausal gerunds* in contemporary PD-syntax is the need for a reappraisal of the conceptualization of *Representational Economy* (Polinsky and Scontras 2020; Scontras et al. 2018) and its role in the development of HL syntax. Based on an abundance of experimental evidence that confirms that bi- and multilingual grammars exist in an integrated modular network (Kroll and Gollan 2014; López 2020; Putnam, Carlson and Reitter 2018), a logical outcome of this architecture is that the basic operations of syntactic computation for both/all source grammars are one and the same, i.e. we're dealing with a "shared syntax" (Branigan and Pickering 2017). All things being equal, the call for a reduction of computational domains to preserve processing resources, i.e. the "shrinking of syntax," is a confirmed axiom of *Representational Economy*; however, situations such as the one reviewed here necessitate a minor reformulation of our current understanding of this concept. I offer a revised definition of *Representational Economy* in (41):

(41) **Representational Economy** (*revised*): In HL syntax, reduce syntactic structure (i.e., computational domains) in order to achieve maximal processing efficiency, *or avoid generating projections that result in significant representational conflict between the two source grammars in the dyad.*

This adjusted definition of *Representational Economy* in (41) places instances of clause-*stretching* on par with clause-*shrinking* to the extent that both strategies are tacitly undertaken as an attempt to free up cognitive resources in order to achieve processing efficiency. Our nuanced view on *Representational Economy* rests on two key factors: first, this shift highlights the importance of the syntactic structures that exist, overlap, and conflict in a particular dyad (see Scontras and Putnam 2020 for a discussion of the importance of dyads in HL research). Second, the only way that we can establish a metric of overlap vs. conflict is through the lens of a consistent and falsifiable formal system of representations. Simply put, generative approaches to syntax establish key landmarks of comparison between languages that otherwise may appear circumstantial or irrelevant. The general trend to avoid projecting Spec,TP in PD explains a number of observations about the syntax of defective domains in this language and provides a straightforward answer as to why *defective clausal gerunds* have been adopted, while *clausal gerunds* have not. This revised version of *Representational Economy* also explains the limits of transfer that have thus far taken place in PD, i.e. why *defective clausal gerunds* have made their way into the grammar, while *clausal gerunds* have not. Adopting the proposal that in *defective clausal gerunds* (also referred to as *verbal gerunds* in the literature) the *-ing* exponent in English which appears on D in a DP-structure contributes to defining the outer aspect properties of these events (Alexiadou 2005, 2013; Alexiadou et al. 2010), the transfer of these structures from English to PD prevent a substantial overhaul of the syntax of the latter (i.e., Spec, TP needn't be a final landing site for nouns that receive nominative Case). This assumption is further supported by the lack of ECM- and raising predicates in PD. What remains a question for future research, however, is the relationship between the exponency for non-finiteness that marks both infinitives and gerunds, *-e*, and the overt functional morpheme that marks progressive aspect, *an*.

6.6 Conclusion

In conclusion, the development of *defective clausal gerunds* in contemporary Pennsylvania Dutch is an adaptation that does not require a massive overhaul

of other core properties of its syntax. This returns us to the central research question (RQ) posed in the introduction directed toward understanding instances of clausal-stretching (vs. shrinking). The observed tendency for *Representational Economy* in HL syntax (Polinsky 2018; Polinsky and Scontras 2020; Scontras et al. 2018) should not only be interpreted as a "shrinking" of computational domains (Putnam 2020b), but also as a constraint that avoids the creation of representations that would upend and clash with the syntax of one of the source languages in a dyad. Avoiding projecting Spec,TP – or, at the very least, the same sort of T-head that occurs in English – is essential to maintaining the other German(ic) characteristics observed in PD syntax (e.g., C/TP-expletives, scrambling, (in)coherent infinitives, sluicing, asymmetric V2, etc.). The "shrinking" and "stretching" of computational domains achieve the desired goal of envisaging *Representational Economy* as a means to ease processing efficiency.

As we have seen, barring a few minor exceptions, the syntax of defective domains has remained relatively sturdy and German(ic)-like in PD. However, to be clear, the projection of Spec,TP certainly *could* take place in PD, however, we should expect this change to be gradual and incremental if/when it takes place. A handful of the youngest L1-PD speakers consulted for acceptability judgments found examples such as (29c), repeated here as (42) for the sake of the reader, to be marginally acceptable.

(42) Ich meind(e) noch vun [$_{XP}$?ihn/*Sally mit uns zu die
I remember still P him/Sally mit us to the
Gmee geh.
church go.NF
'I remember him/Sally going to church with us.'

If structures such as (42) with an accusative pronoun become more common and represent a growing trend, it would be the first step toward the licensing of clausal gerunds in PD-syntax, and perhaps the gradual acceptance of ECM-verb-like structures.

The most important point I wish to emphasize at the conclusion of this chapter is the primacy of mental representations in research on HL syntax. D'Alessandro et al. (2021) make the case for how a consistent and falsifiable theory of linguistic representations is a critical component that will guide this research program moving forward. The rationale provided here for how clausal-*stretching* can be subsumed into a (slightly) revised definition of *Representational Economy* only makes sense when one assumes

the interconnectivity of multiple phenomena that are tied to an individual syntactic projection, in this case, TP. In sum, generative approaches to syntax still have much to offer for research on the intricacies and nuances of HL syntactic development.

Acknowledgments

Special thanks are in order to the participants of HLS2 for thoughtful questions on an earlier version of this presentation and discussion of these data. The comments and questions of two anonymous reviewers also provided thoughtful questions and guidance. I also owe a debt of gratitude to the following individuals for discussions and questions involving my analysis: Roberta D'Alessandro, Theresa Biberauer, Hanna Fischer, Vera Lee-Schoenfeld, Terje Lohndal, Masha Polinsky, Kari Kinn, Elena Soare, and Silvia Terenghi. Lastly, a special thanks to Rose Fisher, Mark Louden, Alan Raber, and all of my other PD-speaking colleagues, friends, and informants. The usual disclaimers apply.

PART IV
THE DP

7

Parallel changes in pronominal clitic systems

A view from heritage Romance and Slavic

Alberto Frasson

7.1 Introduction

This study discusses changes in the use of pronominal clitics in heritage Venetan and Bulgarian, adopting a perspective in which native competence in heritage varieties of a language may diverge from native competence in the homeland variety of the same language (Kupisch and Rothman 2018; Pires and Rothman 2007). Therefore, the two heritage varieties will be treated as autonomous and complete grammars that follow their own grammatical rules and constraints in the distribution of pronominal clitics, which are the focus of the present study. The discussion carried out in this study also builds on recent works in the field of heritage language syntax that deal with innovation in heritage languages (D'Alessandro and Frasson 2023; Frasson 2022a,b); with respect to the role of the dominant languages (Brazilian Portuguese for heritage Venetan speakers; English for heritage Bulgarian speakers), it will be shown that the type of change identified in Venetan and Bulgarian clitic paradigms does not depend on cross-linguistic influence or transfer.

The final goal of the present chapter is to identify a common pattern of language change for clitics in heritage languages. In particular, it will be shown that heritage Venetan displays a distribution of subject clitics that is not captured by previous analyses of homeland varieties of the language; similarly, heritage Bulgarian (as described in Ivanova-Sullivan 2019) displays a behavior for object clitics that contrasts with previous analyses of homeland Bulgarian. Despite their different grammatical functions and placement restrictions, previous analyses of Venetan subject clitics and Bulgarian object clitics predict that they will behave as agreement markers, rather than pronominal arguments. However, such clitics in heritage varieties of the two languages do not

Alberto Frasson, *Parallel changes in pronominal clitic systems*. In: *Heritage Languages and Syntactic Theory*.
Edited by: Roberta D'Alessandro et al., Oxford University Press. © Alberto Frasson (2025).
DOI: 10.1093/9780191987731.003.0008

display clear clitic properties, as their distribution resembles more closely that of phrasal pronouns. This behavior challenges previous models of structural deficiency of functional words such as the one presented in Cardinaletti and Starke (1999), according to which clitics lack parallel sets of syntactic, morphological, prosodic, phonological, and semantic properties and are therefore analyzed as reduced elements with limited internal structure.

Section 7.2 presents the aspects of Romance and Slavic clitics that will be relevant in the discussion of heritage languages: I will first discuss previous approaches to the structure and role of clitics, moving to placement properties of clitics. Section 7.3 discusses data from heritage Venetan and Bulgarian, showing a change in the complex set of constraints ruling the distribution of clitics in non-heritage varieties of the languages. In Section 7.4, this change will be attributed to the argumental or agreement-like nature of clitics, in that only the latter allows for cliticization in heritage languages; I also reflect on the reasons behind the change, providing a possible explanation consistent with previous analyses of heritage language change. Section 7.5 concludes the chapter.

7.2 Argumental and agreement-like clitics in Romance and Slavic

7.2.1 Cross-linguistic parallels among different types of clitics

Pronominal clitics in Romance and Slavic belong to the class of special clitics (Zwicky 1977), in that they are reduced counterparts of tonic pronouns and do not occur in the same position and contexts as their corresponding non-clitic forms. At the phonological and prosodic level, clitics are known to lack lexical stress and require a stressed host word to attach to.

The relevant property of clitics addressed in this study regards their structure and ambiguous behavior as arguments or agreement-like morphemes. Chomsky (1994) defines clitics as elements that are at the same time phrases and heads and therefore display an ambiguous XP/X^0 status.[1]

[1] The application of Chomsky's definition to clitics is discussed in Bošković (2002), where clitics are defined as "non-branching" elements. This structural representation, which is assumed to hold cross-linguistically, builds on Chomsky's proposal that a phrase structure should allow for the existence of elements that are at the same time phrases and heads. The idea is formally captured for clitics by assuming that they do not head XPs, but are rather realized in a Spec,XP, while the head X of XP is null. This way, it is X rather than the clitic that takes a complement, so the clitic remains an ambiguous non-branching XP/X^0 element. Notice also that Bošković's analysis is not compatible with some of

Following Uriagereka (1995) and Bleam (1999), I assume that cliticization is not a unitary process obeying a single principle. Crucially, different clitics may have different structural properties that are particularly relevant in the definition of doubling contexts for different types of clitics: a well-known property of some pronominal clitics is their ability to double a strong pronoun, while other clitics generally disallow doubling. Uriagereka (1995) correlates this behavior with the different structures of clitics, which are divided into "strong" and "weak" ones: while strong clitics are XP-like elements and cannot double another XP argument, weak clitics are X-like elements and can double an XP argument. This idea was further developed by Bleam (1999), who proposes a more fine-grained clitic classification, exemplified by different doubling contexts with dative and accusative third-person clitics: Romance accusative clitics, unlike dative clitics, trigger specificity effects that constrain the contexts in which they can double a DP. Bleam generalizes that some clitics can be analyzed as determiners, as they realize a D-layer in their structure, where specificity is encoded;[2] conversely, other clitics do not encode specificity and are therefore better analyzed as agreement markers.

I maintain that clitics may belong to different classes in view of the type of features they encode; I show that the presence of certain features in a clitic reflects on their structure and their ability to double arguments or, conversely, behave as arguments themselves. The crucial point in this respect is that clitics, despite being often analyzed as representing the same class of elements, may correspond to quite different syntactic representations. In the remainder of this section, I will briefly show how different syntactic representations affect doubling contexts in some Romance and Slavic languages. Among Slavic languages, Slovenian accusative pronominal clitics, on the one hand, have an argumental function, in that they cannot co-occur with their strong counterparts (1); Macedonian object pronominal clitics, on the other hand, double strong pronouns and are therefore more properly defined as agreement markers (2). The same distinction emerges in Romance, if we consider object clitics in Italian, where clitics appear in complementary distribution with strong object pronouns (3), and Spanish, where clitics double strong object pronouns (4).

the examples by Cardinaletti and Starke (1999) discussed in Section 7.2.2, supporting the idea that different clitics should receive separate dedicated analyses.

[2] A reviewer pointed out that this may be true only for some languages, like Spanish varieties discussed in Bleam (1999) or Romanian. Other languages, like Italian, freely allow clitics to resume non-specific XPs.

(1) Slovenian
Ona me je pogledala (*mene).
she me.ACC-CL is. CL look.PRT me.ACC-PRON
'She looked at me.'

(2) Macedonian
Taa me pogledna mene.
she me.CL look.PST me.PRON
'She looked at me.'

(3) Italian
Lei mi ha guardato (*a me).
she me.CL has look.PRT to me.PRON
'She looked at me.'

(4) Spanish
Ella me miró a mí.
she me.CL look.PST to me.PRON
'She looked at me.'

Examples (1)–(4) reveal a quite striking difference in the distribution of clitics: on the one hand, Slovenian and Italian object clitics act as arguments of the verb;[3] on the other hand, Macedonian and Spanish clitics double full pronominal arguments of the verb.[4] Despite this fact, all clitics in examples (1)–(4) are often taken to represent the same class of elements.

A relevant structural analysis of clitics is presented in the seminal work by Cardinaletti and Starke (1999); they introduce a tripartite model of pronominal structural deficiency, in which Slavic and Romance clitics are analyzed as the same type of element. This model assumes a strong correspondence of deficiency at all levels of linguistic analysis: the two classes of deficient pronouns (weak and clitic) are defined in view of the amount of syntactic, morphological, phonological, and semantic properties they lack with respect to the class of strong pronouns. The structure of the three classes of pronouns in Cardinaletti and Starke's model is summarized in (5).

[3] Consider that, especially in Italian, object drop is not a possibility in the context of (3).

[4] A reviewer pointed out that, according to Kayne's generalization (in Jaeggli 1982), objects are doubled by a clitic only if they are preceded by a preposition. A discussion of this generalization is beyond the goals of the present study; however, notice that (Balkan) Slavic languages are known to violate Kayne's generalization, in that the constraint on the co-occurrence of clitic doubling and prepositional objects does not hold (Kallulli 2019).

(5) (Adapted from Cardinaletti and Starke 1999: 86–87)
(a) strong pronouns: [CPL [ΣPL [IPL [LP]]]]
(b) weak pronouns: [ΣPL [IPL [LP]]]
(c) clitic pronouns: [IPL [LP]]

Strong pronouns (5a) are compared to verbal CPs, they encode a nominal CPL layer (with L identifying any lexical category) associated with specific referential and interpretive properties and therefore represent phrases. Weak pronouns lack such layer, and encode only ΣPL, associated by the authors with prosody and lexical stress. Clitics represent proper subsets of the syntactic structure of strong (and weak) pronouns, and they represent pure agreement markers (inflection-like elements in Cardinaletti and Starke's definition); this property is reflected in the lack of one or more morphemes, of stress and of certain interpretive properties. An alternative way of understanding this proposal is to assume that each morpheme corresponds to a head in syntax, hence strong pronouns realize more heads in view of their more complex morphological make-up. At the same time, clitics, as morphologically and syntactically deficient elements, lack the morpheme corresponding to the highest functional projection in syntax, the one associated with the semantic notions of reference and interpretation.

Even though Cardinaletti and Starke's model elegantly captures the correspondence of clitic behavior at various levels of linguistic analysis, the data discussed in this study present a potential challenge for their approach, regarding the interpretation of clitics in Slovenian and Italian, on the one hand, and Macedonian and Spanish, on the other (examples (1)–(4)). Macedonian and Spanish *me* in examples (2) and (4) are correctly predicted to lack semantic information about the referent: they are simple agreement markers and they double the phi-features of the strong pronoun, which represents the real argument and, as such, carries the referential information. Slovenian *me* (1) and Italian *mi* (3), however, do not double a strong pronoun; that is to say, they represent arguments of the verb and have to encode the information relative to the referent. Since they are assumed to lack the portion of the structure encoding semantic information about interpretation, it is not clear where such information is realized and how.

This section aimed at showing a potential problem with our current understanding of classes of pronouns, providing evidence that clitics may represent different types of deficiency, even though they superficially allow for parallel analyses: the strict correspondence between morphological, syntactic, and semantic deficiency proposed by Cardinaletti and Starke (1999) cannot be a

general property of all clitics. Before discussing how heritage languages can shed light on this matter, I will briefly discuss different placement properties of clitics in Bulgarian and Venetan.

7.2.2 Clitic placement in Bulgarian and Venetan

Aside from their status as arguments or agreement markers, the literature focused on the fact that clitics are generally placed in a position that does not correspond to their base thematic position inside the sentence.

Pronominal clitics in Bulgarian obey the Tobler–Mussafia law, a constraint that disallows clitics in the first position because of their phonologically weak nature. Bulgarian clitics have a high degree of host selectivity, in that they need to attach to a verb. Pronominal clitics are proclitic to the verb (6a) and do not need to appear in the second position (Bošković 2001; Halpern 1995); the only restriction on their placement is represented by the impossibility of appearing in the first clausal position (6b).

(6) Bulgarian (Bošković 2001: 5)

a. Petko mi go dade včera.
Petko me.DAT-CL it.ACC-CL give.PST yesterday

b. *Mi go dade Petko včera.
me.DAT-CL it.ACC-CL give.PST Petko yesterday
'Petko gave it to me yesterday.'

Franks and Rudin (2005) analyze Bulgarian clitics as agreement markers. The main evidence in this respect comes from the fact that Bulgarian object clitics are involved in doubling constructions: rather than appearing in complementary distribution with strong pronouns and lexical objects, a clitic can co-occur with an associate nominal, which represents the real argument of the verb.

A similar analysis was proposed for subject clitics in different Romance languages. Starting from Brandi and Cordin (1981) and Rizzi (1986a), it was shown that subject clitics in northern Italo-Romance varieties, such as Venetan, have peculiar placement properties that distinguish them from clitic subject pronouns of the French type: these properties regard the placement of subject clitics with respect to preverbal negation (7) and their placement in coordinated structures (8).

(7) a. Ils ne savaient pas parler de certaines choises.
they.CL NEG know.PST NEG speak.INF of certain things

b. No i savea parlar de serte robe.
NEG they.CL know.PST speak.INF of certain things
'They did not know how to speak about certain things.'

(8) a. Elle chante et danse.
she.CL sing.3SG and dance.3SG

b. La canta e la bala.
she.CL sing.3SG and she.CL dance.3SG
'She is singing and dancing.'

The French subject pronoun in (7a) is realized before preverbal negation *ne*, while the Venetan subject clitic *i* (7b) is realized after preverbal negation *no*. In the coordination context in (8), the French subject pronoun *elle* is realized in the first coordinated conjunct, while the Venetan subject clitic *la* is repeated in both coordinated conjuncts. On the basis of such distributional facts, Venetan subject clitics have been analyzed as the reflex of quite different structural properties with respect to their French counterparts. The analysis presented in Brandi and Cordin (1981) and Rizzi (1986a), as well as subsequent studies (see, in particular, Benincà 1994; Poletto 2000; Roberts 2010) defines subject clitics of the Venetan type as heads that cliticize on the verb, rather than phrasal pronouns. Conversely, Rizzi's study shows that French subject clitics are real phrasal arguments of the verb and their clitic nature is only a phonological property. Subject clitics of the Venetan type are rather part of the inflection, on par with agreement morphology on the verb, and are therefore agreement markers doubling any type of subject (overt or null).

In conclusion, both Bulgarian object clitics and Venetan subject clitics have been analyzed as agreement-like heads, rather than argument clitics (see Franks and Rudin 2005 on Bulgarian; and Poletto 2000, Rizzi 1986a, Roberts 2010 on Venetan[5]) because of their peculiar distribution and their participation in doubling constructions. In the remainder of the chapter, I will show how these properties are affected in heritage varieties of Venetan and Bulgarian.

[5] The participation in doubling phenomena shows that an analysis of Bulgarian object clitics and Venetan subject clitics has to be preferred. See however Bošković (2001) on Bulgarian and Poletto (1993) on Venetan for alternative analyses of the same types of clitics as argumental heads.

7.3 Clitics in bilingual and heritage speakers

In this section, I will show how structural and distributional properties of Venetan and Bulgarian clitics are affected in heritage varieties of these languages. Even though the literature on language change has touched upon the realization of clitics in bilingual and heritage speakers, the number of studies dealing with their structural properties is still scarce. Besides, the majority of the studies on clitics mainly focused on cross-linguistic influence and processing costs at the interface level, whereas the question of how the structure and syntactic distribution of clitics are affected in bilingual speakers of heritage languages has somehow been eluded so far.

A study by Pérez-Leroux et al. (2011), for instance, shows that bilingual Spanish-English children tend to avoid proclitic (preverbal) objects and extend the use of enclitic (postverbal) ones in Spanish; they attribute this change to a priming effect, in that English does not allow for preverbal objects, which in turn leads to a shift in lexical items. The authors further propose that lexical activation is not language-selective, so both the Spanish and the English lexicons remain active, allowing for cross-linguistic priming from the dominant language (English). Working in a similar framework, Rinke and Flores (2014) analyzed the acquisition of clitics in heritage speakers of European Portuguese and found that, even though there is a tendency to accept ungrammatical clitic placement, there is no overgeneralization of enclisis to contexts that require proclisis and that priming cannot explain the differences.

In fact, the data in this section assume that native competence in heritage varieties of a language may diverge from native competence in the homeland variety of the same language (Kupisch and Rothman 2018). Therefore, the varieties spoken by heritage bilingual speakers of Venetan and Bulgarian are not treated as incomplete systems that develop under the constant influence of a dominant language and are not compared to their homeland (monolingual) counterparts;[6] the two varieties are treated as autonomous and complete grammars that follow their own grammatical rules and constraints, which are the focus of the present study. This approach inspired recent works in the field of heritage language syntax that deal with innovation in heritage languages (D'Alessandro and Frasson 2022; Frasson 2022a, 2022b); building on

[6] In this respect, I follow Rothman et al. (2022), who show that monolingual varieties of a language cannot be taken for granted as an appropriate comparison. The approach to language change I adopt here is neutral with respect to the relationship between homeland and heritage varieties. I compare heritage and homeland clitic systems, and analyze the change in linguistic terms, without resorting in any way to "monolingual normativity."

such works, the present study will test the hypothesis that certain changes in the distribution of pronominal clitics in heritage languages do not depend on cross-linguistic influence, but represent a common path of change in heritage languages, regardless of their contact language.

In particular, the data discussed in the study show that pronominal clitics in heritage Venetan and Bulgarian do not exhibit clear agreement-like properties, but they rather behave as arguments.

7.3.1 Subject clitics in heritage Venetan

The data presented in this section are of heritage Venetan spoken in Brazil. This variety is a northern Italo-Romance language, currently spoken in Brazil by almost half a million people, as a result of the migration waves from northern Italy at the end of the 19th century. All the speakers of Venetan in Brazil are bilingual with Brazilian Portuguese as a dominant language. The data discussed in this section, originally presented in Frasson (2022b), show that heritage Venetan subject clitics display a behavior that is not captured by previous analyses. Recall that the analysis of Venetan subject clitics as agreement markers discussed in Section 7.2 builds on examples such as (9), where the subject clitic doubles a strong pronoun.

(9) Venetan

Ti	te	ga	ciamà	to	fradel.
you.PRON	you.CL	have.2SG	call.PRT	your	brother

'You called your brother.'

The subject clitic *te* in (9) cannot be dropped and is truly doubling the strong pronoun *ti*. However, as shown in previous studies (Benincà 1994; Frasson 2022b; Poletto 1993), doubling is obligatory in Venetan only for second-person singular. In this study, I will refer to third-person subject clitics, leaving an analysis of doubling in second-person singular for future research. In the case of the third person, previous studies show that the subject clitic is not obligatory unless the subject XP is left dislocated. This discrepancy in the behavior of second- and third-person subject clitics becomes even more evident in the heritage variety of Venetan spoken in Brazil. In this variety, third-person subject clitics are generally banned in sentences where a strong subject pronoun is present, as shown in (10). Besides, Frasson (2022b) shows that subject clitics can be dropped even if no other overt subjects are present,

giving rise to a canonical null subject configuration (11) and that they are in complementary distribution with strong pronouns (12):

(10) Lore (*le) ga ciamà so fradel.
they.PRON they.CL have.3PL called.PRT their brother
'They called their brother.'

(11) (La) ze vegnesta sola.
she.CL is come.PRT alone
'She came alone.'

(12) Ela / La ga ciamà me pare.
she.PRON she.CL have.3SG call.PRT my father
'She called my father.'

If subject clitics are in complementary distribution with strong pronouns and lexical subjects, as in (12), and there is no doubling function (10) nor obligatory realization of subject clitics (11), then their analysis as agreement-like clitics is weakened.

Besides, there is another property displayed by heritage Venetan subject clitics that makes their analysis as agreement markers impossible: third-person subject clitics can be separated from the verbal host by non-clitic material, like adverbs (13).

(13) El pena ze rivà.
it.CL just is arrive.PRT
'It has just arrived.'

A syntactic analysis of Venetan subject clitics as agreement-like heads cannot be maintained for elements that behave as in (10)–(13). Despite being phonologically reduced, the subject clitics in (10)–(13) behave as regular subject pronouns, rather than obligatory agreement markers: they appear in complementary distribution with other types of subjects, they can be dropped given the appropriate discourse conditions, and adverbs can be interpolated between them and the verb.

The idea that at least some Venetan subject clitics can have an argumental function is not absent in literature on the language (see Poletto 1993). However, the heritage data provide final evidence that they not only are argumental,

but they are most likely not distinguished from other subject pronouns, as proposed in Frasson (2022b).

Finally, it should be noted that it is not possible to establish a correlation between the distribution of subject clitics in heritage Venetan and the contact with Brazilian Portuguese, the dominant language of the speakers. As shown in Holmberg (2005), Brazilian Portuguese is a partial null subject language, which does not allow for referential third-person null subjects. Conversely, heritage Venetan in Brazil is a canonical null subject language, it allows for null referential subjects and allows for subject clitic drop (11), making the hypothesis of a direct effect of contact between the two varieties not plausible.

7.3.1.1 A note on heritage Venetan object clitics

Before moving to the discussion of Bulgarian object clitics, one may wonder whether the same type of behavior is attested in heritage Venetan object (accusative) and indirect object (dative) clitics.[7] Data from Brazilian Venetan show that this is not the case: it is impossible to draw a parallel between the behavior of subject and object/indirect object clitics in heritage Venetan. Evidence in this respect comes again from the obligatory realization of a doubling clitic in a dedicated preverbal position (14) and the impossibility of adverb interpolation (15).

(14) *(Te) vardo sempre.
you.CL watch.1SG always
'I am always watching at you.'

(15) Ghe (*pena) go (pena) dato el libro a ela.
to.her.CL just have.1SG just give.PRT the book to her.PRON
'I have just given her the book.'

The object clitic *te* (14) and the indirect object clitic *ghe* (15) are found in a dedicated position; besides, the clitic in (15) doubles a strong pronoun that is realized postverbally, in its canonical argumental position. Finally, the clitic in (15) cannot be separated from the verb by non-clitic material, showing that it

[7] For reasons of space, Venetan object clitics were not discussed in Section 7.2, where I decided to focus on elements that displayed different behaviors in heritage varieties with respect to previous descriptions of the same language. Object clitics do not differ significantly in heritage Venetan and the literature on Romance object clitics is vast; it would not have been possible to include all the different approaches to the phenomenon in this study. I refer to the recent work by Pescarini (2021) for an overview of previous literature and a novel approach to the study of Romance object clitics.

is indeed an agreement marker realized on a verbal head via head movement (see Roberts 2010 in this respect).

In conclusion, only subject clitics in heritage Venetan allow for an analysis as argumental clitics, which further allow for a peculiar strong pronominal use; heritage Venetan object clitics, conversely, are better analyzed as agreement-like clitics, supporting the need for a more fine-grained classification of clitics.

7.3.1.2 Object clitics in heritage Bulgarian

An interesting behavior, parallel to the one identified in Frasson (2022b) for heritage Venetan subject clitics, is attested for object clitics in heritage Bulgarian. Changes in the distribution of object clitics were noticed by Ivanova-Sullivan (2019) in heritage Bulgarian spoken in the United States.[8] For the sake of clarity, it should be noted that the Venetan data discussed in Section 7.3.1 refer to bilingual adults, while the Bulgarian data discussed in the present section refer to bilingual children. Despite the difference in the speakers' profiles, it will be shown that the two cases show a striking similarity.

As shown in Section 7.2, pronominal clitics in Bulgarian are adverbal and obey the Tobler–Mussafia law: they are generally proclitic to the verb, but they can be enclitic in contexts in which they would otherwise end up in the first position in the sentence. The data presented in Ivanova-Sullivan (2019), however, show that this is not the case for heritage Bulgarian. Heritage speakers produce mainly enclitic objects, even if the verb is not the first constituent in the sentence:

(16) Heritage Bulgarian (adapted from Ivanova-Sullivan 2019)

Toj	vze	go.
he	took	it.CL

'He took it.'

What is relevant in (16) is that the clitic does not climb: the generalization of a postverbal position for object clitics, means that they are realized in a canonical object position where strong object pronouns would be otherwise realized:

[8] In this chapter, I report the data in Ivanova-Sullivan (2019) because of their relevance for the analysis of structural and distributional properties of clitics in heritage languages and for the study of heritage grammars in general. I refer to her original work for all implications related to the acquisition of clitics in heritage speakers.

(17) Toj vze nego.
he took it.PRON
'He took it.'

The strong object pronoun *nego* in (17) cannot be distinguished from the clitic *go* in (16) on the basis of their distributional properties.[9]

The realization of both strong and clitic objects in a postverbal position allows to predict that the two types of objects will be in complementary distribution in heritage Bulgarian. This prediction is confirmed in Ivanova-Sullivan's (2019) data: heritage Bulgarian object clitics pronouns can be used interchangeably.

(18) (Adapted from Ivanova-Sullivan 2019)
Ne zakačaj men / me.
not bother me.PRON me.CL
'Do not bother me.'

The complementary distribution of strong *men* and clitic *me*, rather than their co-occurrence, allows for an even stronger parallel with heritage Venetan subject clitics: heritage Bulgarian object clitics do not show the typical behavior of agreement-like clitics, but that of argumental clitics and further allow for a parallel strong pronominal use.

Just as in the case of heritage Venetan, one may hypothesize that the change in the placement of clitics in heritage Bulgarian is correlated with the distribution of object clitics in English, the dominant language of the speakers. In this case too, it should be noticed that object realization and placement in heritage Bulgarian do not converge toward the pattern displayed by English. As shown in (18), heritage Bulgarian retains the original alternation between the two object forms, tonic and clitic, but their distribution is different from what is found in homeland Bulgarian. As in the case of Venetan-Portuguese bilinguals in Brazil, I do not take this difference to signal the influence of English, but rather the development of independent conditions on the distribution of object clitics in heritage Bulgarian.

In conclusion, data presented in this section show that Venetan subject clitics and Bulgarian object clitics in heritage varieties of the languages have quite

[9] Strong pronouns are expected to have different interpretive properties in Bulgarian, as predicted by Cardinaletti and Starke (1999). In absence of more detailed data on the interpretation of strong object and clitic pronouns, I remain agnostic regarding the possibility for the same interpretive properties to be present in heritage Bulgarian as well.

similar distributional properties, in that they exhibit a tendency to appear in positions in which their tonic counterparts would normally appear, rather than in the position normally expected for clitics.

Besides, the complementary distribution of clitics and strong pronouns can be interpreted as a tendency to behave as arguments of the verb, rather than appearing in doubling constructions. In other words, clitics in heritage Venetan and Bulgarian represent argumental, rather than agreement-like clitics.

7.4 The syntactic status of pronominal clitics in heritage languages

The data presented in Section 7.3 showed that the two heritage varieties under analysis do not distinguish clitics from their strong counterparts based on their distributional properties. This applies equally to heritage Venetan subject clitics and to heritage Bulgarian object clitics. At this point, one may wonder whether it is necessary to maintain the analysis of special clitics proposed in previous studies for Venetan and Bulgarian. In the present section, I will show that the analysis of clitics proposed in Cardinaletti and Starke (1999) cannot be maintained for argumental clitics in heritage Venetan and Bulgarian; I will also show that agreement-like clitics too allow for a different analysis, supporting the discussion developed in Section 7.2 on the crucial relevance of the contrast between argumental and agreement-like clitics. I will first consider the case of heritage Venetan subject clitics and heritage Bulgarian object clitics and finally move to the contrast with heritage Venetan object clitics.

In view of the data presented in Section 7.3, I propose that heritage Venetan subject clitics and heritage Bulgarian object clitics have the status of pronouns, rather than clitics. It was shown that they are in complementary distribution with strong pronouns and therefore should be analyzed as arguments of the verb. Besides, their distribution and position inside the sentence make it unnecessary to analyze them as special clitics. Crucially, a clitic analysis along the lines discussed in Section 7.2 should be limited to elements that display clear clitic syntactic and distributional properties. In the case of heritage Venetan and Bulgarian, clitics are morphologically reduced with respect to their strong counterparts, but there is no conclusive evidence that the lack of certain morphological properties systematically corresponds to the lack of a portion of syntactic structure. Recall that, according to the model discussed in Cardinaletti and Starke (1999), pronouns are divided into three different

types: strong, weak, or clitic. Each pronoun type is associated with a syntactic projection. The model, already presented in (5), is repeated here as (19).

(19) (Adapted from Cardinaletti and Starke 1999: 86–87)

(a)	strong pronouns:	[CPL [ΣPL [IPL [LP]]]]
(b)	weak pronouns:	[ΣPL [IPL [LP]]]
(c)	clitic pronouns:	[IPL [LP]]

In view of the distributional properties displayed by heritage Venetan and Bulgarian clitics, I propose that argumental clitics such as the ones presented in Section 7.3 encode at least the ΣPL layer associated with prosodic properties and lexical stress, which also explains their lack of dependence on the verbal host. The structure of heritage Venetan and Bulgarian clitics, compared to that of clitics in the homeland varieties in the two languages, is represented in (20).

(20)

(a)	heritage clitics:	[ΣPL ***clitic*** [IPL [LP]]]
(b)	homeland clitics:	[IPL ***clitic*** [LP]]]

The discussion of the complementary distribution of clitics and strong pronouns in heritage Venetan and Bulgarian could also suggest that clitics do not simply behave like weak pronouns, but like strong ones. I will not delve into this possibility, which is explored in Frasson (2022b), who shows the distribution of the two types of subject pronouns in heritage Venetan is that of strong pronouns (DP-pronouns, in the tripartite model discussed by Déchaine and Wiltschko 2002[10]). The question of whether a parallel analysis could be applied to heritage Bulgarian object clitics is a matter for future research.

Before moving to the question of language change, I will briefly address the problem presented in Section 7.3.1.1: it was shown that heritage Venetan object clitics (unlike subject clitics) display a behavior consistent with the analysis as agreement-like elements discussed in Section 7.2. Besides, the distribution of object clitics in doubling constructions in heritage Venetan is consistent with previous accounts of the phenomenon (Bonet 1991; Kayne 1991; Manzini and Savoia 2000). An interesting proposal about the nature of Romance object clitics is discussed in Pescarini (2021). He argues that the structure of object clitics is as complex as that of strong pronouns

[10] Frasson (2022b) adopts Déchaine and Wiltschko's (2002) model, rather than Cardinaletti and Starke's (1999) structural deficiency. Both proposals distinguish three classes of pronouns, although the way such classes are identified is not exactly the same in the two systems. For the purposes of the present study, I adopt Cardinaletti and Starke's (1999) proposal for the sake of representation.

(DP-pronouns in Déchaine and Wiltschko 2002). The special distributional properties of clitics derive from the fact that an object pronoun is attracted and frozen in a dedicated topical position, which the author identifies with the Wackernagel position. The pronoun cannot receive sentence stress in this position, so it undergoes a process of phonological reduction and may, subsequently, undergo incorporation into the verb. In this approach, incorporation is intended in Matushansky's (2006) terms as a morphological operation resulting in the conflation of two adjacent elements into a single morphophonological unit. Incorporation explains the ban on interpolation in examples (14)–(15), whose structure is represented in (21).

(21) (Adapted from Pescarini 2021: 177)
[C… [I … [Z clitic + V] Adv]]

The Z-position in (21) was originally proposed by Ledgeway and Lombardi (2005) and is analyzed as an evolution of the original Wackernagel position; clitics move to Z as phrasal constituents and are then incorporated into a verb moving through it.

Object clitic placement in heritage Venetan allows for an analysis along the lines of Pescarini (2021), while the same type of approach cannot be maintained for subject clitics as well as heritage Bulgarian object clitics. The distinction between argumental and agreement-like clitics is crucial in this respect. I propose that cliticization as defined in Pescarini (2021) is available in heritage languages for pronouns that participate in doubling constructions. Heritage Venetan subject clitics and heritage Bulgarian object clitics do not have this function and therefore do not climb to a dedicated position or incorporate into the verb.

Heritage data are relevant in two respects. On the one hand, they confirm that clitics do not lack a portion of internal syntactic structure and can therefore behave as free pronouns. In other words, the case of heritage Venetan and Bulgarian supports an analysis in the style of Pescarini (2021) and challenges Cardinaletti and Starke's (1999) model, in that there is no consistent correspondence between phonological, morphological, and syntactic constraints on the distribution of clitics. On the other hand, heritage data support the idea that different clitics belong to different classes, as evidenced by the contrast between the behavior of argumental and agreement-like clitics in heritage languages. Argumental clitics do not display a clitic behavior in heritage Venetan and Bulgarian; the distributional properties of argumental clitics in the two varieties allow to analyze them on par with strong pronouns. However,

heritage Venetan object clitics do not have different distributional properties with respect to previous descriptions based on the homeland variety of the language, as they retain their clitic-like behavior. As a last point, I will address the question of why only argumental clitics behave differently and what makes such a difference in heritage languages so evident.

7.4.1 Clitics in heritage languages: A hypothesis on language change

In the course of this chapter, I have shown that more than one clitic class exists and that it is not possible to include all clitic types in the same category because of the variety of structural and distributional properties they display. There are a number of potential factors that can tell apart different types of clitics; in this contribution, I focused on the argumental vs. agreement-like distinction. This contrast is particularly relevant in the case of heritage Venetan subject clitics and Bulgarian object clitics, which, unlike their homeland counterparts, do not exhibit clear clitic properties. At this point, one may wonder what the heritage perspective can tell us about the fact that agreement-like clitics tend to retain their properties, while argumental clitics do not.

In this section, I build on Polinsky (2018) and Scontras et al. (2018) and explore a hypothesis on the reorganization of heritage grammars.

Polinsky (2018) discusses two possible triggers for heritage language change that are compatible with the case presented in this chapter. The first possible trigger concerns specifically different working-memory costs for the processing of arguments and agreement markers. The second possible trigger regards the syntactic status of clitics as heads and their reanalysis as phrases in heritage languages.

As far as working-memory costs are concerned, Polinsky's (2018) approach predicts argumental clitics to be more resistant to processes of change typical of heritage languages (e.g., attrition; see Montrul 2008) than agreement markers; this difference depends on the fact that the type of relationship established by agreement markers is more costly than the one established by arguments; this approach captures the case of heritage Venetan and Bulgarian, in that argumental clitics are, at least in some cases, preferred to agreement-like ones because they do not need to establish an agreement relationship with an element they double. Agreement markers, on the other hand, need to match features that are located on the constituents they double, which are often not so close to them in the syntactic structure.

As for the status of clitics as heads, Polinsky (2018: 57) discusses cases in which heritage speakers of English exhibit a tendency to reanalyze verbal particles of phrasal verbs (heads) as phrasal adverbs. As a consequence, particles display distributional properties of free adverbs, in that they do not comply with any placement restrictions and allow for variability in word order. A similar process was discussed for Venetan in Frasson (2021) for subject clitics and emerges from the heritage data discussed in the present chapter as well.

The two hypotheses discussed by Polinsky can be considered in a unitary perspective. The starting point is that the doubling of arguments via extra agreement markers realized on the verb presents speakers with extra costs for working memory. According to Polinsky (2018), this may have to do with transparency and compositionality (see also Ramchand and Svenonius 2002 in this respect): as long as newly introduced structures (in the case under analysis, argumental clitics instead of agreement-like clitics) are transparent and compositional, they do not pose any problem for speakers of heritage languages. Compositional constructions, whose parts can be clearly identified and therefore easily separated (such as heritage Venetan subject clitics) or moved more freely (such as heritage Bulgarian object clitics), are morphologically transparent and are preferred to constructions that behave as single units, formed by different inseparable but identifiable parts (such as the clitic-verb complex). This tendency to avoid morphological complexity results in the reanalysis of clitics as phrases: heritage speakers do not substitute clitic forms with strong forms, but they do not distinguish the two forms at a structural level. This reanalysis allows for a more free distribution of clitics with respect to their placement. Subject clitics in heritage Venetan and object clitics in heritage Bulgarian still exhibit some clitic properties at the morphological and phonological level, but they resemble the "strong" clitics described by Uriagereka (1995), the argumental clitics. Such clitics are analyzed as phrasal elements in heritage languages, in that they impose less constraints on their placement than clitics.

Polinsky's approach explains why heritage Venetan subject clitics and heritage Bulgarian object clitics are reanalyzed as free pronouns: the ambiguous XP/X^0 status of clitics is resolved in favor of an XP realization, which is less costly for working memory and more transparent at the morphological level. However, this approach still does not explain why heritage Venetan object clitics do not undergo a parallel change. While the data analyzed in the present chapter do not allow for a conclusive answer in this respect, in the remainder of this section I will draw some preliminary hypotheses. A first possibility is to analyze the pattern displayed by heritage Venetan as an ongoing change; such

process started from subject clitics because of a mechanism that is reminiscent of McCloskey's (1990) Highest Subject Restriction, which disfavors the doubling of immediately subjacent subjects. Therefore, the restriction on subject doubling, as well as the subsequent change attested in heritage Venetan subject clitics, is intended as the first step in a process of change that is not yet completed. Alternatively, it is possible to interpret the contrast between heritage Venetan subject and object clitics as a different strategy to reduce working-memory costs. I have previously proposed, building on Polinsky (2018), that the tendency to reduce processing costs led subject clitics to be reinterpreted as phrasal pronouns. Scontras et al. (2018) propose that the same tendency may have the opposite outcome, in that agreement-like clitics could further reduce in view of a principle of Representational Economy; this approach predicts that features of object clitics in heritage Venetan are not valued independently, but bundle together with other features of the verb and are retrieved in memory as one single unit, in which single sub-units are no longer independently identified. In Scontras et al. (2018), such a process leads to an amalgamation of features and a reinterpretation as a single-valued constituent, finally resulting in a loss of feature specification of the clitic. A conclusive answer on whether either of these hypotheses is plausible for heritage Venetan object clitics cannot be provided based on the data presented in this chapter; I leave this question open for further investigation.

This section showed that some processes of change and reanalysis in heritage grammars are transparent and easily captured, while others are more mysterious and are still poorly understood. The reasons behind the change attested in the systems of heritage Venetan subject clitics and heritage Bulgarian object clitics can be captured by combining syntactic and processing approaches to language change; it is however more difficult and elusive to understand why heritage languages also allow for a complexification of the system of clitics as the one displayed by heritage Venetan object clitics (see also Frasson 2022b in this respect).

7.5 Conclusion

Analyzing clitics has posed a number of challenges for linguistic theories in the past decades. In Section 7.2, I attempted to capture a little portion of the complexity represented by the phenomenon of cliticization in Romance and Slavic languages. If anything, Section 7.3 showed that things become even more complex in heritage languages, where clitics have unpredictable

behaviors, pointing to different change paths even within the same variety; conversely, they may display converging properties in quite distant varieties. The former case is represented by heritage Venetan subject and object clitics; the latter is represented by heritage Venetan subject clitics and heritage Bulgarian object clitics. I proposed that this last case can be analyzed as an instance of change from agreement-like to argumental clitics, triggered by a combination of processing and structural factors. Finally, I tentatively proposed that similar considerations could also capture the different behavior of heritage Venetan object clitics, possibly in the perspective on an ongoing change.

The case presented in this study does not presume to represent an exhaustive discussion of how clitic systems change in heritage varieties. The main goal of this contribution is to shed light on formal properties of clitics, regardless of their level of similarity (as in the case of Venetan and Bulgarian), and identify a predictable change path in heritage languages, regardless of extra-linguistic factors. The account proposed here for linguistic facts in heritage languages moves away from the perspective in which heritage languages are viewed as strictly dependent on their homeland and contact varieties.

In conclusion, heritage languages exhibit intriguing change patterns that can challenge or confirm previously established principles of heritage grammars and provide clear answers to more general questions on the nature of clitics cross-linguistically.

Chapter 8

The DP layer in heritage Norwegian

Vulnerability and nominal architecture

Terje Lohndal and Yvonne van Baal

8.1 Introduction

An important objective in the study of heritage languages is to identify what properties are stable under the conditions of reduced input and language use. A lot of research has shown that some properties indeed are stable, whereas other properties can be quite vulnerable to omission or restructuring (Benmamoun 2021; Flores et al. 2022; Montrul 2008, 2016; Polinsky 2018; Rinke and Flores 2014). Even so, it is clear that heritage speakers have a complete grammar (Polinsky 2018: 25, 350) and that moribund heritage languages maintain a high level of complexity (Bousquette and Putnam 2020).

Generalizing over several heritage languages, Polinsky (2018: 64–65) claims that salient aspects of language are more likely to be maintained in heritage languages, while non-salient elements are not. Tense and determiners are mentioned explicitly as salient and stable categories. In this chapter, we will provide data which demonstrate that D is not necessarily resilient and probe some of the reasons for why this may be the case. We will do so based on data from heritage Norwegian. Specifically, we will study data from North American Norwegian (henceforth, NAmNo), which is a heritage language spoken in the United States by third to fifth generation immigrants. It has a rich history, well documented in the landmark study by Haugen (1953). Recently, a big corpus of spoken NAmNo has been collected at the Text Laboratory at the University of Oslo (see Johannessen 2015a). The present chapter will use data from this corpus as well as elicited speech.

This chapter is organized as follows. Section 8.2 provides the relevant background for the chapter, which includes an outline of the structure of DPs in Norwegian. We also present basic facts about North American Norwegian and some previous research. In Section 8.3, we present the puzzle of the

Terje Lohndal and Yvonne van Baal, *The DP layer in heritage Norwegian*. In: *Heritage Languages and Syntactic Theory*. Edited by: Roberta D'Alessandro, et al., Oxford University Press.
DOI: 10.1093/9780191987731.003.0009

chapter in more detail before providing an analysis of NAmNo in Section 8.4. Section 8.5 presents a comparison with other domains, in particular the T- and C-domains. Concluding remarks appear in Section 8.6.

8.2 Background

In this section, we provide some background on the object of study, namely NAmNo. We present a brief overview of the language, and we also discuss some core issues relating to data availability and baseline comparison. Then we turn to a brief overview of the DP structure in homeland Norwegian, which serves as our point of reference when considering the DP structure in NAmNo.

8.2.1 Basic facts North American Norwegian

Norwegians first started emigrating to the United States in the 1830s and quite soon researchers became interested in documenting and studying the Norwegian that these settlers spoke (Flom 1903; Haugen 1953; Hjelde 1992). These speakers came from many different areas in Norway, and as a result, they brought with them a wide variety of dialects (for more on the language and dialect situation in Norway, see Haugen 1976). In the first generations, they retained the ability to read Norwegian too, but this varies significantly in the later generations. This means that it is not trivial to identify the baseline for these speakers. Ideally, the baseline would be the input they received, but given the dialectal variability, this input could often contain quite some variation. The quality of the sociolinguistic background information for each of the speakers varies a lot, which means that it is not always clear what their dialect background is. Often, they only know approximately from where their family emigrated, and sometimes they do not remember or even misremember. For those speakers where we have reliable information, we can use older descriptions (when available) of these homeland dialects to estimate what the baseline would be like. In the present chapter, we use present-day homeland Norwegian as a point of comparison, but we do not consider this the target for NAmNo. We furthermore consider the data at a group level and do not delve into the quite substantial individual variation that exists.

Until the end of World War II, many of the NAmNo speakers lived in close-knit communities where Norwegian was in daily use. Many speakers were probably more dominant in Norwegian than in English. However, in

the aftermath of the war, these communities quite quickly dissolved, although some families and speakers continued using the language till the present day (see Lovoll 1999 for a sociographic history of the language and its communities and see Natvig 2022 for a case study of a community in language shift). With the establishment of the spoken Corpus of American Nordic Speech (CANS; Johannessen 2015a) in 2010, it has become possible to investigate the language of a large number of speakers. Currently, the Norwegian part of the corpus consists of 246 individuals and roughly 730,000 words. It consists of interviews with and conversations between heritage speakers of Norwegian, and it contains sound and video files together with transcriptions. Older recordings from Haugen and others have also been embedded into the corpus, but in this chapter, we will focus on the data collected since 2010 (170 speakers, roughly 650,000 words). This more recent data collection involved speakers who all are above 70 years of age, and their proficiency in Norwegian varies widely. As a result, they often mix Norwegian and their majority language English. An exciting new development is the emergence of research on Latin American Norwegian, which complements the data from NAmNo since the majority language in this case is Spanish and not English. However, since this work is just starting, we will limit ourselves to NAmNo in this chapter.

The NAmNo corpus has generated a lot of research into NAmNo, in particular on the grammatical side. It has also inspired more experimentally oriented research where scholars have used various elicitation experiments (see especially Rødvand 2017, 2018, and van Baal 2020). In terms of grammatical phenomena in the nominal domain, the areas in (1) have been investigated so far.

(1) a. Definiteness marking (Anderssen, Lundquist, and Westergaard 2018; Johannessen 2015b; Riksem 2017; van Baal 2020)
    b. Grammatical gender (Grimstad et al. 2018; Johannessen and Larsson 2015; Lohndal and Westergaard 2016; Riksem 2018; Riksem et al. 2019; Rødvand 2017, 2018)
    c. Adjectival inflection (Johannessen and Larsson 2015; Riksem, Lohndal, and Åfarli 2021)
    d. Possessives (Anderssen, Lundquist, and Westergaard 2018)

Since we are primarily concerned with the status of the top layer of the nominal domain, we will not survey all the findings in this chapter. However, before we can look at the relevant findings, we need to introduce some core aspects of Norwegian DP structure. We do that in Section 8.2.2 before turning to the relevant NAmNo data in Section 8.2.3.

8.2.2 Norwegian DP structure

Scandinavian nominal phrases have been extensively studied since the 1980s where multiple alternative analyses have been presented and debated. In our view, Julien (2005) represents the state of the art and has become the current, commonly accepted analysis. Based on this, van Baal, Eik, and Lohndal (to appear) provide the general structural overview in (2) for Norwegian nominal phrases, illustrated with the example in (3).

(2) determiner/possessive/demonstrative > numeral > adjective > noun + article > possessive

(3) de fire stor-e bil-ene mine
DEF.PL four big-DEF car-DEF.PL my.PL
'my four big cars'

The majority of elements are prenominal, and several elements have co-occurrence restrictions. For instance, determiners, possessives, and demonstratives cannot appear together at the same time. Adjectives are positioned between these three elements and the noun. As for possessives, they can appear either prenominally or postnominally. A definite suffix attaches to the noun in definite phrases, and as (3) illustrates, Norwegian also has double definiteness, which emerges in the context of nouns modified by an adjective or numeral. As Julien (2005) explains, the semantics of the two definiteness markers is not identical: The prenominal definiteness marker encodes inclusiveness, and the postnominal marker encodes specificity (see also Anderssen 2006, 2012; Kester 1996). Note that double definiteness is obligatory, leaving either determiner out results in ungrammaticality, as shown in (4).

(4) a. den rød-e bil-en
DEF.SG red-DEF car-DEF.M.SG
'the red car'

b. *den rød-e bil
DEF.SG red-DEF car

c. *rød-e bil-en
red-DEF car-DEF.M.SG

There are cases whereby the prenominal determiner and the definite suffix can be left out in Norwegian, but we won't review these cases here (see Anderssen 2006; Julien 2005; van Baal 2020, 2024).

The examples in (3) and (4) also show that Norwegian has nominal concord. The various elements agree with each other in definiteness, number, and also grammatical gender. Traditionally, Norwegian dialects exhibit a three-gender system consisting of masculine, feminine, and neuter (Haugen 1976) and this can be considered the baseline for NAmNo as well. Gender is marked by determiners, possessives, demonstratives, and adjectives in the singular. Note that the definiteness suffix also has a three-way distinction, which often corresponds to the three genders, but whether or not this is a true gender marker or rather a marker of declension class remains a contested issue (see Lohndal and Westergaard 2021 and van Baal, Eik, and Lohndal to appear for discussion).

8.2.3 North American Norwegian DP structure

Previous research shows that NAmNo speakers master core aspects of the overall syntactic structure in nominal phrases (e.g., van Baal 2020). That is, when an element is produced, they typically place it in its right position. In (5)–(6), we provide some examples of DPs in NAmNo that have the same structure as their homeland Norwegian counterparts. The speaker code identifies the speaker from the corpus, and in the case the data come from production experiments, they are also marked with "exp" at the end of the code.

(5) a. en stor boks (albert_lea_MN_01gk)
DEF.SG big box
'a big box'

b. ei lit-a stund (coon_valley_WI_07gk)
INDEF.SG small-INDEF.F while
'a little while'

(6) a. den først-e gang-en (billings_MT_01gm)
DEF.SG first-DEF time-DEF.M.SG
'the first time'

b. de sist-e år-a (blair_WI_01gm)
DEF.PL last-DEF year-DEF.N.PL
'the last years'

However, there are also important deviations from homeland Norwegian. Van Baal (2020) argues that a typical modified definite phrase in American Norwegian lacks the prenominal determiner whereas the postnominal determiner is in place. Examples of this are provided in (7)–(8).

(7) blå skjorte (hendricks_MN_07gk_exp)
blue shirt
'a blue shirt' (van Baal 2020: 108–109)

(8) a. hvit-e hest-en (westby_WI_06gm_exp)
white-DEF horse-DEF.M.SG
'the white horse' (van Baal 2020: 113–114)

b. tre ung-ene (sunburg_MN_15gm_exp)
three child-DEF.PL
'the three children' (van Baal 2020: 114)

In (7), the indefinite determiner is missing, whereas the definite determiner is missing in (8). Note that for all three examples, the noun ending is just like homeland Norwegian. Furthermore, the examples also demonstrate that there is a D-layer present in these phrases despite not being phonetically realized. The concord between the noun and the adjective is as predicted if there had been a phonetically realized determiner: There is no inflection on the adjective in (7), whereas the definite phrase in (8a) has the required adjectival inflection.

Furthermore, even though the postnominal determiner is generally in place, a small subgroup of speakers also omits this determiner. That is, they produce utterances like the ones in (9).

(9) a. den blå-e bok (sunburg_MN_11gk_exp)
DEF.SG blue-DEF book
'the blue book' (van Baal 2020: 108)

b. det sist-e år (blair_WI_04gk)
DEF.SG.N last-DEF year

In these examples, the noun should have been inflected such that it would have been *boka* and *året*, respectively. It has been argued that speakers who produce such phrases are less proficient than speakers who use the suffix consistently and therefore show more signs of cross-linguistic influence from English (Anderssen, Lundquist, and Westergaard 2018).

Lastly, there are also cases where speakers use an English determiner instead of a Norwegian determiner, a topic of great interest in recent work (Alexiadou and Lohndal 2018; Alexiadou et al. 2015; Grimstad, Lohndal, and Åfarli 2014; Grimstad et al. 2018; Lohndal and Putnam 2021; Riksem 2017, 2018; Riksem et al. 2019). Examples of this from Riksem (2018) are provided in (10).

(10) a. the by (chicago_IL_01gk)
the city
'the city'

b. the penger (albert_lea_MN_01gk)
the money.PL
'the money'

c. the andre dagen (vancouver_WA_01gm)
the other day.DEF
'the other day'

In addition, there are also cases where the Norwegian definite suffix appears on English words, as illustrated in (11) (Riksem 2018).

(11) a. field-a (coon_valley_WI_02gm)
field-DEF.F.SG
'the field'

b. shed-et (westby_WI_06gm)
shed-DEF.N.SG
'the shed'

c. chopper-en (blair_WI_01gm)
chopper-DEF.M.SG
'the chopper'

In both (10) and (11), the syntactic structures are intact, but they demonstrate deviations from the homeland variety in that there is substantial language mixing involved.

We have been assuming that the homeland variety is a representative point of comparison for these heritage speakers. To assess this assumption, van Baal (2020: 41–44) considers data from Haugen (1953) and van Baal (2022) investigates the historical corpus data in CANS to see if the input to the later generations would be likely to contain double definiteness. In both, it is found that modified definite phrases are not very common, but when they occur, double definiteness is used just like in the homeland variety. There are instances where the prenominal determiner does not occur, but these are in line with the exceptions that we find in homeland Norwegian. This means that the homeland variety is a valid point of comparison, and it also suggests that

there has been a change such that the structure of nominal phrases diverges more clearly from the homeland variety in the later (present-day) heritage speakers (see van Baal 2022).

8.3 The puzzle: Stability of DP?

The fact that NAmNo has been well studied on production data (both corpus data and elicitation data) makes it ideal to investigate general claims about vulnerability and stability in the nominal phrase. Based on data from multiple heritage languages, Polinsky (2018: 65, 211) argues that high projections in the syntactic structure are resilient to change or restructuring. In addition, she claims that "heritage speakers are best able to maintain forms and structures that have strong perceptual salience or rely on salient conceptual categories" (Polinsky 2018: 64). This applies to tense in the clausal domain, and determiners and the D-layer in nominal phrases.

These claims about the stability of the D-layer and the determiner seem at odds with the data from NAmNo. As noted in the previous section, several studies have found vulnerability of the determiner in NAmNo. In modified definite phrases, we find frequent and consistent omission of the determiner (Anderssen, Lundquist, and Westergaard 2018; van Baal 2020). Although the D-layer is clearly present in these phrases, it is not overtly realized by a determiner. In addition, we find cases where the English determiner replaces the Norwegian determiner (Riksem 2017, 2018).

In other words, the data from NAmNo present us with a puzzle: is the DP as stable and resilient to change as previously claimed? How can we explain the vulnerability of determiners in NAmNo? This puzzle is at the core of this chapter, and our aim is to solve it in the next section. As we will see there, the NAmNo data provide important nuances and adjustments to Polinsky's claims about the stability of DP.

Before we turn to this, however, the NAmNo data present us with another puzzle. It has been argued that heritage speakers dislike silent elements and prefer overt ones (Laleko and Polinsky 2017; Polinsky 2018: 253–270), which seems at odds with the empty (unrealized) determiner in NAmNo modified definite phrases. However, it seems that the claim about silent elements is more specific. The studies arguing for this "Silent Problem" are based on phenomena where both overt and silent heads occur in the language (e.g., both overt pronouns and pro-drop) and where discourse pragmatics determines which option is acceptable. These studies show that heritage speakers have difficulty

(both in production and in comprehension) with such discourse-licensed silent elements, and then prefer the overt element. However, determiner omission in NAmNo is a different type of phenomenon, as it is not discourse-licensed. In that respect, the data are not necessarily at odds with the claims about difficulty with silent elements. It is also important that the vulnerability of determiners in NAmNo is not only seen in omission of the determiner, but also in replacement with English determiners (see above and Section 8.2). Therefore, we focus in this chapter on the apparent contradiction between the observed vulnerability of determiners in NAmNo and the claimed stability of the DP.

8.4 Analysis

In this section, we aim to solve the puzzle presented above. We do this by looking at two aspects: (i) the role of perceptual versus structural salience, and (ii) the stability of a lower determiner-like position in Norwegian. We argue that Polinsky's (2018) claim can be nuanced or adjusted when we take these aspects into account, and that the NAmNo data are not in conflict with this nuanced take on stability and vulnerability in the nominal domain.

8.4.1 Perceptual versus structural salience

Polinsky (2018: 64) notes that elements with perceptual or structural salience tend to be stable in heritage language grammars. There are two types of support for this: erosion or loss of forms that lack salience, and the maintenance of forms that are salient. The former can be observed in the erosion of the DOM-marker *a* in heritage Spanish (Polinsky 2018: 166–168; see Montrul, Bhatt, and Girju 2015 and Montrul and Sánchez-Walker 2013 for details). The latter, maintenance of salient elements, is for example observed in the stability of determiners in heritage English (Polinsky 2018: 64, 175).

Salience can be defined as "the property of a linguistic item or feature that makes it in some way perceptually and cognitively prominent" (Kerswill and Williams 2002: 81). This definition indicates that there are several ways in which an item can be prominent, as also discussed by Polinsky (2018). On the one hand, there is perceptual salience, meaning that an item is more noticeable, which primarily means "phonologically heavier forms" (Polinsky 2018: 165). A relevant aspect of this is stress, as stress makes an element more

phonologically heavy and hence easier to perceive. The erosion of *a* as a DOM marker in heritage Spanish, mentioned above, is an example of the vulnerability of perceptually non-salient material. On the other hand, there is structural salience, which means that an element occurs high(er) in the syntactic projection. Within the nominal domain, determiners occur in the highest projection (the D-layer) and are therefore structurally salient.

When we consider determiners in (NAm) Norwegian, we see that their level of salience is rather different depending on the type of salience that one investigates. From a structural perspective, the determiner is salient because of its high syntactic position. In fact, they take the highest position in the Norwegian nominal phrase, and are therefore the most salient element in the phrase.[1] Perceptually, however, determiners are not salient, because they are short and unstressed words (*den, det, de* in respectively the masculine, neuter, and plural form). On top of that, the Norwegian determiner could be considered an extrametrical syllable, because it falls outside the typical trochaic prosodic pattern. As illustrated in (12), the determiner adds an unstressed (W) syllable to an otherwise trochaic pattern (of Strong-Weak syllables), which is the preferred metrical template of Norwegian and Swedish. It has been argued that this lack of prosodic salience is one of the reasons why Norwegian and Swedish children acquire the prenominal determiner much later than the definite suffix (Anderssen 2006; Bohnacker 2004; Kupisch et al. 2009).

(12) a. den grønn-e bil-en
W S-W S-W
'the green car'

b. grønn-e bil-en
S-W S-W
'the green car'
(based on Bohnacker 2004: 233)

Taken together, a diverse set of evidence – including L1 acquisition data – suggests that the prenominal determiner and the suffix have different properties in Norwegian, which is important in order to understand their behavior in heritage varieties.

With respect to the puzzle presented in Section 8.3, we believe that (part of) the solution is the fact that there are different types of salience, and that this

[1] Julien (2005) argues for separate projections for demonstratives and strong quantifiers on top of DP. We follow Van Baal, Eik, and Lohndal (to appear) in assuming that demonstratives are located in DP rather than on top of DP. In addition, quantifiers are not obligatory and therefore not present in all phrases. In most phrases, the D-layer is thus the highest projection of the nominal phrase in Norwegian.

has to be taken into account in formulating hypotheses about vulnerability in heritage language grammars. When perceptual and structural salience go hand in hand, the predictions for heritage languages are relatively clear: an element that is salient (in both respects) is predicted to be stable, whereas an element that is non-salient (in both respects) is predicted to be vulnerable. The perhaps most interesting case is where perceptual and structural salience diverge. This is exactly the case for Norwegian prenominal determiners: they are perceptually non-salient, but structurally salient.[2] We have argued that in the case of NAmNo, the lack of perceptual salience "overrides" the structural salience. As we pointed out, the same has been argued for (monolingual) Norwegian acquisition, where the lack of perceptual salience is one of the factors in the late acquisition of the prenominal determiner.[3]

Polinsky (2018) argues that elements "that have strong perceptual salience or rely on salient conceptual categories" are stable, where the "salient conceptual categories" could be interpreted as elements with structural salience. When we take into account the contradictory effects that perceptual and structural salience may have, and facts on how these two types of salience relate to each other in the homeland language, the findings from NAmNo in van Baal (2020) and Riksem (2017, 2018) are not at odds with Polinsky's claim. They nuance that claim by noticing that perceptual salience is a stronger factor in the NAmNo nominal domain than structural salience. As a result, the prenominal determiner in modified definite phrases is vulnerable for omission or replacement by an English determiner. It remains an open question whether the same "hierarchy of saliences" applies to other domains and other (heritage) languages, but the stability of the determiner in heritage English (observed by Polinsky 2018) suggests that in this case, the structural saliency outweighs the lack of perceptual salience.

8.4.2 Stability of a lower position

Norwegian has two types of definiteness marking elements: prenominal determiners (definite and indefinite) which are free-standing morphemes, and definite determiners that occur as suffixes on the noun. In unmodified definite phrases, only the suffix occurs, but modified definite phrases contain both a prenominal determiner and a suffixed determiner. As was noted in Section 8.2,

[2] This is not only the case for Norwegian: determiners are unstressed in many languages (including in English) and hence generally lack perceptual salience.

[3] Another example that phonological salience can override structural salience is left-edge ellipsis (LEE) in English, where clause-initial subjects or determiners may be omitted if they are phonologically weak (see Weir 2012, 2022).

the prenominal determiner is vulnerable in NAmNo, but the definite suffix is generally stable. In unmodified phrases, NAmNo speakers hardly omit the suffixed article (van Baal 2020), and the suffix is also very stable in modified definite phrases. In those, the determiner is frequently left out, but the suffix is only occasionally omitted by a small subgroup of the speakers (Anderssen, Lundquist, and Westergaard 2018, van Baal 2020). The definite suffix is also used on nouns from an English origin (Riksem 2017, 2018; see the examples in (11)) and furthermore stable in terms of inflection for grammatical gender (Johannessen and Larsson 2015; Lohndal and Westergaard 2016).[4]

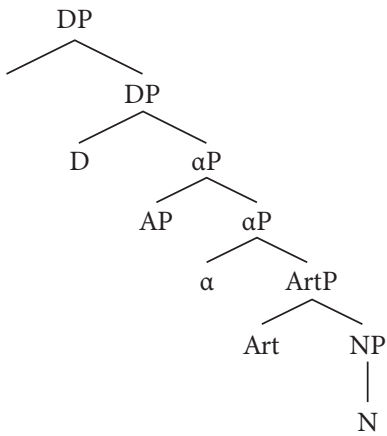
In homeland Norwegian, the definite suffix is also very stable compared to free-standing prenominal determiners. In monolingual acquisition, the definite suffix is acquired (long) before the definite determiner in double definiteness (Anderssen 2006, 2012) and also earlier than the indefinite determiner (Kupisch et al. 2009). Several Norwegian dialects are currently undergoing a change in grammatical gender, where the feminine gender is lost, and traditionally feminine nouns instead occur with masculine forms. In this change, the feminine definite suffixes are retained as a declension class, while the feminine indefinite determiner is replaced by the masculine one (i.e., a change from ei_F *bok, bok-*a_F to en_M *bok, bok-*a_F 'a book, the book'). This asymmetry between prenominal determiner and definite suffix in the loss of feminine gender has been observed in Oslo (Lødrup 2011), Tromsø (Rodina and Westergaard 2015), and Trondheim (Busterud et al. 2019). The stability of the definite suffix is thus not specific for NAmNo, but typical for Norwegian in general. And conversely, the prenominal determiner is weaker in the sense of being late acquired, subject to change, and prosodically non-salient.

The existence of double definiteness has led to much theoretical work on the Scandinavian and Norwegian nominal phrase (e.g., Delsing 1993; Taraldsen 1990; Vangsnes 1999). Since Taraldsen (1990), it is common to assume two "determiner-like" positions in the nominal domain. This position is also taken in Julien (2002, 2005) and the analysis we adopt here. A slightly simplified nominal cartography based on Julien (2005) is given in (13). The high position is labeled DP and occurs on top of the nominal domain, and the lower position is labeled ArtP or *n*P. This lower position is located below the adjectival projection, and is the position where the suffixed article is generated. The

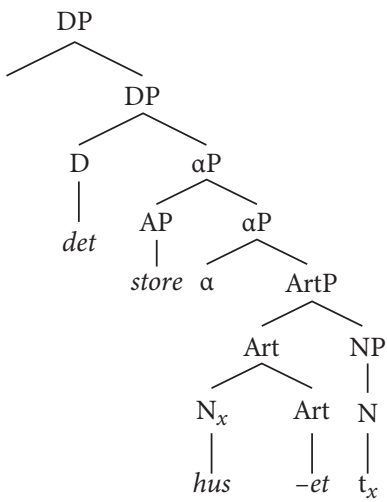
[4] There has been much debate about whether or not the definite suffix should be seen as an exponent of grammatical gender, or rather as a declension class marker (see Lohndal and Westergaard 2021 for a review), and Lohndal and Westergaard (2016) take it to be an expression of declension class while Johannessen and Larsson (2015) consider it grammatical gender. Stepping aside from this debate, it is clear that the inflection of the definite suffix is highly homeland-like in NAmNo.

prenominal determiner, on the other hand, is the realization of the higher DP. An example of a phrase with both the prenominal determiner and the suffixed article is given in (14). Here, the noun moves to merge with Art in order to combine with the definite suffix. In unmodified phrases, where no adjective is present, ArtP moves to Spec-DP, but this movement is blocked if an adjective is present (see Julien 2005: 28 for details).

(13) Syntactic structure of the Norwegian nominal domain (based on Julien 2002, 2005)



(14) Syntactic structure of the phrase *det store huset* 'the big house'



To account for the co-occurrence of the prenominal determiner and the suffixed article, Norwegian (and the other Scandinavian languages; see Julien 2002, 2005) has a more detailed cartography than many other languages. The lower position, ArtP, where the definite suffix is located, is a projection that is not present in most languages. It is for example not necessary to assume this position for English, one of the languages that are mentioned by Polinsky (2018) to have stable determiners in the heritage variety.

The data from Norwegian and NAmNo discussed above show that there is an asymmetry between the two functional projections. The DP-layer is less stable than the ArtP-layer, both when it comes to the heritage language and in acquisition and change in the monolingual homeland variety. This asymmetry is larger in the heritage variety, where the determiner is often omitted. It has been documented that patterns or tendencies present in the homeland language can be extended or amplified in the heritage language (e.g., Polinsky 2018: 167, 274; see also Kupisch and Polinsky 2022 and references therein). The case at hand may be a similar situation, where the determiner – already somewhat vulnerable in Norwegian – becomes more vulnerable in NAmNo. In this case, the vulnerability of the determiner originates in the homeland variety, but, given Polinsky's claim, that still does not explain why the determiner becomes *more* vulnerable in the heritage language.

We believe that an important observation, in this respect, is that there is a stable functional element in the nominal spine of Norwegian. Unlike in other languages, however, this stable functional element is not the determiner but the suffixed definite article. It is thus not the highest functional element that is most stable, but there is a stable "determiner-like" element in the nominal domain, and this projection also expresses definiteness. Polinsky (2018: 64) assumes that salient conceptual categories (like definiteness) occur at the top of their domain. This holds for languages with a D-N dichotomy, but the data from NAmNo show that this does not hold for languages with a more detailed nominal cartography. If we take Polinsky's claim to apply to all heritage languages, the NAmNo data are in conflict with this. However, if we restrict Polinsky's claim to languages with a simple DP-NP structure (like English), the claim does not apply to NAmNo. In languages with a more detailed nominal cartography, the salient conceptual category definiteness need not occur in the structurally most salient position. Rather, it can be expressed in another syntactic position and still be stable in the heritage variety, as the data from NAmNo show.

The discussions in Section 8.4 underline the importance of understanding the homeland (or baseline) variety and its syntactic cartography. Based on

the data from NAmNo, we expect that other heritage languages with multiple determiner-like projections will show stability of one of these projections, but not necessarily the structurally most salient (i.e., highest) projection.

8.5 Comparison with other domains

Polinsky (2018) compares the D-layer with Tense, which is also at the top of the "middle field" of the clause. Her claim is that tense is stable as well (p. 174). Eide and Hjelde (2015) and Lykke (2020) have investigated tense in NAmNo, and they concur with Polinsky (2018). In particular, through a detailed and careful investigation, Lykke (2020) demonstrates that NAmNo generally retains the allomorphy characteristic of Germanic tense systems where verbs are split into regular and irregular classes. That is, the various distinct phonologically unpredictable suffixes, stem alternations, and combinations of the two are generally consistent with the baseline. NAmNo speakers do not seem to be influenced by English in their mastery of the tense system, but their particular system depends on the dialect area from which their ancestors emigrated. The following examples in (15)–(17) illustrate tense inflections that are in line with the baseline, and the syntactic structure is also as we would expect. The examples are taken from Lykke (2020) and follow his way of rendering them phonetically.

(15) fårelldr-a minne domm (sunburg_MN_12gk)
parent-DEF.M.PL mine they
prat-a nåssjt
speak-PAST Norwegian
'My parents, they spoke Norwegian.' (Lykke 2020: 115)

(16) da **bynn-te** e å baka leffse (westby_WI_01gm)
then begin-PAST I to bake lefse
'Then I began to make lefse.' (Lykke 2020: 117)

(17) så æ **skreiv-Ø** såmm dæmm snakkt (webster_SD_01gm)
so I write-PAST like they speak.PAST
'So I wrote like they spoke.' (Lykke 2020: 126)

(18) å så ha snor-a **røk-i** (westby_WI_01gm)
and then have line-DEF.F.SG snap-PART
'And then the line has snapped.' (Lykke 2020: 127)

(15)–(16) exemplify regular inflection whereas (17)–(18) illustrate irregular inflection.

Of course, even though the overall picture is stability, Lykke also finds examples of change compared to the baseline. The most common type is overgeneralization of the regular classes of inflection. In particular, this happens with the regular past tense inflection *-dde* which is extended to previously irregular verbs with roots ending in long vowels. An example is given in (19).

(19) nå vi **gå-dde** der å så (sunburg_MN_03gm)
when we go-PAST there too so

gå-dde vi på aislen er.
go-PAST we on Iceland Air

'When we went there, we too went with Iceland Air.' (Lykke 2020: 172)

There are also instances where the morphosyntactic distribution of forms is different, for instance when present tense forms are used in past tense contexts. In the literature, this is often argued to not occur, cf. Benmamoun, Montrul, and Polinsky (2013: 142). The example in (20), slightly abbreviated from Lykke (2020), illustrates a case where the present tense form is used in a past tense context.

(20) dær e # vår e **livv-i** till # […] (coon_valley_WI_12gm)
there is.PRES where I live-PRES to

'That is where I lived until …' (Lykke 2020: 210)

The target form here would have been *liv-de* 'live-PAST'. Lykke (2020) provides additional examples, demonstrating that (20) is not an isolated occurrence. It is also clear that we are dealing with the *realization* of morphosyntactic features, since the syntax is by and large as expected in these cases.

Summarizing, we see that the differences between NAmNo and the baseline in the tense domain mostly relate to morphophonological properties. There are no cases whereby the tense morpheme is simply left out or other evidence that would point toward a "weakened" T-head in the syntactic structure.

It is also of interest to compare the D-layer to C. Just like D constitutes the top of the nominal phrase, C constitutes the top of the clause (setting aside that this projection may be split into multiple heads as in the cartographic tradition; see Rizzi 1997 and much work since). In NAmNo, the verb second property has been studied carefully. Typically, V2 is defined as the requirement that the finite verb appears in second position, which is generally taken to involve movement to C (though see Westergaard, Lohndal, and Alexiadou 2019 on

potential differences between subject-initial and non-subject-initial declaratives). Although V2 is largely stable in NAmNo, at least verb movement has been shown to be subject to cross-linguistic influence from the majority language (Eide and Hjelde 2015; Westergaard and Kupisch 2020; Westergaard and Lohndal 2019; Westergaard, Lohndal, and Lundquist 2023). Instead of V2, the heritage speakers sometimes place the verb after the subject, as shown in (21).

(21) Og der dem **lage-r** vin.
and there they make-PRES wine
'And there they make wine.'
Target: Og der lager dem vin. (Eide and Hjelde 2015: 89)

Similar findings have also been attested for other Germanic heritage languages (Westergaard and Kupisch 2020; although see Håkansson 1995). This suggests that the C-layer may be more similar to the D-layer in that the syntactic properties of the layer are affected. In the D-layer we saw instances of silence or English determiners, whereas in the C-layer we see instances of verbs not moving high enough in the syntactic structure. Taken together, this points again to the conclusion that properties at the top of a syntactic structure do not have to be stable.

8.6 Conclusion

In this chapter, we have focused on stability in the nominal domain in North American Norwegian (NAmNo) and compared data from NAmNo with the claims in Polinsky (2018) about the stability of the DP and determiners. We have argued that (i) properties at the top of a syntactic structure do not have to be stable and (ii) stable properties do not need to be at the top of the structure.

With respect to our first claim, we showed that prenominal determiners in NAmNo are vulnerable for omission or replacement by an English determiner, despite the fact that they are at the top of the nominal syntactic structure. This aligns with findings about PPs and CPs (see Polinsky 2018: 65). For the prenominal determiner, we argued that its lack of perceptual salience accounts for its vulnerability. Our second claim is supported by data on the definite suffixed article, which is very stable in NAmNo despite its lower syntactic position. Based on the data from NAmNo presented here, we have argued that previous claims on stability of the determiner (Polinsky

2018) do not automatically apply to languages (like Norwegian) with a more detailed nominal cartography.

In general, we believe that our findings point toward the need for more fine-grained predictions regarding the resilience (or vulnerability) of grammatical features and their exponents in heritage language grammars. When such predictions are formulated, it is important to take into account different types of salience (perceptual and structural), as these may have different effects on the stability of an element. Furthermore, we argued that it is important to consider the syntax of the baseline variety and the different syntactic projections it contains. The rather complex cartography of the Norwegian nominal domain allowed us to investigate which syntactic projection is most stable in NAmNo. In addition to these two factors, there are of course also other factors that have to be taken into account in our predictions about resilience in heritage languages. Frequency and age of acquisition (in monolingual acquisition) are examples of such other factors. As we have shown in this chapter, the vulnerability of the determiner (and, in relation, the stability of the definite suffix) is not unique for NAmNo. Instead, this is a recurrent pattern that is also observed in child language acquisition and an ongoing change in homeland Norwegian.

We have argued that the D-layer is not necessarily stable in all heritage languages, and it can thus be vulnerable in other heritage languages than NAmNo as well. We would expect this to happen in languages that have another definiteness-related projection in their nominal phrase, as is the case in Norwegian. An important question is when the determiner becomes vulnerable. As pointed out in Section 8.2.3, research on the use of double definiteness in previous generations of NAmNo speakers has shown that these speakers used the determiner in a stable way (van Baal 2022). The present-day speakers are the first generation to show vulnerability in the determiner. Van Baal (2022) argues that this can be explained by the sociolinguistic context of the speakers, since the present-day NAmNo speakers live in a post-language-shift community where Norwegian is no longer used for daily communication. This illustrates that factors in the sociolinguistic environment of the heritage language should also be considered in our understanding of heritage language syntax. In certain contexts of use, certain elements can be more vulnerable to omission or restructuring. For NAmNo, this vulnerable element is the prenominal determiner while the definite suffixed article is more stable.

References

Aalberse, Suzanne, Ad Backus, and Pieter Muysken (2019) *Heritage languages: A language contact approach*. Amsterdam: John Benjamins.

Aboh, Enoch O. (2010) Information structuring begins with the numeration. *Iberia: An International Journal of Theoretical Linguistics*, 2(1), 12–42.

Aboh, Enoch O. (2015) *The emergence of hybrid grammars: Language contact and change*. Cambridge: Cambridge University Press.

Aboh, Enoch O. (2017) The emergence of hybrid grammars: A rejoinder to Peter Bakker. *WORD*, 63(3), 207–222. https://doi.org/10.1080/00437956.2017.1347316

Aboh, Enoch O. and Michel DeGraff (2017) A null theory of Creole formation based on universal grammar, in I. Roberts (ed.) *The Oxford handbook of universal grammar*. Oxford : Oxford University Press, 401–458. https://doi.org/10.1093/oxfordhb/9780199573776.013.18

Aikhenvald, Alexandra Y. (2006) Reflections on language contact, areal diffusion, mechanisms of linguistic change, in C. Bernard and Z. Petr (eds.) *Sprachbund in the West African Sahel*. Louvain: Peeters, 23–26.

Aikhenvald, Alexandra Y. and Robert M. W. Dixon (2006) *Grammars in contact*. Oxford: Oxford University Press.

Alexiadou, Artemis (2005) Gerund types, the present participle and patterns of derivation, in C. Maienborn and A. Wöllstein-Leisten (eds.) *Event arguments in syntax, semantics and discourse*. Tübingen: Niemeyer, 139–152.

Alexiadou, Artemis (2013) Nominal vs. verbal *-ing* constructions in the development of the English progressive. *English Linguistic Research*, 2(2), 126–140.

Alexiadou, Artemis (2019) A form-function mismatch? The case of Greek deponents, in J. M. M. Brown, A. Schmidt, and M. Wierzba (eds.) *Of trees and birds: A Festschrift for Gisbert Fanselow*. Potsdam: Universitätsverlag Potsdam, 107–117.

Alexiadou, Artemis (2024) Greek deponent verbs and their nominalizations: Genitive case as unmarked case in the nominal domain, in C. Sevdali, D. Mertyris, and E. Anagnostopoulou (eds.) *The place of case in grammar*. Oxford: Oxford University Press, 467–491.

Alexiadou, Artemis and Elena Anagnostopoulou (2004) Voice morphology in the causative-inchoative alternation: Evidence for a non-unified structural analysis of unaccusatives, in A. Alexiadou, E. Anagnostopoulou, and M. Everaert (eds.) *The unaccusativity puzzle: Explorations of the syntax-lexicon interface*. Oxford: Oxford Studies in Theoretical Linguistics, 114–136.

Alexiadou, Artemis, Gianina Iordăchioaia, and Elena Soare (2010) Number/aspect interactions in the syntax of nominalizations: A distributed morphology approach. *Journal of Linguistics*, 46, 537–574.

Alexiadou, Artemis, Elena Anagnostopoulou, and Florian Schäfer (2015) *External arguments in transitivity alternations: A layering approach*. Oxford: Oxford University Press.

Alexiadou, Artemis, Terje Lohndal, Tor A. Åfarli, and Maren Berg Grimstad (2015) Language mixing: A distributed morphology approach, in T. Bui and D. Özyildiz (eds.) *Proceedings of the Fourty-Fifth Annual Meeting of the North East Linguistic Society*. Amherst, MA: GSLA, 25–38.

Alexiadou, Artemis and Terje Lohndal (2018) Units of language mixing: A cross-linguistic perspective. *Frontiers in Psychology*, 9, 1719. DOI: 10.3389/fpsyg.2018.01719

Alexiadou, Artemis and Vasiliki Rizou (2023) The use of periphrasis for the expression of aspect by Greek heritage speakers: A case study of register variation narrowing. *Register Studies*, 5(1), 82–110.

Alexiadou, Artemis, Vasiliki Rizou, and Fotini Karkaletsou (2022) A plural indefinite article in heritage Greek: The role of register, in G. Giusti and A. Cardinaletti (eds.) *Languages: Special issue on variation and optionality in determiner systems across the world languages*, 7(2), 115.

Alter, Kai, Ina Mleinek, and Nicole Richter (2001) Prosodic phrasing and accentuation in Russian, in G. Zybatow, U. Juhghanns, L. Szucsich, and G. Mehlhorn (eds.) *Current issues in formal Slavic linguistics*. Frankfurt: Peter Lang, 317–328.

Alonso-Ovalle, Luis, Susana Fernandez-Solera, Lyn Frazier, and Charles Clifton (2002) Null vs. overt pronouns and the topic-focus articulation in Spanish. *Rivista di Linguistica*, 14(2), 151–169.

Anastasiadis-Symeonidis, Anna and Asimakis Fliatouras (2003) The features [learned] and [- learned] in Greek: Definition and classification, in G. Katsimali, A. Kalokerinos, E. Anagnostopoulou, and I. Kappa (eds.) *Proceedings of the 6th International Conference on Greek Linguistics*. University of Crete, 110–120.

Anderssen, Merete (2006) *The acquisition of compositional definiteness in Norwegian*. Doctoral dissertation, University of Tromsø.

Anderssen, Merete (2012) A spanning approach to the acquisition of double definiteness in Norwegian. *IBERIA: An International Journal of Theoretical Linguistics*, 4, 1–34.

Anderssen, Merete, Björn Lundquist, and Marit Westergaard (2018) Crosslinguistic similarities and differences in bilingual acquisition and attrition: Possessives and double definiteness in Norwegian heritage language. *Bilingualism: Language and Cognition*, 21, 748–764.

Andrews, Avery (1990) Unification and morphological blocking. *Natural Language and Linguistic Theory*, 8(4), 507–557.

Andriani, Luigi and Roberta D'Alessandro (2022) D, T, C drop in heritage Italo-Romance in NYC: Some descriptive generalizations, in Gabriela Alboiu, Daniela Isac, Alexandru Nicolae, Mihaela Tănase-Dogaru & Alina Tigău (eds.) *A life in linguistics: A festschrift for Alexandra Cornilescu on her 75th birthday*. Bucharest, Romania: Bucharest University Press, 41–78.

Andriani, Luigi, Jan Casalicchio, Francesco Ciconte, Roberta D'Alessandro, Alberto Frasson, Brechje van Osch, Luana Sorgini, and Silvia Terenghi (2022a) Documenting Italo-Romance heritage languages in the Americas, in M. Coler and A. Nevins (eds.) *Contemporary research in minority and diaspora languages of Europe*. Berlin: Language Science Press, 9–56. https://langsci-press.org/catalog/view/332/3536/2466-1

Andriani, Luigi, Roberta D'Alessandro, Alberto Frasson, Brechje van Osch, Luana Sorgini, and Silvia Terenghi (2022b) Adding the microdimension to the study of language change in contact: Three case studies. *Glossa: A Journal of General Linguistics*, 7(1), 1–36. https://doi.org/10.16995/glossa.5748

Avram, Larisa, Elena Soare, and Andrew Mardale (2022) Asymmetries in the acquisition of relative clauses in child heritage Romanian. Paper presented at the Language Acquisition Seminar, Goethe-Universität Frankfurt, November 15, 2022.

van Baal, Yvonne (2020) *Compositional definiteness in American heritage Norwegian*. Doctoral dissertation, University of Oslo.

van Baal, Yvonne (2022) New data on language change: Compositional Definiteness in American Norwegian. *Heritage Language Journal*, 19, 1–32. DOI: 10.1163/15507076-bja10005

van Baal, Yvonne (2024) (Den) eneste måten – When the prenominal determiner can be omitted from Norwegian double definite phrases. *Journal of Comparative Germanic Linguistics*, 36, 79–111. DOI:10.1017/S1470542724000011

van Baal, Yvonne, Ragnhild Eik, and Terje Lohndal (to appear) The nominal domain: Norwegian, in K. Hartmann, J. Mursell, and S. Wurmbrand (eds.) *Handbook of Germanic syntax*. Berlin: Mouton de Gruyter.

Bailyn, John Frederick (2003) A (purely) derivational approach to Russian scrambling, in W. Browne, J.-Y. Kim, B. H. Partee, and R. A. Rothstein (eds.) *Formal approaches to Slavic linguistics: The Amherst meeting*. Ann Arbor: Michigan Slavic Publications, 41–62.

Bailyn, John Frederick (1995) *A configurational approach to Russian 'free' word order*. Doctoral dissertation, Cornell University, Ithaca, NY.

Bailyn, John Frederick (2002) Inversion, dislocation and optionality in Russian, in G. Zybatow, U. Junghanns, G. Mehlhorn, and L. Szucsich (eds.) *Current issues in formal Slavic linguistics*. Frankfurt: Peter Lang, 280–293.

Bailyn, John Frederick (2012) *The syntax of Russian*. Cambridge: Cambridge University Press.

Barbosa, Pilar (2000) Clitics: A window into the null subject property, in J. Costa (ed.) *Essays in Portuguese comparative syntax*. New York: Oxford Press, 31–93.

Barbosa, Pilar (2008) Clíticos, deslocação local e linearização cíclica. *Diacrítica, Série Ciências da Linguagem*, 22, 131–156.

Barbosa, Pilar (2014) *Corpus: Sociolinguistic profile of Braga's speech variety*. CEHUM, Universidade do Minho, Portugal. http://cehum.ilch.uminho.pt/

Barker, Chris, Jorge Hankamer, and John Moore (1990) *Wa* and *Ga* in Turkish, in K. Dziwirek, P. Farrell, and E. Mejias-Bikandi (eds.) *Grammatical relations: A cross-theoretical perspective*. Stanford, CA: CSLI Publications, 21–43.

Battye, Adrian and Ian Roberts (eds.) (1995) *Clause structure and language change*. Oxford: Oxford University Press.

Bayram, Fatih, Grazia Di Pisa, Jason Rothman, and Roumyana Slabakova (2021) Current trends and emerging methodologies in charting heritage language grammars, in S. Montrul and M. Polinsky (eds.) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, 545–578.

Bayram, Fatih, Tanja Kupisch, Diego Pascualy y Cabo, and Jason Rothman (2019) Terminology matters on theoretical grounds too! Coherent grammars cannot be incomplete. *Studies in Second Language Acquisition*, 41, 257–264.

Belletti, Adriana (2017) On the acquisition of complex derivations with related considerations on poverty of the stimulus and frequency, in E. Di Domenico (ed.) *Syntactic complexity from acquisition*. Newcastle: Cambridge Scholars Publishing, 28–48.

Belletti, Adriana (2018) On a-marking of object topics in the Italian left periphery, in R. Petrosino, P. C. Cerrone, and H. van der Hulst (eds.) *Beyond the veil of Maya: From sounds to structures*. Berlin: Mouton de Gruyter, 445–466.

Belletti, Adriana and Maria-Teresa Guasti (2015) *The acquisition of Italian: Morphosyntax and its interfaces in different modes of acquisition*. Amsterdam/Philadelphia: John Benjamins.

Belloni, Silvano (1991) *Grammatica veneta*. Este, PD: Zielo.

Benincà, Paola (1994) *La variazione sintattica. Studi di dialettologia romanza*. Bologna: Il Mulino.

Benmamoun, Elabbas (2021) Heritage language research and theoretical linguistics, in S. Montrul and M. Polinsky (eds.) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, 375–398.

Benmamoun, Elabbas, Silvina Montrul, and Maria Polinsky (2013) Heritage languages and their speakers: Opportunities and challenges for linguistics. *Theoretical Linguistics*, 39, 129–181.

Bentzen, Kristine (2015) Mapping information structure and word order, in L. Serratrice and S. E. M. Allen (eds.) *The acquisition of reference*. Amsterdam: John Benjamins, 51–82.

Berger, Tilman, Karl Gutschmidt, Sebastian Kempgen, and Peter Kosta (2009) *The Slavic languages*. Vol. 1. Berlin, Germany: De Gruyter.

Bhatt, Rajesh (2002) The raising analysis of relative clauses: Evidence from adjectival modification. *Natural Language Semantics*, 10, 43–90.

Bianchi, Valentina (1999) *Consequences of antisymmetry: Headed relative clauses*. Berlin: Mouton de Gruyter.

Bianchi, Valentina (2002) Headed relative clauses in generative syntax, Part I and Part 2. *Glot International*, 6(7), 197–204; 8, 235–247.

Biberauer, Theresa (2003) Reconsidering the EPP and Spec-TP in Germanic. *Cambridge Occasional Papers in Linguistics*, 1, 100–120.

Biberauer, Theresa (2018) *Peripheral significance: A phasal perspective on the grammaticalisation of speaker perspective*. Presentation at DiGS 20, York, England.

Biberauer, Theresa (2019) Factors 2 and 3: Towards a principled explanation. *Catalan Journal of Linguistics 2019: Special issue on generative syntax. questions, crossroads, and challenges*, 45–88. DOI: 10.5565/rev/catjl.219

Bickerton, Derek (1989) The child, the bioprogram and the input data: A commentary on Cziko. *First Language*, 9(25), 33–37.

Bickerton, Derek (1990) *Language and species*. Chicago, IL: University of Chicago Press.

Bleam, Tonia (1999) *Leísta Spanish and the syntax of clitic doubling*. Doctoral dissertation, University of Delaware.

Bloomfield, Leonard (1927) Literate and illiterate speech. *American Speech*, 2, 432–439.

Boeckx, Cedric (2014) *Elementary syntactic structures: Prospects of a feature-free syntax*. Cambridge: Cambridge University Press

Bohnacker, Ute (2004) Nominal phrases, in G. Josefsson, C. Platzack, and G. Håkansson (eds.) *The acquisition of Swedish grammar*. Amsterdam: John Benjamins, 195–260.

Börjars, Kersti and Kate Burridge (2011) From preposition to purposive to infinitival marker: The Pennsylvania German *fer … zu* construction, in M. T. Putnam (ed.) *Studies on German-language islands*. Amsterdam: John Benjamins, 385–411.

Bonet, Eulàlia (1991) *Morphology after syntax: Pronominal clitics in Romance*. MIT: Doctoral dissertation.

Bošković, Željko (2001) *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: Elsevier.

Bošković, Željko (2002) Clitics as nonbranching elements and the linear correspondence axiom. *Linguistic Inquiry*, 33(2), 329–340.

Bošković, Željko (2007) On the locality and motivation of Move and Agree: An even more minimal theory. *Linguistic Inquiry*, 38, 589–644.

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, 27–89.

Bossong, Georg (1985) *Differentielle objektmarkierung in den neuiranischen sprachen*. Tübingen: Narr Verlag.

Bossong, Georg (1991) Differential object marking, in D. Wanner and D. A. Kibbee (eds.) *New analyses in Romance linguistic*. Amsterdam: John Benjamins, 143–170. DOI: 10.1075/cilt.69.14bos

Bousquette, Joshua and Michael T. Putnam (2020) Redefining language death: Evidence from moribund grammars. *Language Learning*, 70(S1), 188–225.

Bowles, Melissa A. (2011) Exploring the role of modality: Second language-heritage learner interactions in the Spanish language classroom. *The Heritage Language Journal*, 8(1), 30–65.

Brander, Ellen (2020) A "borderline case" of syntactic variation. *Glossa*, 5(1), 1–34.

Brandi, Luciana and Patrizia Cordin (1981) Two Italian dialects and the null subject parameter, in O. A. Jaeggli and K. J. Safir (eds.) *The null subject parameter*. Dordrecht: Springer, 111–142.

Branigan, Holly P. and Martin J. Pickering (2017) An experimental approach to linguistic representation. *Behavioral and Brain Sciences*, e282, 1–61.

Brown, Joshua R. and Michael T. Putnam (2015) Functional convergence and extension in contact: Syntactic and semantic attributes of the progressive aspect in Pennsylvania Dutch, in J. B. Johannesen and J. C. Salmons (eds.) *Germanic heritage languages in North America: Acquisition, attrition, and change*. Amsterdam: John Benjamins, 135–160.

Brun, Dina (2001) Information structure and the status of NP in Russian. *Theoretical Linguistics*, 27, 109–135.

Bryzgunova, Elena A. (1981) *Zvuk i Intonacija Russkoj Reči*. Moscow: Russkij Jazyk.

Busterud, Guro, Terje Lohndal, Yulia Rodina, and Marit Westergaard (2019) The loss of feminine gender in Norwegian: A dialect comparison. *Journal of Comparative Germanic Linguistics*, 22, 141–167.

Cannizzaro, Gisi and Petra Hendriks (2020) Production before comprehension in the emergence of transitive constructions in Dutch child language. *Frontiers in Psychology*, 11, 546495.

Cardinaletti, Anna and Michal Starke (1999) The typology of structural deficiency: A case study of the three classes of pronouns, in H. van Riemsdijk (ed.) *Clitics in the languages of Europe*. The Hague: Mouton de Gruyter, 145–233.

Cardoso, Adriana (2017) *Portuguese relative clauses in synchrony and diachrony*. Oxford: Oxford University Press.

Carminati, Maria Nella (2002) *The processing of Italian subject pronouns*. Ph.D. dissertation, University of Massachusetts, Amherst.

Carminati, Maria Nella (2005) Processing reflexes of the feature hierarchy (person > number > gender) and implications for linguistic theory. *Lingua*, 115, 259–285.

Cennamo, Michela and Antonella Sorace (2007) Auxiliary selection and split intransitivity in Paduan: Variation, and lexical-aspectual constraints, in R. Aranovich (ed.) *Split auxiliary systems: A cross-linguistic perspective*. Amsterdam: John Benjamins, 65–99. DOI: 10.1075/tsl.69.05cen

Chan, Angel, Si Chen, Stephen Matthews, and Virginia Yip (2017) Comprehension of subject and object relative clauses in a trilingual acquisition context. *Frontiers in Psychology*, 8, 1641.

Chan, Angel, Wenchun Yang, Franklin Chang, and Evan Kidd (2018) Four-year-old Cantonese-speaking children online processing of relative clauses: A permutation analysis. *Journal of Child Language*, 45, 174–203.

Chandra, Pritha, D'Alessandro Roberta, and Michael T. Putnam (2024) Ethical concerns for theoretical research in linguistics: Issues and best practices, in M. Terkourafi (ed.) *Linguistics: Special issue on ethics in linguistics*, DOI: https://doi.org/10.1515/ling-2023-0154.

Chang, Charles B. (2021) Phonetics and phonology of heritage languages, in S. Montrul and M. Polinsky (eds.) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, 581–612.

Chasiotis, Ioannis, Olga Katsiardi- Hering, and Evridiki Abatzi (2006) *Oi Ellines sti diaspora 15os–21os ai.* [Greeks in Diaspora 15th–21st century]. Athens: Vouli ton Ellinon.

Chevalier, Joan F. (2004) *Heritage language literacy: Theory and practice*. UCLA, Center for Chinese Studies. Available online: https://www.international.ucla.edu/ccs/article/18712 (Accessed: 08 November 2022).

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

Chomsky, Noam (1994) *Bare phrase structure. MIT Occasional Papers in Linguistics 5*. MITWPL, Department of Linguistics and Philosophy, MIT, Cambridge, MA.

Chomsky, Noam (1995) *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam (2000) *New horizons in the study of language and mind*. Cambridge/New York: Cambridge University Press.

Chomsky, Noam (2005) Three factors in language design. *Linguistic Inquiry*, 36(1), 1–22.

Chomsky, Noam (2008) On phases, in R. Freidin, C. P. Otero, and M. L. Zubizarreta (eds.) *Foundational issues in linguistic theory*. Cambridge, MA: MIT Press, 134–166.

Chomsky, Noam and Morris Halle (1968) *The sound pattern of english*. New York: Harper and Row.

Chou, Chao-Ting Tim, Tsung-Ying Chen, and Acrisio Pires (2020) Acquisition of null objects in Mandarin Chinese by heritage speakers: Syntax-pragmatics interface knowledge without inflectional morphology. *International Journal of Chinese Linguistics*, 7, 223–269.

Ciconte, Francesco Maria (2018a) Soggetto e oggetto nell'italo-romanze antico. *Studi e Saggi Linguistici*, 56(1), 97–136.

Ciconte, Francesco Maria (2018b) La posizione del soggetto e dell'oggetto nel siciliano antico. *Bollettino del CSFLS*, 29, 177–204.

Cinque, Guglielmo (2020) *The syntax of relative clauses*. Cambridge: Cambridge University Press.

Collins, Chris (2021) *A conversation with Noam Chomsky about formal semantics: Ordinary working grammarian* [Blog]. https://ordinaryworkinggrammarian.blogspot.com/2021/06/a-conversation-with-noam-chomsky-about.html

Cordin, Patrizia. (1997). Tense, mood and aspect in the verb, in M. Maiden and M. Parry (eds.), *The dialects of Italy*. London/New York: Routledge, 87–98

Cordin, Patrizia (2021) Italo-Romance: Venetan, in Michele Loporcaro and Francesco Gardani (eds.) *Oxford encyclopedia of Romance linguistics*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.724

Cordin, Patrizia and Leonardo Degasperi (2020) Cambiamenti morfo-sintattici nel dialetto trentino di Piracicaba. *Vox Romanica*, 79, 41–60.

Costa, João, Alexandra Fiéis, and Maria Lobo (2015) Input variability and late acquisition: Clitic misplacement in European Portuguese. *Lingua*, 161, 10–26. http://dx.doi.org/10.1016/j.lingua.2014.05.009

Costa, João and Maria Lobo (2007) Clitic omission, null objects or both in the acquisition of European Portuguese, in S. Baauw, F. Drijkonongen, and M. Pinto (eds.) *Romance languages and linguistic theory 2005*. Amsterdam: John Benjamins, 59–72. DOI: https://doi.org/10.1075/cilt.291.06cos

Costa, João, Maria Lobo, and Carolina Silva (2009) Null objects and early pragmatics in the acquisition of European Portuguese. *Probus*, 21(2), 143–162. http://doi.org/10.1515/prbs.2009.005

Cyrino, Sônia M. L., Maria Eugênia Duarte, and Mary A. Kato (2000) Visible subjects and invisible clitics in Brazilian Portuguese, in M. A. Kato and E. V. Negrão (eds.) *Brazilian Portuguese and the null subject parameter*. Frankfurt am Main/Vervuert/Madrid: Iberoamericana, 55–73.

D'Alessandro, Roberta (2015) *Microcontact: Language variation and change from the Italian heritage perspective*. ERC project, https://microcontact.sites.uu.nl/project/

D'Alessandro, Roberta (2017) When you have too many features: Auxiliaries, agreement and clitics in Italian varieties. *Glossa: A Journal of General Linguistics*, 2(1), 50. https://doi.org/10.5334/gjgl.102

D'Alessandro, Roberta (2021a) Syntactic change in contact: Romance. *Annual Review of Linguistics*, 7(7), 309–328. https://doi.org/10.1146/annurev-linguistics-011619-030311

D'Alessandro, Roberta (2021b) *Understanding language change – directionality, simplification, or neither?* LINGUAE Lectures, ENS, Paris, December, 2 2021.

D'Alessandro, Roberta (2022a) Case, alignment, agreement, and information structure: The role of DOM in language change. Paper presented at the 31st Colloquium on Generative Grammar, Universitat de les Illes Balears, June 1–3.

D'Alessandro, Roberta (2022b) Crossing domains: Topic marking and doubling in Romance, in N. Boneh, D. Harbour, O. Matushansky and I. Roy (eds.) *Building on Babel's rubble*. Paris: Presses universitaires de Vincennes, 395–409.

D'Alessandro, Roberta (2023) A complex solution to an unsolvable problem? *Linguistic Approaches to Bilingualism*, 13(1), 40–44.

D'Alessandro and Claudio Di Felice (2015) The diachrony of Abruzzese complementation. *Revue Roumaine de Linguistique*, 1–2, 129–145.

D'Alessandro, Roberta and Alberto Frasson (2023) Simplification or complexification: Auxiliary selection and anti-agreement effect in Brazilian Venetan. *Catalan Journal of Linguistics*, 22, 9–27.

D'Alessandro, Roberta, David Natvig, and Michael T. Putnam (2021) Addressing challenges in formal research on moribund heritage languages: A path forward. *Front Psychol*, 12, 700126.

D'Alessandro, Roberta and Ian Roberts (2010) Past participle agreement in Abruzzese: Split auxiliary selection and the null-subject parameter. *Natural Language & Linguistics Theory*, 28(1), 41–72.

D'Alessandro, Roberta and Silvia Terenghi (2024) Non-monotonic functional sequences: A new metric for complexity in heritage languages, in M. Polinsky, M. T. Putnam, and J. Salmons (eds.) *Formal approaches to complexity in heritage language grammars*. Berlin: Language Science Press, 153–179.

Dąbrowska, Eva, Esther Pascual, and Beatriz Macías Gómez-Estem (2022) Literacy improves the comprehension of object relatives. *Cognition*, 224. https://doi.org/10.1016/j.cognition.2021.104958

Dal Picol, Greyce (2013) Novo perfil linguístico dos falantes bilíngues da Região de Colonização Italiana do Nordeste do Rio Grande do Sul: Mudança dialetal e mescla linguística. *Sociodialeto*, 3, 281–297.

Dal Picol, Greyce (2014) *A morfossintaxe na oralidade do vêneto sul-rio-grandense. Perfil dialetal de comunidades rurais da região da 4a légua, Caxias do Sul (RS)*. Universidade Caxias do Sul Doctoral Dissertation.

Daskalaki, Evangelia, Vicky Chondrogianni, Elma Blom, Froso Argyri, and Johanne Paradis (2019) Input effects across domains: The case of Greek subjects in child heritage language. *Second Language Research*, 35, 421–445.

De Houwer, Annick (2015) Harmonious bilingual development: Young families' well-being in language contact situations. *International Journal of Bilingualism*, 14, 422–446.

de Prada Pérez, Ana (2009) *Subject expression in Minorcan Spanish: Consequences of contact with Catalan.* Ph.D. dissertation, Pennsylvania State University.

de Prada Pérez, Ana (2019) Theoretical implications of research on bilingual subject production: The vulnerability hypothesis. *International Journal of Bilingualism*, 23, 670–694.

Déchaine, Rose-Marie, and Martina Wiltschko (2002) Decomposing pronouns. *Linguistic Inquiry*, 33(3), 409–442.

Delsing, Lars-Olof (1993) *The internal structure of noun phrases in the Scandinavian languages*. Doctoral dissertation, University of Lund.

Diessel, Holger (1999) *Demonstratives: Form, function, and grammaticalization*. Amsterdam/ Philadelphia, PA: John Benjamins. DOI: 10.1075/tsl.42

Diessel, Holger and Michael Tomasello (2005) A new look at the acquisition of relative clauses. *Language*, 81, 882–906.

Di Felice, Claudio and Stefano Cesinaro (2016) Documentando l'antico abruzzese, in R. Coluccia, J. Brincat, and F. Möhren (eds.) *Actes du XXVIIe Congrès international de linguistique et de philologie romanes (Nancy, 15–20 juillet 2013). Section 5: Lexicologie,*

phraséologie, lexicographie. Nancy, ATILF, 219–228. http://www.atilf.fr/cilpr2013/actes/section-5.html

Di Tullio, Ángela L., Andrés L. Saab, and Pablo D. Zdrojewski (2019) Clitic doubling in a doubling world: The case of Argentinean Spanish reconsidered, in Angel J. Gallego (ed.) *The syntactic variation of Spanish dialects*. Oxford: Oxford University Press, 215–244.

Dixon, Robert M. W. (1979) Ergativity. *Language*, 55(1), 55–138. DOI: 10.2307/412519

Domínguez, Laura, Glyn Hicks, and Roumyana Slabakova (2019) Terminology choice in generative acquisition research: The case of "incomplete acquisition" in heritage language grammars. *Studies in Second Language Acquisition*, 41, 241–255.

Dosi, Ifigeneia and Despina Papadopoulou (2019) The role of educational setting in the development of verbal aspect and executive functions: Evidence from Greek-German bilingual children. *International Journal of Bilingual Education and Bilingualism*, 23, 964–980.

Douglas, Biber and Susan Conrad (2019) *Register, genre, and style*. 2nd ed. Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press.

Dressler, Wolfgang U. (1991) The sociolinguistic and patholinguistic attrition of Breton phonology, morphology, and morphophonology, in H. W. Seliger and R. M. Vago (eds.) *First language attrition*. Cambridge: Cambridge University Press, 99–112.

Dryer, Matthew (2007) Word order, in T. Shopen (ed.) *Language typology and syntactic description*. 2nd ed., vol. 1. Cambridge: Cambridge University Press, 61–131.

Duarte, Inês and Gabriela Matos (2000) Romance clitics and the minimalist program, in J. Costa (ed.) *Portuguese syntax: New comparative studies*. Oxford: Oxford University Press, 116–142.

Duarte, Inês, Gabriela Matos, and Isabel Hub Faria (1995) Specificity of European Portuguese clitics in Romance, in I. H. Faria and M. J. Freitas (eds.) *Studies on the acquisition of Portuguese*. Lisbon, Portugal: APL/Colibri, 129–154.

Duarte, Inês, Gabriela Matos, and Anabela Gonçalves (2005) Pronominal clitics in European and Brazilian Portuguese. *Journal of Portuguese Linguistics*, 4(2), 79–112. DOI: http://doi.org/https://doi.org/10.5334/jpl.161

Dubinina, Irina and Maria Polinsky (2013) Russian in the USA, in M. Moser and M. Polinsky (eds.) *Slavic languages in migration*. Wien: University of Vienna, 131–160.

Dyakonova, Marina (2009) A phase-based approach to Russian free word order. Doctoral Dissertation. University of Amsterdam, Amsterdam.

Eggins, Suzanne (2004) *An introduction to systemic functional linguistics*. 2nd ed. London: Continuum.

Eguren, Luis, Olga Fernández-Soriano, and Amaya Mendikoetxea (eds.) (2016) *Rethinking parameters*. Oxford: Oxford University Press.

Eide, Kristin Melum and Arnstein Hjelde (2015) Verb second and finiteness morphology in Norwegian heritage language of the American Midwest, in B. R. Page and M. T. Putnam (eds.) *Moribund Germanic heritage languages in North America: Theoretical perspectives and empirical findings*. Leiden: Brill, 64–101.

Embick, David (1997) *Voice and the interfaces of syntax*. PhD thesis, University of Pennsylvania.

Embick, David (1998) Voice systems and the syntax/morphology interface, in *The Proceedings of the Penn/MIT Workshop on Aspect, Argument Structure, and Events*, May. University of Pennsylvania, 1–32.

Faggion, Carmen Maria (2013) Aspectos morfossintáticos do vêneto do Sul do Brasil, in E. C. Herrero and C. C. Rigual (eds.) *Actas del XXVI Congreso Internacional de Lingüística y de Filología Románicas*. Vol. 6. Berlin: De Gruyter, 131–142.

Fanselow, Gisbert and Denisa Lenertová (2011) Left peripheral focus: Mismatches between syntax and information structure. *Natural Language and Linguistic Theory*, 29(1), 169–209.

Fedele, Emili and Elsi Kaiser (2014) Looking back and looking forward: Anaphora and cataphora in Italian. *Proceedings of the 37th Annual Penn Linguistics Conference, Working Papers in Linguistics*, 20(1), 81–90.

Fernández-Dobao, Ana and Julia Herschensohn (2021) Acquisition of Spanish verbal morphology by child bilinguals: Overregularization by heritage speakers and second language learners. *Bilingualism: Language and Cognition*, 24(1), 56–68.

Firbas, Jan (1964) On defining the theme in functional sentence analysis. *Travaux Linguistique de Prague*, 1, 267–280.

Fischer, Olga, Hedrik De Smet, and Wim van der Wurff (2017) *A brief history of English syntax*. Cambridge: Cambridge University Press.

Fitch, Gesche Westphal (2011) Changes in frequency as a measure of language change: Extraposition in Pennsylvania Dutch, in M. T. Putnam (ed.) *Studies on German-language islands*. Amsterdam: John Benjamins, 371–384.

Flom, George T. (1903) The gender of English loan-nouns in Norse dialects in America: A contribution to the study of the development of grammatical gender. *The Journal of English and Germanic Philology*, 5, 1–31

Flores, Cristina and Pilar Barbosa (2014) When reduced input leads to delayed acquisition: A study on the acquisition of clitic placement by Portuguese heritage speakers. *International Journal of Bilingualism*, 18(3), 304–325.

Flores, Cristina and Esther Rinke (2020a) Factors constraining subject expression in European Portuguese spoken in Hamburg: A bi-generational corpus investigation. *Discours. Revue de linguistique, psycholinguistique et informatique. A Journal of Linguistics, Psycholinguistics and Computational Linguistics*, 26. https://doi.org/10.4000/discours.10648

Flores, Cristina and Esther Rinke (2020b) The relevance of language-internal variation in predicting heritage language grammars. *Bilingualism: Language and Cognition*, 23, 25–26.

Flores, Cristina, Esther Rinke, and Cecília Azevedo (2017) Object realization across generations: A closer look on the spontaneous speech of Portuguese first and second generation migrants, in E. Domenico (ed.) *Syntactic complexity from a language acquisition perspective*. Newcastle upon Tyne: Cambridge Scholars, 178–205.

Flores, Cristina, Esther Rinke, and Aldona Sopata (2020) Acquiring the distribution of null and overt direct objects in European Portuguese. *Journal of Portuguese Linguistics*, 19(1), 5. https://doi.org/10.5334/jpl.239

Flores, Cristina, Esther Rinke, Jacopo Torregrossa, and Daniel Weingärtner (2022) Language separation and stable syntactic knowledge: Verbs and verb phrases in bilingual children's narratives. *Journal of Portuguese Linguistics*, 5, 1–30.

Flores, Cristina, Esther Rinke, and Cláudia Wagner (2022) To hón ich imma insistieat: Syntactic stability in heritage Hunsrückisch German spoken in Brazil. *Linguistic Approaches to Bilingualism*, 12(3), 251–279.

Flores, Cristina, Ana Lucia Santos, Alice Jesus, and Rui Markes (2017) Age and input effects in the acquisition of mood in heritage Portuguese. *Journal of Child Language*, 44, 795–828.

Flynn, Suzanne, Claire Foley, and Inna Vinnitskaya (2004) The cumulative-enhancement model for language acquisition: Comparing adults' and children's patterns of development in first, second and third language acquisition. *International Journal of Multilingualism*, 1, 3–17. DOI: 10.1080/14790710408668175

Fotiadou, Georgia and Ianthi Tsimpli (2010) The acquisition of transitivity alternations in Greek: Does frequency count? *Lingua*, 120(11), 2605–2626.

Franks, Steven and Catherine Rudin (2005) Bulgarian Clitics as K^0 heads, in F. Steven, F. Gladney, and M. Tasseva-Kurktchieva (eds.) *Proceedings of FASL 13: The South Carolina Meeting.* Ann Arbor: Michigan Slavic Publications, 104–116.

Frasson, Alberto (2021) Clitics are not enough: On agreement and null subjects in Brazilian Venetan. *Glossa: A Journal of General Linguistics*, 6(1), 86.

Frasson, Alberto (2022a) *The syntax of subject pronouns in heritage languages: Innovation and complexification*. Amsterdam: LOT.

Frasson, Alberto (2022b) Structural change in relative clauses and the autonomy of heritage languages. *Languages*, 7(2), 106.

Frasson, Alberto, Roberta D'Alessandro, and Brechje van Osch (2021) Null subjects in contact: A case study. *Heritage Language Journal*, 18, 1–36.

Friedmann, Naama and Rama Novogrodsky (2004) The acquisition of relative clause comprehension in Hebrew: A study of SLI and normal development. *Journal of Child Language*, 31, 661–668.

Frosi, Vitalina Maria and Mioranza Ciro (1983) *Dialetos Italianos. Um perfil lingüístico dos Ítalo-Brasileiros do nordeste do Rio Grande do Sul*. Caxias do Sul: EDUCS.

Fuchs, Zuzanna, Maria Polinsky, and Gregory Scontras (2024) Explaining gender: Lessons from heritage Spanish, in K. Grohmann (ed.) *Multifaceted multilingualism*. Amsterdam: John Benjamins, 292–326.

Fukuda, Shin (2012) Aspectual verbs as functional heads: Evidence from Japanese aspectual verbs. *Natural Language and Linguistic Theory*, 30, 965–1026.

Givón, Talmy (2009) *The genesis of syntactic complexity: Diachrony, ontogeny, neurocognition, evolution*. Amsterdam/Philadelphia: John Benjamins.

Gould, Isaac (2017) *Choosing a grammar: Learning paths and ambiguous evidence in the acquisition of syntax*. Amsterdam: John Benjamins.

Green, David W. and Jubin Abutalebi (2013) Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25(5), 515–530.

Greenberg, Joseph (1963) Some universals of grammar with particular reference to the order of meaningful elements, in J. Greenberg (ed.) *Universals of language*. Cambridge, MA: MIT Press, 58–90.

Grestenberger, Laura (2018) Deponency in non-finite contexts. *Language*, 94, 487–526.

Grey, Christina (2020) *The acquisition of transitivity alternations by bilingual children: A comparative study*. PhD thesis, Humboldt University of Berlin.

Grimstad, Maren Berg, Terje Lohndal, and Tor A. Åfarli (2014) Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd*, 41, 213–237.

Grimstad, Maren Berg, Brita Ramsevik Riksem, Terje Lohndal, and Tor A. Åfarli (2018) Lexicalist vs. exoskeletal approaches to language mixing. *The Linguistic Review*, 35, 187–218.

Grosjean, François (2008) *Studying bilinguals*. Oxford: Oxford University Press.

Gundel, Jeanette K. (1988) *The role of topic and comment in linguistic theory*. New York: Garland.

Gundel, Jeanette K. and Thorstein Fretheim (2004) Topic and focus, in L. Horn and G. Ward (eds.) *Handbook of pragmatics*. Oxford: Blackwell Publishing, 175–196.

Gürel, Ayşe (2004) Selectivity in L2-induced attrition: A psycholinguistic account. *Journal of Neurolinguistics*, 17, 53–78.

Haider, Hubert (2010) *The syntax of German*. Cambridge: Cambridge University Press.

Haider, Hubert and Luka Szucsich (2022) Slavic languages – "SVO" languages without SVO qualities? *Theoretical Linguistics*, 48(1–2), 1–39.

Håkansson, Gisela (1995) Syntax and morphology in language attrition: A study of five bilingual expatriate Swedes. *International Journal of Applied Linguistics*, 5, 153–171.

Halliday, Michael Alexander Kirkwood (1978) *Language as social semiotic*. London: Edward Arnold.

Halpern, Aaron (1995) *On the placement and morphology of clitics*. Stanford: Center for the Study of Language and Information.

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

Han, Chung-hye, Julien Musolino, and Jeffrey Lidz (2016) Endogenous sources of variation in language acquisition. *PNAS*, 113, 942–947.

Harbour, Daniel (2016) *Impossible persons*. Cambridge, MA: The MIT Press. DOI: 10.7551/mitpress/9780262034739.001.0001

Haugen, Einar (1953) *The Norwegian language in America: A study in bilingual behavior*. Vol. I–II. Pennsylvania: University of Pennsylvania Press.

Haugen, Einar (1976) *The Scandinavian languages: An introduction to their history*. London: Faber.

Hejná, Michaela and George Walkden (2022) *A history of English*. Berlin: Language Science Press. DOI: 10.5281/zenodo.6560337

Hickey, Raymond (ed.) (2020) *The handbook of language contact*. London: Wiley Blackwell.

Higginbotham, James (2009) *Tense, aspect, and indexicality*. Oxford: Oxford University Press.

Hjelde, Arnstein (1992) *Trøndsk talemål i Amerika*. Trondheim: Tapir.

Holmberg, Anders (2005) Is there a little pro? Evidence from Finnish. *Linguistic Inquiry*, 36–4; 533–564.

Hoot, Bradley and Tania Leal (2023) Resilience and vulnerability of discourse-conditioned word order in heritage Spanish. *Applied Psycholinguistics*, 44, 668–698.

Hopp, Holger, Michael T. Putnam, and Nora Vosburg (2019) Derivational complexity vs. transfer effects: Longdistance *wh*-movement in heritage and L2 grammars. *Linguistic Approaches to Bilingualism*, 9, 341–375.

Horne, Abraham Reeser (1905) *Pennsylvania German manual*. Allentown, PA: T.K. Horne.

Hu, Shenai, Anna Gavarró, Mirta Vernice, and Maria-Teresa Guasti (2016) The acquisition of Chinese relative clauses: Contrasting two theoretical approaches. *Journal of Child Language*, 43, 1–21.

Huang, C.-T. James (1984) On the distribution and reference of empty pronouns. *Linguistic Inquiry*, 15(4), 531–574.

Huettig, Falk and Robert J. Hartsuiker (2010) Listening to yourself is like listening to others: External, but not internal, verbal self-monitoring is based on speech perception. *Language and Cognitive Processes*, 25, 347–374.

Huffines, Marion Lois (1986) The function of aspect in Pennsylvania German and the impact of English. *Yearbook for German-American Studies*, 21, 137–154.

Huffines, Marion Lois (1990) Contact phenomena in language maintenance and shift: The Pennsylvania German infinitive construction. *American Journal of Germanic Linguistics and Literatures*, 2(2), 95–108.

Hulsey, Sarah and Uli Sauerland (2006) Sorting out relative clauses. *Natural Language Semantics*, 14, 111.

Iemmolo, Giorgio (2009) La marcatura differenziale dell'oggetto in siciliano antico. *Archivio Glottologico Italiano*, 94(2), 185–225.

Iemmolo, Giorgio (2010) Topicality and Differential Object Marking: Evidence from Romance and beyond. *Studies in Language*, 34(2), 239–272. DOI: 10.1075/sl.34.2.01iem

Ionin, Tania, Maria Goldshtein, Tatiana Luchkina, and Sofya Styrina (2023) Who did what to whom, and what did we already know? Word order and information structure in heritage and L2 Russian. *Linguistic Approaches to Bilingualism*, 13(3), 343–371.

Ionin, Tania and Tatiana Luchkina (2024) Linguistic knowledge and proficiency assessment in heritage and L2 Russian: Classroom implications of experimental findings, in O. Kisselev, O. Laleko, and I. Dubinina (eds.) *Russian as a heritage language: From research to classroom application*. Abingdon & New York: Routledge, 25–45.

Isurin, Ludmila and Tanya Ivanova-Sullivan (2008) Lost in between: The case of Russian Heritage speakers. *Heritage Language Journal*, 6(1), 72–104.

Ivanova-Sullivan, Tanya (2014) *Theoretical and experimental aspects of syntax-discourse interface in heritage grammars*. Leiden: Brill.

Ivanova-Sullivan, Tanya (2019) Acquisition of clitics in Heritage Bulgarian in the USA: Initial variability and reduction of optionality. *Linguistique Balkanique*, 58(1), 19–28.

Jack, George (1988) The origins of the English gerund. *NOWELE*, 12, 15–75.

Jaeggli, Osvaldo (1982) *Topics in Romance syntax*. Dordrecht: Foris.

Jany, Carmen (2008) Relativization versus nominalization strategies in Chimariko, in J. Kiester and V. Muñoz-Ledo (eds.) *Proceedings from the Eleventh Workshop on American Indigenous Languages. Santa Barbara Papers in Linguistics*, 19, 40–50.

Jasinskaja, Katja (2016) Information structure in Slavic, in C. Féry and S. Ishihara (eds.) *The Oxford handbook of information structure*. Oxford: Oxford University Press, 709–732.

Jia, Ruiting and Johanne Paradis (2018) The acquisition of relative clauses by Mandarin heritage language children. *Linguistic Approaches to Bilingualism*, 10, 153–183.

Johannessen, Janne Bondi (2015a) The Corpus of American Norwegian Speech (CANS), in B. Megyesi (ed.) *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)*. Linköping: NEALT Proceedings Series 23, 297–300.

Johannessen, Janne Bondi (2015b) Attrition in an American Norwegian heritage speaker?, in J. B. Johannessen and J. C. Salmons (eds.) *Germanic heritage languages in North America*. Amsterdam: John Benjamins, 46–71.

Johannessen, Janne Bondi and Ida Larsson (2015) Complexity matters: On gender agreement in heritage Scandinavian. *Frontiers in Psychology*, 6: 1842. DOI: 10.3389/fpsyg.2015.01842

Jourdain, Morgane, Emmanuelle Canut, and Karen Lahousse (2020) Item-based acquisition of dislocation in early child French. *Journal of Monolingual and Bilingual Speech*, 2(1): 44–72. https://doi.org/10.1558/jmbs.15679.

Julien, Marit (2002) Determiners and word order in Scandinavian DPs. *Studia Linguistica*, 56, 246–314.

Julien, Marit (2005) *Nominal phrases from a Scandinavian perspective*. Amsterdam: John Benjamins.

Junghanns, Uwe and Gerhild Zybatow (1997) Syntax and information structure of Russian clauses, in Wayles E. Browne (ed.) *Proceedings of the Annual Workshop on Formal Approaches to Slavic Linguistics (FASL): The Cornell Meeting*. Ann Arbor, MI: Michigan Slavic Publications, 289–319.

Kallestinova, Elena (2007) *Aspects of word order in Russian*. PhD dissertation, University of Iowa.

Kallulli, Dalina (2019) Balkan clitic doubling revisited: Micro-variation, typological generalizations, and a true Universal, in I. Krapova and B. Joseph (eds.) *Balkan syntax and (universal) principles of grammar*. Berlin/Boston: De Gruyter Mouton, 192–217.

Kaltsa, Maria, Ianthi Tsimpli, and Jason Rothman (2015) Exploring the source of differences and similarities in L1 attrition and heritage speaker competence: Evidence from pronominal resolution. *Lingua*, 164, 266–288.

Kato, Mary A. and Eduardo P. Raposo (2005) Obje(c)tos e artigos nulos. Similaridades e diferenças entre o português europeu e o português brasileiro, in M. Denilda Moura (ed.) *Reflexões sobre a sintaxe do português*. Maceió: Edufal, 73–96.

Kauhanen, Henri, Sarah Einhaus, and George Walkden (2023) Language structure is influenced by the proportion of non-native speakers: A reply to Koplenig (2019). *Journal of Language Evolution*, 20, 1–12. DOI: 10.1093/jole/lzad005

Kayne, Richard S. (1991) Romance clitics, verb movement, and PRO. *Linguistic Inquiry*, 22, 647–686.

Kayne, Richard S. (1993) Toward a modular theory of auxiliary selection. *Studia Linguistica*, 47, 3–31. https://doi.org/10.1111/j.1467-9582.1993.tb00837.x

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

Keating, Gregory, Jill Jegerski, and Bill VanPatten (2016) Online processing of subject pronouns in monolingual and heritage bilingual speakers of Mexican Spanish. *Bilingualism: Language and Cognition*, 19, 36–49.

Keating, Gregory, Bill VanPatten, and Jill Jegerski (2011) Who was walking on the beach? Anaphora resolution in Spanish heritage speakers and adult second language learners. *Studies in Second Language Acquisition*, 33, 193–221. https://doi.org/10.1017/S0272263110000732

Keenan, Edward L. and Bernard Comrie (1977) Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8, 63–99.

Kehayov, Petar (2017) *The fate of mood and modality in language death: Evidence from Minor Finnic*. Berlin/Boston: De Gruyter Mouton.

Keiser, Steven Hartman (2012) *Pennsylvania German in the American Midwest*. Durham, NC: Duke University Press.

van Kemenade, Ans and Nigel Vincent (eds.) (1997) *Parameters of morphosyntactic change*. Cambridge: Cambridge University Press.

Kerswill, Paul and Ann Williams (2002) "Salience" as an explanatory factor in language change: Evidence from dialect levelling in urban England, in M. C. Jones and E. Esch (eds.) *Language change: The interplay of internal external and extra-linguistic factors*. Berlin/New York: Mouton de Gruyter, 81–110.

Kester, Ellen-Petra (1996) *The nature of adjectival inflection*. Doctoral dissertation, University of Utrecht.

Kidd, Evan (ed.) (2011) *The acquisition of relative clauses: Processing, typology and function*. Amsterdam: John Benjamins.

Kim, Ji-Hye, Silvina Montrul, and James Yoon (2009) Binding interpretation of anaphors by Korean heritage speakers. *Language Acquisition*, 16, 3–35.

Kim, Ji-Hye, Silvina Montrul, and James Yoon (2010) Dominant language influence in acquisition and attrition of binding: Interpretation of the Korean reflexive *caki*. *Bilingualism: Language and Cognition*, 13, 73–84.

King, Tracy Holloway (1995) *Configuring topic and focus in Russian*. Stanford, CA: CSLI Publications.

Kondrashova, Natalija Ju (1996) The syntax of existential quantification. Doctoral Dissertation. University of Wisconsin-Madison.

Kovtunova, Irina I. (1976) Sovremennyj russkij jazyk. Porjadok slov i aktual'noe členenie predloženija [Modern Russian: Word order and the communicative structure of a sentence]. Moscow: Prosveščenie.

Kroll, Judith F. and Tamar H. Gollan (2014) Speech planning in two languages: What bilinguals tell us about language production, in M. Goldrick, V. Ferreira, and M. Miozzo (eds.) *The Oxford handbook of language production*. Oxford: Oxford University Press, 165–181.

Krylova, Olga A. and Serafima A. Khavronina (1986) *Poryadok slov v russkom jazyke [Word order in Russian]*. Moscow: Russkij Jazyk.

Kučerová, Ivona (2012) Grammatical marking of givenness. *Natural Language Semantics*, 20(1), 1–30.

Kupisch, Tanja (2014) Adjective placement in simultaneous bilinguals (German–Italian) and the concept of cross-linguistic overcorrection. *Bilingualism: Language and Cognition*, 17(1), 222–233.

Kupisch, Tanja (2019) 2L1 simultaneous bilinguals as heritage speakers, in M. S. Schmid and B. Köpke (eds.) *The Oxford handbook of language attrition*. Oxford: Oxford University Press, 458–469.

Kupisch, Tanja, Merete Anderssen, Ute Bohnacker, and Neil Snape (2009) Article acquisition in English, German, Norwegian, and Swedish, in R. P. Leow, H. Campos, and D. Lardiere (eds.) *Little words: Their history phonology, syntax, semantics, pragmatics, and acquisition*. Washington, DC: Georgetown University Press, 223–235.

Kupisch, Tanja and Maria Polinsky (2022) Language history on fast forward: Innovations in heritage languages and diachronic change. *Bilingualism: Language and Cognition*, 25(1), 1–12. DOI: 10.1017/S1366728921000997

Kupisch, Tanja and Jason Rothman (2018) Terminology matters! Why difference is not incompleteness and how early child bilinguals are heritage speakers. *International Journal of Bilingualism*, 22, 564–582.

Lakshmanan, Usha (2000) The acquisition of relative clauses by Tamil children. *Journal of Child Language*, 27, 587–617.

Laleko, Oksana (2011) Restructuring of verbal aspect in heritage Russian: Beyond lexicalization. *International Journal of Language Studies*, 5(3), 13–26.

Laleko, Oksana (2015) From privative to equipollent: Incipient changes in the aspectual system of heritage Russian, in G. Zybatow, P. Biskup, M. Guhl, C. Hurtig, O. Mueller-Reichau, and M. Yastrebova (eds.) *Slavic grammar from a formal perspective*. Frankfurt am Main: Peter Lang, 273–286.

Laleko, Oksana (2021) Discourse and information structure in heritage languages, in S. Montrul and M. Polinsky (eds.) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, 691–727.

Laleko, Oksana (2022) Word order and information structure in heritage and L2 Russian: Focus and unaccusativity effects in subject inversion. *International Journal of Bilingualism*, 26(6), 749–766.

Laleko, Oksana (2024) The complexity of word order change in a flexible system: On stability and variation in Heritage Russian word order, in M. Polinsky and M. Putnam (eds.) *Formal approaches to complexity in heritage language grammars*. Berlin: Language Science Press, 73–99.

Laleko, Oksana and Irina Dubinina (2018) Word order production in Heritage Russian: Perspectives from linguistics and pedagogy, in S. Bauckus and S. Kresin (eds.) *Connecting across languages and cultures: A heritage language Festschrift in Honor of Olga Kagan*. Bloomington, IN: Slavica, 191–215.

Laleko, Oksana and Olesya Kisselev (2023) Embracing linguistic variation in shift ecologies: The dominant language varies too. *Linguistic Approaches to Bilingualism*, 13(1), 60–64.

Laleko, Oksana and Yana Miroshnychenko (2022) Grammars in contact: A linguistic study of Russian in Brighton Beach, New York. *Russian Language Journal*, 72, 7–27.

Laleko, Oksana and Maria Polinsky (2017) Silence is difficult: On missing elements in bilingual grammars. *Zeitschrift für Sprachwissenschaft*, *36*, 135–163.

Laleko, Oksana and Gregory Scontras (2021) On the many dimensions of complexity in heritage languages. *Heritage Language Journal*, 18(2), 1–37.

Lambrecht, Knud (1994) *Information structure and sentence form. Topic, focus, and the mental representation of discourse referents*. Cambridge: Cambridge University Press.

Landau, Idan (2013) *Control in generative grammar: A research companion*. Cambridge: Cambridge University Press.

Lardiere, Donna (2008) Feature-assembly in second language acquisition, in J. Liceras, H. Zobl, and Helen Goodluck (eds.) *The role of formal features in second language acquisition*. New York: Lawrence Erlbaum Associates, 106–140.

Lasnik, Howard and Tim Stowell (1991) Weakest crossover. *Linguistic Inquiry*, 22, 687–720.

Lau, Elaine and Nozomi Tanaka (2021) The subject advantage in relative clauses: A review. *Glossa: A Journal of General Linguistics*, 6(1), 34.

Ledgeway, Adam (2009) *Grammatica diacronica del napoletano*. Tübingen: Max Niemeyer Verlag. https://doi.org/10.1515/9783484971288

Ledgeway, Adam (2012) *From Latin to Romance: Configurationality, functional categories and head-marking*. Oxford/New York: Oxford University Press.

Ledgeway, Adam and Alessandra Lombardi (2005) Verb movement, adverbs and clitic positions in Romance. *Probus*, 17, 79–113.

Ledgeway, Adam and John C. Smith (2016) Deixis, in A. Ledgeway and M. Maiden (eds.) *The Oxford Guide to the Romance languages*. Oxford/New York: Oxford University Press, 879–896. DOI: 10.1093/acprof:oso/9780199677108.003.0054

Ledgeway, Adam, Norma Schifano, and Silvestri Giuseppina (2019) Differential object marking and the properties of D in the dialects of the extreme south of Italy. *Glossa: A Journal of General Linguistics*, 4(1), 51. https://doi.org/10.5334/gjgl.569

Lee-Schoenfeld, Vera (2007) *Beyond coherence: The syntactic opacity of German*. Amsterdam: John Benjamins.

Lehmann, Christian (1984) *Der Relativsatz. Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik*. Tübingen: G. Narr.

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

Leivada, Evelina (2020) Mid-level generalizations of generative linguistics: Experimental robustness, cognitive underpinnings and the interdisciplinarity paradox. *Zeitschrift fürSprachwissenschaft*, 39(3), 357–374.

Leivada, Evelina and Elliot Murphy (2021) Mind the (teminological) gap: 10 misused, ambiguous, or polysemous terms in linguistics. *Ampersand*, 8, 1–8.

Lepschy, Giulio (1984) Costruzioni impersonali con SE in veneziano, in M. Cortelazzo (ed.) *Guida ai dialetti veneti VI*. Padua: CLEUP, 69–79.

Levine, Robert D. (2012) Auxilaries: *To*'s company. *Journal of Linguistics*, 48(1), 187–203.

Levinson, Stephen C. (2018) Introduction. Demonstratives: Patterns in diversity, in S. C. Levinson, S. Cutfield, M. J. Dunn, N. J. Enfield, and Sérgio Meira (eds.) *Demonstratives in cross-linguistic perspective*. Cambridge: Cambridge University Press, 1–42. DOI: 10.1017/9781108333818.002

Lidz, Jeffrey and Annie Gagliardi (2015) How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics*, 1(1), 333–353.

Lightfoot, David W. (2020) *Born to parse: How children select their languages*. Cambridge, MA: MIT Press.

Lipski, John M. (1985) Creole Spanish and vestigial Spanish: Evolutionary parallels. *Linguistics*, 23, 963–984.

Lipski, John M. (1996) Patterns of pronominal evolution in Cuban-American bilinguals, in A. Roca and J. B. Jensen (eds.) *Spanish in contact: Issues in bilingualism*. Somerville, MA: Cascadilla, 159–186.

Lipski, John M. (1999) Creoloid phenomena in the Spanish of transitional bilinguals, in A. Roca and J. Lipski (eds.) *Spanish in the United States: Linguistic contact and diversity*. Berlin: Mounton de Gruyter, 155–182.

Lobo, Maria and Carolina Silva (2016) Ambiguidade pronominal em orações adverbiais do português europeu: Crianças vs. adultos. *Revista da Associação Portuguesa de Linguística*, 2, 319–338.

Lødrup, Helge (2011) Hvor mange genus er det i Oslo-dialekten? *Maal og Minne*, 2, 120–136.

Lohndal, Terje (2013) Generative grammar and language mixing. *Theoretical Linguistics*, 39(3–4), 215–224.

Lohndal, Terje (2020) Predicting outcomes in heritage grammars. *Bilingualism: Language and Cognition*, 23(11), 31–32. https://doi.org/10.1017/S1366728919000403

Lohndal, Terje (2021) Syntax of heritage languages, in S. Montrul and M. Polinsky (eds.) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, 644–667.

Lohndal, Terje and Michael T. Putnam (2021) The tale of two lexicons: Decomposing complexity across a distributed lexicon. *Heritage Language Journal*, 18(2), 1–29.

Lohndal, Terje and Michael T. Putnam (2024) Expanding structures while reducing mappings: Morphosyntactic complexity in aggultinating heritage languages, in M. Polinsky and M. T. Putnam (eds.) *Formal approaches to complexity in heritage language grammars*. Berlin: Language Science Press, 101–132.

Lohndal, Terje and Marit Westergaard (2016) Grammatical gender in American Norwegian heritage language: Stability or attrition? *Frontiers in Psychology*, 7, 344. DOI: 10.3389/fpsyg.2016.00344

Lohndal, Terje and Marit Westergaard (2021) Grammatical gender: Acquisition, attrition, and change. *Journal of Germanic Linguistics*, 33, 95–121.

López, Luis (2012) *Indefinite objects: Scrambling, choice functions, and differential marking*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/9165.001.0001

López, Luis (2020) *Bilingual grammar: Toward an integrated model*. Cambridge: Cambridge University Press.

Loporcaro, Michele (2007) On triple auxiliation in Romance. *Linguistics*, 44(1), 173–222. DOI: 10.1515/LING.2007.005

Loporcaro, Michele (2016) Auxiliary selection and participial agreement, in A. Ledgeway and M. Maiden (eds.) *The Oxford guide to the Romance languages*. Oxford: Oxford University Press, 802–818.

Loporcaro, Michele and Maria Teresa Vigolo (1995) Ricerche sintattiche sul confine veneto-trentino in Valsugana. l'Accordo del participio passato, in E. Banfi, G. Bonfadini, P. Cordin, and M. Iliescu (eds.) *Italia settentrionale. Crocevia di idiomi romanzi*. Tübingen: Niemeyer, 87–101.

Louden, Mark L. (1988) *Bilingualism and syntactic change in Pennsylvania German* (Unpublished doctoral dissertation). Cornell University

Louden, Mark L. (1997) Linguistic structure and sociolinguistic identity in Pennsylvania German society, in J. R. Dow and M. Wolff (eds.) *Languages and lives: Essays in honour of Werner Enninger*. Bern: Peter Lang, 79–92.

Louden, Mark L. (2005) Grundzüge der Pennsylvaniadeutschen Satzstruktur, in E. Eggers, J. E. Schmidt, and D. Stellmacher (eds.) *Moderne Dialekte - Neue Dialektologie*. Stuttgart: Franz Steiner Verlag, 253–265.

Louden, Mark L. (2016a) *Infinitivkonstruktionen im Pennsylvaniadeutschen*. Presentation at Deutscher Sprachatlas, Marburg.

Louden, Mark L. (2016b) *Pennsylvania Dutch: The story of an American language*. Baltimore: Johns Hopkins University.

Louden, Mark L. (2019) The English 'infusion' in Pennsylvania German, in R. Hickey (ed.) *English in the German-speaking world*. Cambridge: Cambridge University Press, 384–407.

Lovoll, Odd S. (1999) *The promise of America: A history of the Norwegian-American people*. Minneapolis: University of Minnesota Press.

Luchkina, Tatiana and Jennifer S. Cole (2021) Perception of word-level prominence in free word order language discourse. *Language and Speech*, 64(2), 381–412.

Luján, Marta and Claudia Parodi (1996) Clitic-doubling and the acquisition of agreement in Spanish, in J. Gutiérrez-Rexach and L. S. Villar (eds.) *Perspectives on Spanish linguistics*. Vol. 1. Los Angeles, CA: UCLA, Department of Linguistics: UCLA, Department of Linguistics, 119–138.

Luk, Gigi and Ellen Bialystok (2013) Bilingualism is not a categorial variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology*, 25, 605–621.

Luk, Gigi and Jason Rothman (2022) Experience-based individual differences modulate language, mind, and brain outcomes in multilinguals. *Brain and Language*, 228, 105107.

Lykke, Alexander K. (2020) *Variation and change in the tense morphology of heritage Norwegian in North America*. Doctoral dissertation, University of Oslo.

Łyskawa, Paulina and Naomi Nagy (2019) Case marking variation in heritage Slavic languages in Toronto: Not so different. *Language Learning*, 70(SI), 122–156. https://doi.org/10.1111/lang.12348

Lyons, John (1977) *Semantics*. vol. 2. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511620614

Manzini, M. Rita and Leonardo M. Savoia (2000) The syntax of object clitics: Si in Italian dialects, in G. Salvi and G. Cinque (eds.) *Current studies in Italian syntax: Essays offered to Lorenzo Renzi*. Amsterdam: Elsevier, 225–255.

Manzini, M. Rita and Leonardo M. Savoia (2005) *I dialetti italiani e romanci. Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.

Marcato, Gianna and Flavia Ursini (1998) *Dialetti veneti. Grammatica e storia*. Padua: Unipress.

Marinis, Theo and Duygu Özge (2020) How do bilingual children acquire complex syntax in their heritage vs. the majority language: Turkish-English speaking children in the UK. Presentation in the series "The Multilingual Mind: Lecture series on multilingualism across disciplines."

Mathesius, Vilém (1947) *Čeština a obecný jazykozpyt [Czech and general linguistics]*. Prague: Melantrich.

Matsumoto, Yoshiko (1989) *Grammar and semantics of adnominal clauses in Japanese*. Ph.D. dissertation, Berkeley.

Matushansky, Ora (2006) Head movement in linguistic theory. *Linguistic Inquiry*, 37(1), 69–109.

Mayer, Elisabeth and Liliana Sánchez (2021) Emerging DOM patterns in clitic doubling and dislocated structures in Peruvian-Spanish contact varieties, in J. Kabatek, P. Obrist, and A. Wall (eds.) *Differential object marking in Romance*. Berlin/Boston, MA: De Gruyter, 103–137.

McCloskey, James (1990) Resumptive pronouns, A-bar binding and levels of representation in Irish, in R. Hendrick (ed.) *Syntax of the modern celtic languages*. San Diego: Academic Press, 199–248.

McCloskey, James (2017) Resumption, in M. Everaert and H.C. Riemsdijk (eds.) *The Wiley Blackwell companion to syntax*. 2nd ed. Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118358733.wbsyncom105

McWhorter, John (2018) *The creole debate*. Cambridge: Cambridge University Press.

Meir, Natalia and Sharon Armon-Lotem (2017) Independent and combined effects of socioeconomic (SES) and bilingualism on children's vocabulary and verbal short-term memory. *Frontiers in Psychology*, 8, 1442.

Meir, Natalia and Bibi Janssen (2021) Child heritage language development: An interplay between cross-linguistic influence and language-external factors. *Frontiers in Psychology*, 12, 651730. https://doi.org/10.3389/fpsyg.2021.651730

Meisel, Jürgen (2011) *First and second language acquisition: Parallels and differences*. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511862694

Meisel, Jürgen (2013) Sensitive phases in successive language acquisition: The critical period hypothesis revisited, in C. Boeckx and K. K. Grohmann (eds.) *The Cambridge handbook of biolinguistics*. Cambridge: Cambridge University Press. 69–85.

Miller, Jim and Regina Weinert (1998) *Spontaneous spoken language: Syntax and discourse*. Oxford: Oxford University Press.

Miyagawa, Shigeru (2005) On the EPP, in M. McGinnis and N. Richards (eds.) *Perspectives on phases*. Cambridge, MA: MIT Working Papers in Linguistics, 49, 201–236.

Momma, Shota and Colin Phillips (2018) The relationship between parsing and generation. *Annual Review of Linguistics*, 4, 233–254.

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

Montrul, Silvina (2004) Subject and object expression in Spanish heritage speakers: A case of morphosyntactic convergence. *Bilingualism: Language and Cognition*, 7(2), 125–142.

Montrul, Silvina (2008) *Incomplete acquisition in bilingualism: Re-examining the age factor*. Amsterdam: John Benjamins.

Montrul, Silvina (2016) *The acquisition of heritage languages*. Cambridge: Cambridge University Press.

Montrul, Silvina, Rakesh Bhatt, and Roxana Girju (2015) Differential object marking in Spanish, Hindi and Romanian as heritage languages. *Language*, 91, 564–610.

Montrul, Silvina and Melissa Bowles (2009) Back to basics: Differential object marking under incomplete acquisition in Spanish heritage speakers. *Bilingualism: Language and Cognition*, 12(3), 363–383. DOI: 10.1017/S1366728909990071

Montrul, Silvina and Maria Polinsky (2011) Why not heritage speakers? *Linguistic Approaches to Bilingualism*, 1, 58–62.

Montrul, Silvina and Maria Polinsky (2019) Introduction to heritage language development, in M. S. Schmid and B. Köpke (eds.) *The Oxford handbook of language attrition*. Oxford: Oxford University Press, 418–433.

Montrul, Silvina and Maria Polinsky (eds.) (2021) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press.

Montrul, Silvina, Rebecca Foote, and Silvia Perpiñán (2008) Gender agreement in adult second language learners and Spanish heritage speakers: The effects of age and context of acquisition. *Language Learning*, 58, 503–553.

Montrul, Silvina and Noelia Sánchez-Walker (2013) Differential object marking in child and adult Spanish heritage speakers. *Language Acquisition*, 20(2), 109–132. DOI: 10.1080/10489223.2013.766741

Morgado, Sara (2013) Qual a importância da informação semântica na resolução de sujeitos pronominais no processamento offline? *Textos Selecionados do XXVIII Encontro Nacional da Associação Portuguesa de Linguística*, 493–507.

Moro, Andrea (2008) *The boundaries of Babel: The brain and the enigma of impossible languages*. Cambridge, MA: MIT Press.

Mufwene, Salikoko (2020) Grammaticization in the development of creoles, in R. Vossen and C. König (eds.) *Patterns of linguistic convergence in Africa, a Festschrift for Bernd Heine*. Köln: Rüdiger Köppe, 153–169.

Müller, Natascha (2017) Different sources of delay and acceleration in early child bilingualism. *Zeitschrift Für Sprachwissenschaft*, 36, 7–30. DOI: 10.1515/zfs-2017-0002

Müller, Natascha and Aafke Hulk (2001) Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4, 1–21.

Myers-Scotton, Carol (2002) *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.

Nagy, Noemi (2016) A sociolinguistic view of null subjects and VOT in Toronto heritage languages. *Lingua*, 164, 309–327. https://doi.org/10.1016/j.lingua.2014.04.012

Natvig, David (2022) The Great Change and the shift from Norwegian to English in Ulen, Minnesota, in J. R. Brown (ed.) *The verticalization model of language shift: The Great change in American communities*. Oxford: Oxford University Press, 85–113.

Neeleman, Ad and Tanya Reinhart (1998) Scrambling and the PF interface, in M. Butt and W. Geuder (eds.) *The projection of arguments: Lexical and compositional factors*. Stanford, CA: CSLI Publications, 309–353.

Neeleman, Ad and Kriszta Szendrői (2004) Superman sentences. *Linguistic Inquiry*, 35, 149–159.

van Ness, Silke (1994) Pennsylvania German, in E. König and J. van der Auwera (eds.) *The Germanic languages*. London: Routledge, 420–438.

Nevins, Andrew (2022) *When minoritized languages change linguistic theory*. Cambridge: Cambridge University Press.

Nishigauchi, Taisuke (1986) *Quantification in syntax (logical form, Japanese, wh-constructions)*. Ph.D. Dissertation, University of Massachussetts at Amherst.

Novogrodsky, Rama and Naama Friedmann (2006) The production of relative clauses in syntactic SLI: A window to the nature of the impairment. *Advances in Speech-Language Pathology*, 8, 364–375.

Nunes, Jairo (2004) *Linearization of chains and sideward movement*. Cambridge, MA: MIT Press.

O'Grady, William, Miseon Lee, and Miho Choo (2001) The acquisition of relative clauses by heritage and non-heritage learners of Korean as a second language: A comparative study. *Journal of Korean Language Education*, 12, 283–294.

Oikonomou, Despina and Artemis Alexiadou (2022) Voice syncretism crosslinguistically: The view from minimalism. *Philosophies*, 7(1), 19.

Otheguy, Ricardo, Ana Celia Zentella, and David Livert (2007) Language and dialect contact in Spanish in New York: Toward the formation of a speech community. *Language*, 83, 770–803.

Ouhalla, Jamal (2004) Semitic Relatives. Linguistic Inquiry, 35(2), 288–300. http://www.jstor.org/stable/4179275

Paradis, Johanne (2011) Internal and external factors in child second language acquisition. *Linguistic Approaches to Bilingualism*, 1, 213–237. DOI: 10.1075/lab.1.3.01par

Paradis, Johanne and Fred Genesee (1996) Syntactic acquisition in bilingual children. *Studies in Second Language Acquisition*, 18, 1–25. DOI: 10.1017/S0272263100014662

Pashkova, Tatiana, Wintai Tsehaye, Shanley Allen, and Rosemarie Tracy (2022) Syntactic optionality in heritage language use: Clause type preferences of German heritage speakers in a majority English context. *Heritage Language Journal*, 19(1), 1–41.

Paspali, Anastasia, Fotini Karkaletsou, and Artemis Alexiadou (2022) When morphology is not enough: Structural representations in the acquisition and processing of voice in monolingual and bilingual children. *Studies in Greek Linguistics*, 41, 185–198.

Perez-Cortes, Silvia, Michael T. Putnam, and Liliana Sánchez (2019) Differential access: Asymmetries in accessing features and building representations in heritage language grammars. *Languages*, 4(4), 81.

Pérez-Leroux, Ana Teresa (1993) *Empty categories and the acquisition of wh-movement.* Amherst, MA: GLSA.

Pérez-Leroux, Ana Teresa, Alejandro Cuza, and Danielle Thomas (2011) Clitic placement in Spanish–English bilingual children. *Bilingualism: Language and Cognition*, 14(2), 221–232.

Pérez-Leroux, Ana Teresa, Mihaela Pirvulescu, and Yves Roberge (2017) *Direct objects and language acquisition*. Cambridge: Cambridge University Press. http://doi.org/10.1017/9781139086264

Pescarini, Diego (2021) *Object clitics*. Oxford: Oxford University Press.

Philippaki-Warburton, Irene (1998) Functional categories and modern Greek syntax. *The Linguistic Review* 15(2–3), 159–186.

Pintzuk, Susan, George Tsoulas, and Anthony Warner (eds.) (2000) *Diachronic syntax: Models and mechanisms*. Oxford: Oxford University Press.

Pires, Acrisio (2007) *The minimalist syntax of defective domains: Gerunds and infinitives*. Amsterdam: John Benjamins.

Pires, Acrisio and Jason Rothman (2007) Disentangling sources of incomplete acquisition: An explanation for competence divergence across heritage grammars. *International Journal of Bilingualism*, 13(2), 211–238.

Poletto, Cecilia (1993) *La sintassi del soggetto nei dialetti italiani settentrionali*. Padova: Unipress.

Poletto, Cecilia (2000) *The higher functional field: Evidence from northern Italian dialects*. Oxford: Oxford University Press.

Polinsky, Maria (1995) American Russian: Language loss meets language acquisition, in W. Brown, E. Dornisch, N. Kondrashova, and D. Zec (eds.) *Formal approaches to Slavic linguistics*. Ann Arbor, MI: Michigan Slavic Publications, 370–407.

Polinsky, Maria (2000) A composite linguistic profile of a speaker of Russian in the USA, in O. Kagan and B. Rifkin (eds.) *The learning and teaching of Slavic languages and cultures*. Bloomington: Slavica, 437–466.

Polinsky, Maria (2005) Word class distinctions in an incomplete grammar, in D. Ravid and H. Bat-Zeev Shylgkrot (eds.) *Perspectives on language and language development*. Dordrecht: Kluwer Academic Press, 419–436.

Polinsky, Maria (2006) Incomplete acquisition: American Russian. *Journal of Slavic Linguistics*, 14(2), 191–262.

Polinsky, Maria (2011) Reanalysis in adult heritage language: New evidence in support of sttrition. *Studies in Second Language Acquisition*. 33(2): 305–328. DOI: 10.1017/S027226311000077X

Polinsky, Maria (2016) Structure vs. use in heritage language. *Linguistics Vanguard*, 2(1), 1–14. https://doi.org/10.1515/lingvan-2015-0036

Polinsky, Maria (2018) *Heritage languages and their speakers*. Cambridge: Cambridge University Press. DOI: 10.1017/9781107252349

Polinsky, Maria and Olga Kagan (2007) Heritage languages: In the 'wild' and in the classroom. *Language and Linguistics Compass*, 1(5), 368–395.

Polinsky, Maria and Michael T. Putnam (eds.) (2024) *Formal approaches to complexity in heritage languages*. Berlin: Language Science Press.

Polinsky, Maria and Gregory Scontras (2020) Understanding heritage languages. *Bilingualism: Language and Cognition*, 23(1), 4–20. https://doi.org/10.1017/S1366728919000245

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

Postal, Paul (1993) Remarks on weak crossover effects. *Linguistic Inquiry*, 24, 539–556.

Postma, Gertjan (2014) A Pomeraian language island in Brazil: New syntactic data and some considerations on 'weak T.' *Language Typology and Universals*, 67(4), 627–663.

Postma, Gertjan (2018) *A constrastive grammar of Brazilian Pomeranian*. Amsterdam: John Benjamins.

van Pottelberge, Jeroen (2004) *Der am-Progressive. Struktur und parallele Entwicklung in den kontinentalwestgermanischen Sprachen*. Tübingen: Gunter Narr.

van Pottelberge, Jeroen (2007) Defining grammatical constructions as a linguistic sign: The case of periphrastic progressives in the Germanic languages. *Folia Linguistica*, 41, 99–134.

Poynton, Cate (1985) *Language and gender: Making the difference*. Geelong, Victoria: Deakin University Press.

Prévost, Philippe (2009) *The acquisition of French: The development of inflectional morphology and syntax in L1 acquisition, bilingualism, and L2 acquisition*. Amsterdam: John Benjamins.

Putnam, Michael T. (2019) The (in)stability of grammars. *Studies in Second Language Acquisition*, 41(2), 275–278.

Putnam, Michael T. (2020a) One feature – one head: Features as functional heads in language acquisition and attrition, in P. G. Fuentes and C. Suárez-Gómez (eds.) *New trends in language acquisition within the generative perspective*. Dordrecht: Springer, 1–23.

Putnam, Michael T. (2020b) Separating vs. shrinking. *Bilingualism: Language and Cognition*, 23(1), 41–42.

Putnam, Michael T. and Birna Arnbjörnsdótir (2015) Minimizing (interface) domains: The loss of longdistance binding in North American Icelandic, in B. R. Page and M. T. Putnam

(eds.) *Moribund Germanic heritage langauges in North America: Theoretical perspectives and empirical findings*. Leiden: Brill, 203–223.

Putnam, Michael T., Matthew Carlson, and David Reitter (2018) Integrated, not isolated: Defining typological proximity in an integrated multilingual architecture. *Frontiers in Psychology*, 8, 2212.

Putnam, Michael T. and Andrew D. Hoffman (2021) Overextension in Gottscheerisch (negative) imperatives: Proclisis at the edge of the first phase. *Journal of Comparative Germanic Linguistics*, 24, 185–219.

Putnam, Michael T., Silvia Perez-Cortes, and Liliana Sánchez (2019) 3 language attrition and the feature reassembly hypothesis, in M. S. Schmid and B. Köpke (eds.) *The Oxford handbook of language attrition*. Oxford: Oxford University Press, 18–24.

Putnam, Michael T. and Maike H. Rocker (2019) Aspectualizers and complementation in Pennsylvania Dutch: The case of schtaerte, in K. Biers and R. J. Brown (eds.) *WILA 9 Proceedings*. Somerville, MA: Cascadilla Press, 20–27.

Putnam, Michael T. and Liliana Sánchez (2013) What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism*, 3(4), 478–508.

Putnam, Michael T., Lara Schwarz, and Andrew D. Hoffman (2021) Morphology of heritage languages, in S. Montrul and M. Polinsky (eds.) *The Cambridge handbook of heritage languages and linguistics*. Cambridge: Cambridge University Press, 613–643.

Putnam, Michael T. and Åshild Søfteland (2022) Mismatches at the syntax-semantics interface: The case of non-finite complementation in American Norwegian. *Nordic Journal of Linguistics*, 45, 310–347.

Putnam, Michael T. and Åshild Søfteland (2024) When covert modality sneaks into your grammar: *Wh*-infinitives in American Norwegian. *Second Language Research*, 40(1), 171–194.

Ramchand, Gillian and Peter Svenonius (2002) The lexical syntax and lexical semantics of the verb-particle construction, in L. Mikkelsen and C. Potts (eds.) *WCCFL 21 Proceedings*. Somerville, MA: Cascadilla Press, 387–400.

Raña-Risso, Rocío and Carolina Barrera-Tobón (2018) On the relationship between subject placement and overt pronouns in the Spanish of New York City bilinguals. *Journal of Language Contact*, 11(2), 324–347.

Raposo, Eduardo (2004) Objetos nulos e CLLD. Uma teoria unificada. *Revista da ABRALIN*, 3(1–2), 41–73.

Reinhart, Tanya (1995) *Interface strategies*. Ms., Utrecht University.

Reinhart, Tanya (1998) Wh-in-situ in the framework of the minimalist program. *Natural Language Semantics*, 6, 29–56.

Reiss, Charles (in press) Research methods in armchair linguistics, in G. Dupre, R. Nefdt, and K. Stanton (eds.) *Oxford handbook of philosophy of linguistics*. Oxford: Oxford University Press

Requena, Pablo (2023) Variation versus deviation: Early bilingual acquisition of Spanish differential object marking. *Linguistic Approaches to Bilingualism*, 13(6), 801–829. https://doi.org/10.1075/lab.21001.req

Riksem, Brita Ramsevik (2017) Language mixing and diachronic change: American Norwegian noun phrases then and now. *Languages*, 2, 29. DOI: 10.3390/languages2020003

Riksem, Brita Ramsevik (2018) *Language mixing in American Norwegian noun phrases: An exoskeletal analysis of synchronic and diachronic patterns*. Doctoral dissertation, Norwegian University of Science and Technology (NTNU).

Riksem, Brita Ramsevik, Maren Berg Grimstad, Terje Lohndal, and Tor A. Åfarli (2019) Language mixing within verbs and nouns in American Norwegian. *Journal of Comparative Germanic Linguistics*, 22, 189–209.

Riksem, Brita Ramsevik, Terje Lohndal, and Tor A. Åfarli (2021) Adjektivkongruens i amerikanorsk, in K. Hagen, G. Kristoffersen, and Øystein A. Vangsnes (eds.) *Språk i arkiva. Ny forsking om eldre talemål frå LIA-prosjektet*. Oslo: Novus, 273–293.

Rinke, Esther (in press) Animacy restrictions on null objects in European Portuguese, in P. Barbosa, C. Flores, A. C. Silva, and E.-M. Roessler (eds.) *Null objects from a crosslinguistic and developmental perspective*. Amsterdam: Benjamins.

Rinke, Esther and Cristina Flores (2014) Morphosyntactic knowledge of Portuguese heritage bilinguals. *Bilingualism: Language and Cognition*, 17, 681–699.

Rinke, Esther and Cristina Flores (2018) Another look at the interpretation of overt and null pronominal subjects in bilingual language acquisition: Heritage Portuguese in contact with German and Spanish. *Glossa: A Journal of General Linguistics*, 3(1), 68. https://doi.org/10.5334/gjgl.535

Rinke, Esther and Cristina Flores (2021) Portuguese as heritage language in Germany – a linguistic perspective, in T. Kupisch (ed.) *Languages: Special issue on heritage languages in Germany*, 6(1), 10.

Rinke, Esther, Cristina Flores, and Pilar Barbosa (2018) Null objects in the spontaneous speech of monolingual and bilingual speakers of European Portuguese. *Probus*, 30(1), 93–119. https://doi.org/10.1515/probus-2017-0004

Rinke, Esther, Cristina Flores, and Aldona Sopata (2019) Heritage Portuguese and heritage Polish in contact with German: More evidence on the production of objects. *Languages*, 4(3), 53. https://doi.org/10.3390/languages4030053

Rinke, Esther, Cristina Flores, and Jacopo Torregrossa (2024) How different types of complexity can account for difficult structures in bilingual and monolingual language acquisition, in M. Polinsky and M. Putnam (eds.) *Formal approaches to complexity in Heritage Grammars*. Berlin: Language Science Press, 43–71.

Rizzi, Luigi (1986a) On the status of subject clitics in Romance, in O. A. Jaeggli and C. Silva-Corvalan (eds.) *Studies in Romance linguistics*. Dordrecht: Foris, 391–419.

Rizzi, Luigi (1986b) Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*, 17, 501–557.

Rizzi, Luigi (1997) The fine structure of the left periphery, in H. Liliane (ed.) *Elements of grammar: Handbook in generative syntax*. Dordrecht: Springer, 281–337. DOI: 10.1007/978-94-011-5420-8_7

Roberts, Ian (2009) A deletion analysis of null subjects, in T. Biberauer, A. Holmberg, I. Roberts, and M. Sheehan (eds.) *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press, 88–124.

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

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

Rodina, Yulia and Marit Westergaard (2015) Grammatical gender in Norwegian: Language acquisition and language change. *Journal of Germanic Linguistics*, 27, 145–187.

Rødvand, Linn Iren (2017) *Empirical investigations of grammatical gender in American heritage Norwegian*. Master's thesis, University of Oslo.

Rødvand, Linn Iren (2018) Systematic variation in the gender system in American Norwegian, in J. H. Petersen and K. Kühl (eds.) *Selected Proceedings of the 8th Workshop on Immigrant Languages in the Americas (WILA 8)*. Somerville, MA: Cascadilla Proceedings Project, 88–95.

Romanowski, Piotr (2021) *Family language policy in the Polish diaspora: A focus on Australia*. Abingdon/New York: Routledge.

Rothman, Jason (2010) On the typological economy of syntactic transfer: Word order and relative clause high/low attachment preference in L3 Brazilian Portuguese. *International Review of Applied Linguistics in Teaching (IRAL)*, 48(2/3), 245–273.

Rothman, Jason (2011) L3 Syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, 27(1), 107–127. DOI: 10.1177/0267658310386439

Rothman, Jason (2013a) Cognitive economy, non-redundancy and typological primacy in L3 acquisition: Evidence from initial stages of L3 Romance, in S. Baauw, F. Dirjkoningen and M. Pinto (eds.) *Romance languages and linguistic theory 2011*. Amsterdam: John Benjamins, 217–248.

Rothman, Jason (2013b) Linguistic and cognitive motivations for the Typological Primacy Model (TPM) of third language (L3) transfer: Timing of acquisition and proficiency considered. *Bilingualism: Language And Cognition*, 18(2), 179–190.

Rothman, Jason and Jennifer Cabrelli Amaro (2010) What variables condition syntactic transfer? A look at the L3 initial state. *Second Language Research*, 26(2), 189–218. DOI: 10.1177/0267658309349410

Rothman, Jason, Jorge González Alonso, and Eloi Puig-Mayenco (2019) *Third language acquisition and linguistic transfer*. Cambridge: Cambridge University Press.

Rothman, Jason and Jeanine Treffers-Daller (2014) A prolegomenon to the construct of the native speaker: Heritage speaker bilinguals are natives too! *Applied Linguistics*, 35, 93–98.

Rothman, Jason, Fatih Bayram, Vincent DeLuca, Grazia Di Pisa, Jon A. Duñabeitia, Khadij Gharibi, … , Stefanie Wulff (2022) Monolingual comparative normativity in bilingualism research is out of "control": Arguments and alternatives. *Applied Psycholinguistics*, First View: 1–14. DOI: 10.1017/S0142716422000315

Ruys, Eddy G. (2000) Weak crossover as a scope phenomenon. *Linguistic Inquiry*, 31, 513–539.

Saab, Andrés (2018) *Case conflicts and A-movement: DOM, doubling and clitics in Rioplatense Spanish*. Lecture series held at Utrecht University (December).

Safir, Ken (1984) Multiple variable binding. *Linguistic Inquiry*, 15, 603–638.

Safir, Ken (1986) Relative clauses in a theory of binding and levels. *Linguistic Inquiry*, 17, 663–689.

Safir, Ken (1996) Derivation, representation, and resumption: The domain of weak crossover. *Linguistic Inquiry*, 27, 313–339.

Safir, Ken (2017) Weak crossover, in M. Everaert and H. C. van Riemsdijk (eds.) *The Wiley Blackwell companion to syntax*. 2nd ed. Malden, MA: Wiley-Blackwell.

Sánchez-Walker, Noelia (2012) Comprehension of subject and object relative clauses in Spanish heritage speakers and L2 learners of Spanish. Poster presented at the Sixth Heritage Language Institute, UCLA, June 2012.

Santos, Ana Lúcia (2017) Alguns aspetos da aquisição de orações subordinadas completivas, in M. J. Freitas and A. L. Santos (eds.) *A aquisição de língua materna e não materna*. Berlin: Language Science Press, 249–273.

Sasse, Hans-Jürgen (1992) Language decay and contact-induced change: Similarities and differences, in M. Brenzinger (ed.) *Language death: Factual and theoretical explorations with special reference to East Africa*. Berlin: Mouton de Gruyter, 59–79.

Sauder, Ben (1955) *Der nachbar an de schtroas*. St. Jacobs: Pennsylvania German Folklore Society of Ontario.

Schäfer, Florian (2008) *The syntax of anti-causatives*. Amsterdam: John Benjamins.

Schwartz, Bonnie D. and Rex A. Sprouse (1996) L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12(1), 40–72.

Scontras, Gregory, Zuzanna Fuchs, and Maria Polinsky (2015) Heritage language and linguistic theory. *Frontiers in Psychology*, 6, 1545.

Scontras, Gregory, Maria Polinsky, and Zusanne Fuchs (2018) In support of representational economy: Agreement in heritage Spanish. *Glossa*, 3(1), 1. https://doi.org/10.5334/gjgl.164

Scontras, Gregory and Michael T. Putnam (2020) Lesser-studied heritage languages: An appeal to the dyad. *Heritage Language Journal*, 17(2), 152–155. DOI: 10.46538/hlj.17.2.2

Scontras, Gregory, Cheng-Yu Edwin Tsai, Kenneth Mai, and Maria Polinsky (2014) Chinese scope: An experimental investigation, in U. Etxeberria et al. (eds.) *Proceedings of Sinn und Bedeutung* 18. University of the Basque Country, 396–414.

Seaman, David (1972) *Modern Greek and American English in contact*. The Hague: Mouton De Gruyter.

Sekerina, Irina A. (1997) *The syntax and processing of split scrambling constructions in Russian*. Doctoral Dissertation, City University of New York.

Sekerina, Irina A. (2003) Scrambling processing: Dependencies, complexity, and constraints, in S. Karimi (ed.) *Word order and scrambling*. Malden, MA: Blackwell, 301–324.

Serratrice, Ludovica (2007) Cross-linguistic influence in the interpretation of anaphoric and cataphoric pronouns in English–Italian bilingual children. *Bilingualism: Language and Cognition*, 10, 225–238.

Serratrice, Ludovica, Antontella Sorace, and Sandra Paoli (2004) Crosslinguistic influence at the syntax–pragmatics interface: Subjects and objects in English–Italian bilingual and monolingual acquisition. *Bilingualism: Language and Cognition*, 7, 183–205.

Sevcenco, Anca, Larisa Avram, and Ioana Stoiescu (2013) Subject and object relative clause production in child Romanian, in L. Avram and A. Sevcenco (eds.) *Topics in language acquisition and child language learning in a Romanian cpontext*. Bucharest: Editura Universitatii din Bucuresti, 23–57.

Sheehan, Michelle (2006) *The EPP and null subjects in Romance*. Newcastle University Doctoral Dissertation.

Sherkina-Lieber, Marina. (2015) Tense, aspect, and agreement in heritage Labrador Inuttitut: Do receptive bilinguals understand functional morphology? *Linguistic Approaches to Bilingualism*, 5, 30–61.

Shin, Naomi, Alejandro Cuza, and Liliana Sánchez (2022) Structured variation, language experience, and crosslinguistic influence shape child heritage speakers' Spanish direct objects. *Bilingualism: Language and Cognition*, 1–13. https://doi.org/10.1017/S1366728922000694

Siewierska, Anna and Ludmila Uhlířová (2010) An overview of word order in Slavic languages, in A. Siewierska (ed.) *Constituent order in the languages of Europe*. Berlin: Mouton De Gruyter, 105–150.

Silva-Corvalán, Carmen (1994) *Language contact and change: Spanish in Los Angeles*. Oxford: Clarendon.

Šimík, Radek and Marta Wierzba (2017) Expression of information structure in West Slavic: Modeling the impact of prosodic and word-order factors. *Language*, 93(3), 671–709.

Silverstein, Michael (1976) Hierarchy of features and ergativity, in R. W. M. Dixon (ed.) *Grammatical categories in Austrialian languages*. Canberra: Australian Institute of Aboriginal Studies, 112–171.

Slabakova, Roumyana (2009) Features or parameters: Which one makes second language acquisition easier, and more interesting to study? *Second Language Research*, 25(2), 313–324.

Slabakova, Roumyana, Tania Leal, Amber Dudley, and Micah Stack (2020) *Generative second language acquisition*. Cambridge: Cambridge University Press.

Slioussar, Natalia (2007) *Grammar and information structure: A study with reference to Russian*. Doctoral Dissertation, Utrecht University.

Smith, Helena (2015) A brief history of German-Greek relations. *The Guardian*. Available online: https://www.theguardian.com/world/2015/mar/23/brief-history-german-greek-relations (accessed November 7, 2022).

Soares, Carla (1998) *As Categorias Funcionais no Processo de Aquisição do Português Europeu (Estudo longitudinal da Produção Espontânea de uma Criança de 1;2.0 aos 2;2.17 anos)*. Master's dissertation: Universidade de Lisboa, Portugal.

Sorace, Antonella (1995) Acquiring argument structures in a second language: The unaccusative/ unergative distinction, in L. Eubank, L. Selinker, and M. S. Smith (eds.) *The current state of interlanguage*. Amsterdam: John Benjamins, 153–175.

Sorace, Antonella (2011) Pinning down the concept of interface in bilingualism. *Linguistic Approaches to Bilingualism*, 1(1), 1–33. DOI 10.1075/lab.1.1.01sor issn 1879–9264/e-issn 1879–9272

Sorace, Antonella and Francesca Filiaci (2006) Anaphora resolution in near-native speakers of Italian. *Second Language Research*, 22(3), 339–368.

Sorace, Antonella, Ludovica Serratrice, Francesca Filiaci, and Michela Baldo (2009) Discourse conditions on subject pronoun realization: Testing the linguistic intuitions of older bilingual children. *Lingua*, 119(3), 460–477. DOI: 10.1016/j.lingua.2008.09.008

Sorgini, Luana (2022) *One way for accusative case: Theme and experiencer objects*. Paper presented at the Syntax Interface Lectures, Utrecht University.

Stawinski, Alberto V. (1987) *Dicionário do dialeto vêneto sul-rio-grandense–português*. Porto Alegre: EST S. Lourenço de Brindes; Caxias do Sul: EDUCS.

Stepanov, Arthur (2021) Experimental syntax and Slavic languages, in G. Goodall (ed.) *The Cambridge handbook of experimental syntax*. Cambridge University Press, 534–560.

Svetozarova, Natalia (1998) Intonation in Russian, in D. Hirst and A. Di Christo (eds.) *Intonation systems: A survey of twenty languages*. Cambridge University Press, 261–274.

Szwedek, Aleksander (2011) *The thematic structure of the sentence in English and Polish: Sentence stress and word order*. Frankfurt am Main: Peter Lang.

Taraldsen, Knut Tarald (1990) D-projections and N-projections in Norwegian, in J. Mascaró and M. Nespor (eds.) *Grammar in progress*. Dordrecht: Foris, 419–431.

Terenghi, Silvia (2021) (Non-)monotonicity effects in the person domain. Talk presented at the 2021 Annual Meeting of the Linguistic Association of Great Britain (LAGB), online, 6–9 September 2021.

Terenghi, Silvia (2022) Demonstrative systems are not affected by contact: Evidence from heritage southern Italo-Romance. *Languages*, 7(3), 201. https://doi.org/10.3390/languages7030201

Terenghi, Silvia (2023a) *Missing person: Structure and change in Romance demonstratives*. Utrecht University Doctoral Dissertation. https://www.lotpublications.nl/Documents/635_fulltext.pdf

Terenghi, Silvia (2023b) Last in, first out: Patterns of reduction in Romance demonstrative systems. *Journal of Historical Syntax*, 7(11), 1–40. https://doi.org/10.18148/hs/2023.v7i6-19.157

Thompson, Sandra A., Joseph Sung-Yul Park, and Charles N. Li (2006) *A reference grammar of Wappo*. Berkeley: University of California Press.

Titov, Elena (2012) *Information structure of argument order alternations*. Doctoral Dissertation, University College London.

Toosarvandani, Maziar (2014) Two types of deverbal nominalization in Northern Paiute. *Language*, 90, 786–833.

Torrego, Esther. 1998. *The dependencies of objects*. Cambridge, MA: The MIT Press. https://doi.org/10.7551/mitpress/2337.001.0001

Torregrossa, Jacopo, Cristina Flores, and Esther Rinke (2023) What modulates the acquisition of difficult structures in a heritage language? A study on Portuguese in contact with French, German and Italian. *Bilingualism: Language and Cognition*, 26(1), 179–192. https://doi.org/10.1017/S1366728922000438

Tsai, C.-Y. Edwin, Gregory Scontras, Kenneth Mai, and Maria Polinsky (2014) Prohibiting inverse scope: An experimental study of Chinese vs. English. *Empirical Issues in Syntax and Semantics*, 10, 305–322.

Tsimpli, Ianthi (2006) The acquisition of voice and transitivity alternations in Greek as native and second language, in S. Unsworth, T. Parodi, A. Sorace, and M. Young-Scholten (eds.) *Paths of development in L1 and L2 acquisition: In honor of Bonnie D. Schwartz*. Amsterdam: John Benjamins, 15–55.

Tsimpli, Ianthi (2014) Early, late or very late? *Linguist Approaches Biling*, 4, 283–313. DOI: 10.1075/lab.4.3.01tsi

Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock, and Francesca Filiaci (2004) First language attrition and syntactic subjects: A study of Greak and Italian near native speakers of English. *International Journal of Bilingualism*, 8(3), 257–277.

Tsimpli, Ianthi, Antonella Sorace, Caroline Heycock, Francesca Filiaci, and Maria Bouba (2003) Subjects in L1 attrition: Evidence from Greek and Italian near-native speakers of English. *BUCLD 27. Proceedings of the Annual Boston University Conference on Language Development*. Boston, MA: Cascadilla, 787–797.

Tuttle, Edward F. (1986) The spread of ESSE as universal auxiliary in Central Italo-Romance. *Medioevo Romanzo*, 11(2), 229–287.

Unsworth, Sharon (2013) Assessing age of onset effects in (early) child L2 acquisition. *Language Acquisition*, 20, 74–92. DOI: 10.1080/10489223.2013.766739

Unsworth, Sharon (2016) Early child L2 acquisition: Age or input effects? Neither, or both? *Journal of Child Language*, 43, 608–634. DOI: 10.1017/S030500091500080X

Unsworth, Sharon, Froso Argyri, Leonie Cornips, Aafke Hulk, Antonella Sorace, and Ianthi Tsimpli (2011) Bilingual acquisition of Greek voice morphology and Dutch gender: What do they have in common?, in N. Danis, K. Mesh, and H. Sung (eds.) *Proceedings of the 35th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press, 590–602.

Unsworth, Sharon, Froso Argyri, Leonie Cornips, Aafke Hulk, Antonella Sorace, and Ianthi Tsimpli (2014) The role of age of onset and input in early child bilingualism in Greek and Dutch. *Applied Psycholinguistics*, 35, 765–805. DOI: 10.1017/S0142716412000574

Uriagereka, Juan (1995) Aspects of the syntax of clitic placement in western romance. *Linguistic Inquiry*, 26(1), 79–123.

Vallduví, Enric and Engdahl, Elisabet (1995) The linguistic realization of information packaging. *Linguistics*, 34(3), 459–519.

Vangsnes, Øystein A. (1999) *The identification of functional architecture*. Doctoral dissertation, University of Bergen.

Varlokosta, Spyridoula, Adriana Belletti, João Costa, Naama Friedmann, Anna Gavarró, Kleanthes K. Grohmann, Maria Teresa Guasti, Laurice Tuller, Maria Lobo, Darinka Anđelković, Núria Argemí, Larisa Avram, Sanne Berends, Valentina Brunetto, Hélène Delage, María-José Ezeizabarrena, Iris Fattal, Ewa Haman, Angeliek van Hout, Kristine Jensen de López, Napoleon Katsos, Lana Kologranic, Nadezda Krstić, Jelena Kuvac Kraljevic, Aneta Miękisz, Michaela Nerantzini, Clara Queraltó, Zeljana Radic, Sílvia Ruiz, Uli Sauerland, Anca Sevcenco, Magdalena Smoczyńska, Eleni Theodorou, Heather van der Lely, Alma Veenstra, John Weston, Maya Yachini, and Kazuko Yatsushiro (2016) A cross-linguistic study of the acquisition of clitic and pronoun production. *Language Acquisition*, 23(1), 1–26. DOI: https://doi.org/10.1080/10489223.2015.1028628

de Vries, Mark (2002) *The syntax of relativization*. Ph.D. Dissertation, University of Amsterdam. LOT Dissertation Series 53.

de Vries, Mark (2018) Relative clauses in syntax. *Oxford Research Encyclopedia of Linguistics*. https://doi.org/10.1093/acrefore/9780199384655.013.56

van der Wal, Jenneke (2016) Diagnosing focus. *Studies in Language*, 40, 259–301.

Walkden, George and Anne Breitbarth (2019) Complexity as L2-difficulty: Implications for syntactic change. *Theoretical Linguistics*, 45(3–4), 183–209. DOI: 10.1515/tl-2019-0012

Wang, Qi, Diane Lillo-Martin, Catherine T. Best, and Andrea Levitt (1992) Null subject versus null object: Some evidence from the acquisition of Chinese and English. *Language Acquisition: A Journal of Developmental Linguistics*, 2(3), 221–254. DOI: https://doi.org/10.1207/s15327817la0203_2

Weil, Henri (1844) *De l'ordre des mots dans les langues anciennes compares aux langues modernes*. Paris: Didier Érudition.

Weir, Andrew (2012) Left-edge deletion in English and subject omission in diaries. *English Language and Linguistics*, 16(1), 105–129.

Weir, Andrew (2022) Fragments and left-edge ellipsis: The division of labour between syntax, semantics, and prosody, in G. Güneş and A. Lipták (eds.) *The derivational timing of ellipsis*. Oxford: Oxford University Press, 253–290.

Westergaard, Marit (2009a) *The acquisition of word order: Micro-cues, information structure, and economy*. Amsterdam: John Benjamins.

Westergaard, Marit (2009b) Item-based vs. rule-based learning: The acquisition of word order in wh-questions in English and Norwegian. *Journal of Child Language*, 36, 1023–1051.

Westergaard, Marit (2013) The acquisition of linguistic variation: Parameters vs. micro-cues, in T. Lohndal (ed.) *In search of universal grammar: From Old Norse to Zoque*. Amsterdam: John Benjamins, 275–298.

Westergaard, Marit and Tanja Kupisch (2020) Stable and vulnerable domains in Germanic heritage languages, in K. Hagen, A. Hjelde, K. Stjernholm, and Ø. A. Vangsnes (eds.) *Bauta: Janne Bondi Johannessen in memoriam*. Vol. 11. Oslo: Oslo Studies in Language, 503–526.

Westergaard, Marit and Terje Lohndal (2019) Verb second word order in Norwegian heritage language: Syntax and pragmatics, in D. Lightfoot and J. Havenhill (eds.) *Variable properties in language: Their nature and acquisition*. Washington, DC: Georgetown University Press, 91–102.

Westergaard, Marit, Terje Lohndal, and Artemis Alexiadou (2019) The asymmetric nature of V2: Evidence from learner languages, in K. R. Christensen, H. Jørgensen, and J. L. Wood (eds.) *The Sign of the V – Papers in Honour of Sten Vikner*. Aarhus: Department of English, School of Communication & Culture, Aarhus University, 709–733.

Westergaard, Marit, Terje Lohndal, and Björn Lundquist (2023) Variable V2 in Norwegian heritage language: An effect of crosslinguistic influence? *Linguistic Approaches to Bilingualism* 13(2), 133–162.

Wiese, Heike (2017) Language situations: A method for capturing variation within speakers' repertoires, in A. Yoshiyuki (ed.) *Methods in Dialectology XVI*. Bamberg Studies in English Linguistics. Frankfurt a. M.: Peter Lang, 105–117.

Wiese, Heike, Artemis Alexiadou, Shanley Allen, Oliver Bunk, Natalia Gagarina, Kateryna Iefremenko, Esther Jahns, Martin Klotz, Thomas Krause, Annika Labrenz, Anke Lüdeling, Maria Martynova, Katrin Neuhaus, Tatiana Pashkova, Vicky Rizou, Tracy Rosemarie, Christoph Schroeder, Luka Szucsich, Wintai Tsehaye, … Yulia Zuban (2021) RUEG Corpus (0.4.0) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.5808870

Wiese, Heike, Artemis Alexiadou, Shanley Allen, Oliver Bunk, Natalia Gagarina, Kateryna Iefremenko, Maria Martynova, Tatiana Pashkova, Vicky Rizou, Christoph Schröder, et al. (2022) Heritage speakers as part of the native language continuum. *Frontiers in Psychology*, 12, 717973.

Wilder, Christopher (1989) The syntax of German infinitives. Unpublished doctoral dissertation, University of London (UCL).

Wiltschko, Martina (2021) *The grammar of interactional language*. Cambridge: Cambridge University Press.

Witkoś, Jacek (2007) Polish and A-type scrambling, in Peter Kosta and Lilia Schuercks (eds.) *Linguistic investigations into Formal Description of Slavic Languages*. Frankfurt am Main: Peter Lang, 165–180.

Witkoś, Jacek (2008) On the correlation between A-type scrambling and lack of weak crossover effects. *Studia Anglica Posnaniensia: International Review of English Studies*, 44, 297–328.

Wurmbrand, Susanne (2001) *Infinitives: Restructuring and clause structure*. Berlin: Mouton de Gruyter.

Wurmbrand, Susanne (2014) Tense and aspect in English infinitives. *Linguistic Inquiry*, 45(3), 403–447.

Wurmbrand, Susanne and Christos Christopoulos (2020) Germanic infinitives, in M. T. Putnam and B. R. Page (eds.) *The Cambridge handbook of Germanic linguistics*. Cambridge: Cambridge University Press, 389–412.

Wurmbrand, Susanne and Magdalena Lohninger (2023) An implicational universal in complementation – theoretical insights and empirical progress, in J. M. Hartmann and A. Wöllstein (eds.) *Propositional Arguments in Cross-Linguistic Research: Theoretical and Empirical Issues*. Tübingen: Gunter Narr Verlag, 183–229.

Yager, Lisa, Nora Hellmold, Hyoun-A Joo, Michael T. Putnam, Eleonora Rossi, Catherine Stafford, and Joseph Salmons (2015) New structural patterns in moribund grammar: Case marking in Heritage German. *Frontiers in Psychology*, 6, 1716.

Yang, Charles (2016) *The price of linguistic productivity: How children learn to break the rules of language*. Cambridge, MA: MIT Press.

Yang, Charles (2018) A formalist perspective on language acquisition. *Linguistic Approaches to Bilingualism*, 8(6), 665–706.

Yokoyama, Olga T. (1986) *Discourse and word order*. Amsterdam: John Benjamins.

Zombolou, Katerina and Artemis Alexiadou (2014) The canonical function of the deponent verbs in Greek, in F. Rainer, F. Gardani, H. C. Luschützky, and W. U. Dressler (eds.) *Morphology and meaning: Selected papers from the 15th International Morphology Meeting, Vienna, February 2012*. Amsterdam: John Benjamins, 331–344.

Zombolou, Katerina and Artemis Alexiadou (2012) Incomplete acquisition of anticausatives in Heritage speakers of Spanish-Greek and English-Greek bilingual adults. Paper presented at the workshop *Variation in Language Acquisition*, Münster, February 10–11, 2012.

Zombolou, Katerina, Spyridoula Varlokosta, and Artemis Alexiadou (2010) Acquiring anticausatives versus passives in Greek, in K. Franich, K. M. Iserman, and L. L. Keil (eds.) *Proceedings of the 34th annual Boston University Conference on Language Development 2009*. vol. 2. Somerville, MA: Cascadilla Press, 515–524.

Zubizarreta, Maria Luisa (1998) *Prosody, focus, and word order*. Cambridge, MA: MIT Press.

Zwicky, Arnold (1977) *On clitics*. Bloomington: Indiana University Linguistics Club.

Index