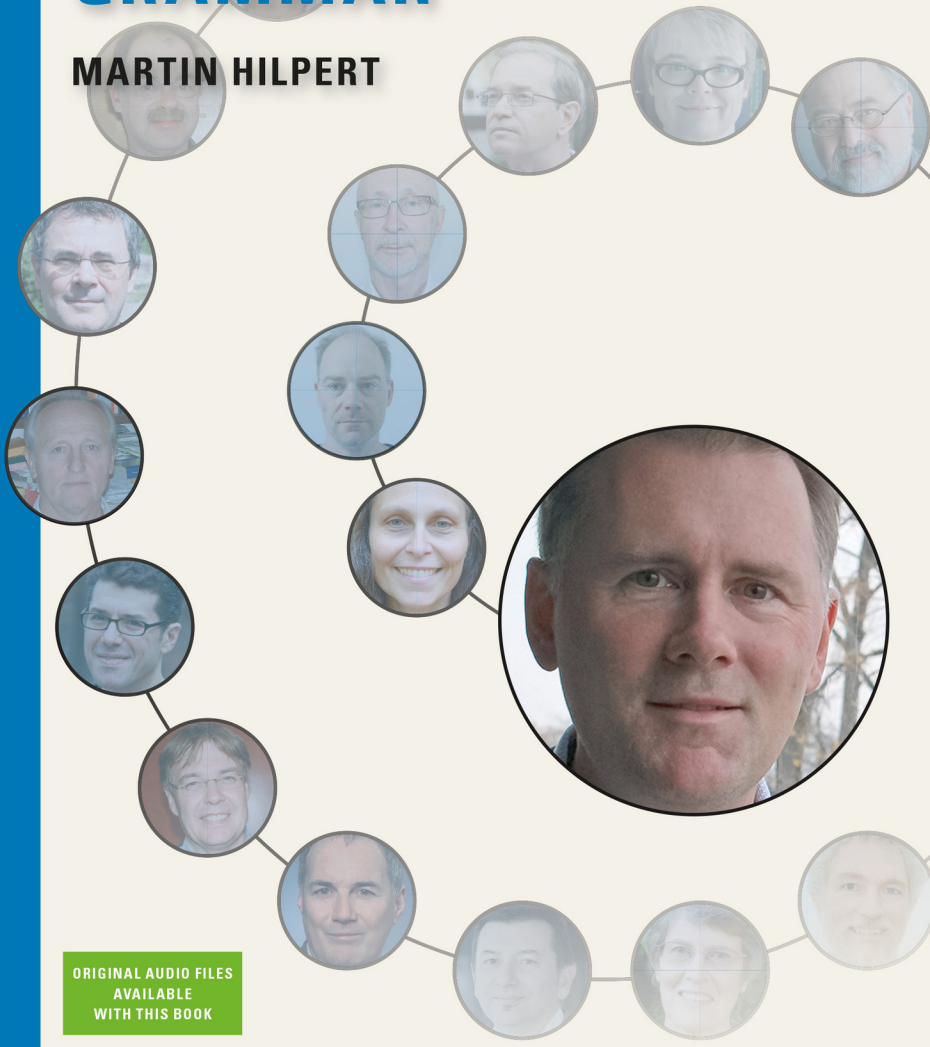
TEN LECTURES ON DIACHRONIC CONSTRUCTION GRAMMAR

MARTIN HILPERT



ORIGINAL AUDIO FILES AVAILABLE WITH THIS BOOK

DISTINGUISHED LECTURES IN COGNITIVE LINGUISTICS

BRILL

Ten Lectures on Diachronic Construction Grammar

Distinguished Lectures in Cognitive Linguistics

Distinguished Lectures in Cognitive Linguistics publishes the keynote lectures series given by prominent international scholars at the China International Forum on Cognitive Linguistics since 2004. Each volume contains the transcripts of 10 lectures under one theme given by an acknowledged expert on a subject and readers have access to the audio recordings of the lectures through links in the e-book and QR codes in the printed volume. This series provides a unique course on the broad subject of Cognitive Linguistics. Speakers include George Lakoff, Ronald Langacker, Leonard Talmy, Laura Janda, Dirk Geeraerts, Ewa Dąbrowska and many others.

The titles published in this series are listed at *brill.com/dlcl*

Ten Lectures on Diachronic Construction Grammar

By

Martin Hilpert

BRILL

LEIDEN | BOSTON

Ouvrage publié avec le soutien de la Faculté des lettres et sciences humaines de l'Université de Neuchâtel.
This work has been published with the support of the Faculty of Humanities at the University of Neuchâtel.

The Library of Congress Cataloging-in-Publication Data is available online at http://catalog.loc.gov

Typeface for the Latin, Greek, and Cyrillic scripts: "Brill". See and download: brill.com/brill-typeface.

ISSN 2468-4872
ISBN 978-90-04-44678-6 (hardback)
ISBN 978-90-04-44679-3 (e-book)

This book is printed on acid-free paper and produced in a sustainable manner.

Contents

Note on Supplementary Material

All original audio-recordings and other supplementary material, such as hand-outs and PowerPoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via a QR code for the print version of this book. In the e-book both the QR code and dynamic links will be available which can be accessed by a mouse-click.

The material can be accessed on figshare.com through a PC internet browser or via mobile devices such as a smartphone or tablet. To listen to the audio recording on hand-held devices, the QR code that appears at the beginning of each chapter should be scanned with a smart phone or tablet. A QR reader/scanner and audio player should be installed on these devices. Alternatively, for the e-book version, one can simply click on the QR code provided to be redirected to the appropriate website.

This book has been made with the intent that the book and the audio are both available and usable as separate entities. Both are complemented by the availability of the actual files of the presentations and material provided as hand-outs at the time these lectures were given. All rights and permission remain with the authors of the respective works, the audio-recording and supplementary material are made available in Open Access via a CC-BY-NC license and are reproduced with kind permission from the authors. The recordings are courtesy of the China International Forum on Cognitive Linguistics (http://cifcl.buaa.edu.cn/), funded by the Beihang University Grant for International Outstanding Scholars.

The complete collection of lectures by Martin Hilpert can be accessed by scanning this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.c.5289337.

Preface by the Series Editor

The present text, entitled Ten Lectures on Diachronic Construction Grammar by Martin Hilpert, is a transcribed version of the lectures given by Professor Hilpert in November 2019 as the forum speaker for the 19th China International Forum on Cognitive Linguistics. Martin Hilpert (PhD 2007) is professor of English Linguistics at the University of Neuchâtel (Switzerland). He holds a PhD from Rice University (USA) and did postdoctoral research at the International Computer Science Institute in Berkeley and at the Freiburg Institute for Advanced Studies. He is interested in cognitive linguistics, language change, construction grammar, and corpus linguistics.

He is the author of *Germanic Future Constructions: A Usage-based Approach to Language Change* (2008, John Benjamins), *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax* (2013, Cambridge University Press), and *Construction Grammar and its Application to English* (2014/2019, Edinburgh University Press). He is editor of the journal *Functions of Language* and associate editor of *Cognitive Linguistics*.

The China International Forum on Cognitive Linguistics (http://cifcl.buaa.edu.cn/) provides a forum for eminent international scholars to give lectures on their original contributions to the field of cognitive linguistics. It is a continuing program organized by several prestigious universities in Beijing. The following is a list of organizers for CIFCL 19.

Organizer:

Fuyin (Thomas) Li: PhD/Professor, Beihang University

Co-organizers:

Yihong Gao: PhD/Professor, Peking University
Baohui Shi: PhD/Professor, Beijing Forestry University
Yuan Gao: PhD/Professor, University of Chinese Academy of Sciences
Xu Zhang: PhD/Associate Professor, Beijing Language and Culture University

The text is published, accompanied by its audio disc counterpart, as one of the *Distinguished Lectures in Cognitive Linguistics*. The transcriptions of the video, proofreading of the text and publication of the work in its present book form have involved many people's strenuous efforts. The initial transcripts were

completed by Na Liu, Xiaoran Zhou, Quting Zhang, Jing Du, Mengmin Xu, Lin Yu, Guannan Zhao and Junjie Lu. Guannan Zhao and Lin Yu made revisions to the whole text. As the editors, we then made word-by-word and line-by-line revisions. To improve the readability of the text, we have deleted the false starts, repetitions, and fillers like *now, so, you know, OK, and so on, again, of course, if you like, sort of*, etc. Occasionally, the written version needs an additional word to be clear, a word that was not actually spoken in the lecture. We have added such words within double brackets [[...]]. To make the written version readable, even without watching the film, we've added a few "stage directions", in italics also within double brackets: [[...]]. These describe what the speaker was doing, such as pointing at a slide, showing an object, etc. Professor Hilpert made final revisions to the transcriptions and the published version is the final version approved by the speaker.

Thomas Fuyin Li
Beihang University
thomasli@buaa.edu.cn

Na Liu
liuna0317@foxmail.com

Preface by the Author

First and foremost, I want to express my sincere gratitude to prof. Fuyin (Thomas) Li for inviting me to present these ten lectures in Beijing in the winter of 2019. Coming to Beijing has been a wonderful and unforgettable experience. Meeting the participants of the Cognitive Linguistics Forum has been a great privilege. The ten lectures that I gave as part of the Forum were based on my own previous work as well as on research that has informed my general outlook on language, which is cognitive, usage-based, connected to corpus-based and experimental research methods, and indebted to the idea that many characteristics of language can be understood more fully once they are framed in a diachronic perspective. Understanding how language works in the here-and-now requires us to think about how it came to be that way, and conversely, examining the cognitive and social pressures that shape language use can help us understand why language changes diachronically in the way it does. The invitation to give the Beijing Forum lectures gave me the exciting opportunity to draw together the main lines of my research and to try and communicate the bigger picture that unites the individual studies. I thank the students and colleagues in the audience for their questions and their feedback, and I hope that the ideas presented in the lectures will help them in coming up with answers to the questions that arise in their own research, and in our common endeavor of understanding language and languages. Now that the lectures have been turned into a book, I hope the same for its readership.

I wholeheartedly thank the team of student volunteers who have been supporting the organization of the lectures, transcribing them, and helping to turn the transcripts into the present book: Na (Selina) Liu, Xiaoran (Kara) Zhou, Quting (Daisy) Zhang, Jing (Milly) Du, Mengmin (Amy) Xu, Lin (Joyce) Yu, Guannan (Vivian) Zhao and Junjie (Jim) Lu. Not only were they highly dedicated and successful in their efforts to make the lectures run smoothly, they were also kind and considerate in showing me around on campus and in the city, as well as helping me find my own way when I felt like it. Moreover, they were very knowledgeable, open and invigorating in their interactions with professors and speakers. I wish them all lots of success in their careers, and I look forward to seeing them again very soon.

I would like to acknowledge the generous funding from the Swiss National Science Foundation that was provided for research that is reported on in this book (SNF Grant 100015_149176/1, SNF Grant 100012L/169490/1).

Martin Hilpert
25 February, 2020

LECTURE 1

What Is Construction Grammar?

Many thanks for inviting me to this wonderful event. It is a great honor for me to be here and speak to you. I wish to express my sincere gratitude to Professor Thomas Li and to all volunteers who have been involved in the preparations for this meeting. I have never been to the China International Forum on Cognitive Linguistics (CIFCL) before, but I am certainly not a stranger to it. Since many of the recorded lectures are available on the internet, I have been able to download and listen to wonderful colleagues like Ewa Dąbrowska, Stefan Th. Gries, or Mark Turner. It is an immense privilege to be invited to follow in their footsteps, and I want you to know that I truly appreciate the honor.

What is this lecture series all about? First of all, the lectures will address the general issue of language change, but they will do so by adopting a perspective that differs from other approaches to historical linguistics. Specifically, I will draw on the theoretical framework of Construction Grammar, in order to explore how a constructional view can help us understand certain aspects of language change that other frameworks find difficult to explain. How can a constructional view help us make progress in the study of how languages develop over time?

In this context, it will be necessary to spell out how a constructional view differs in its assumptions from other theoretical approaches to language change. In Cognitive Linguistics, we have a long tradition of comparing our views to generative and formalist views, and often this goes along with a narrative of conflict between theories. That is not exactly what I will be after in these lectures. I think there are differences worth exploring between approaches that lie within the cognitive-functional enterprise. For example, how does Diachronic Construction Grammar differ in its assumptions from an approach such as grammaticalization theory? It is a worthwhile enterprise to spell out the assumptions and to discuss aspects that are problematic and at this stage unresolved.

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13690930.

In my talks, I also want to discuss methodological aspects. What methods can be used for the constructional study of language change? I will be discussing a range of different techniques, mostly corpus-based techniques, but also some experimental methods, that can be exploited within the project of Diachronic Construction Grammar. Of course, a large part of that will be a discussion of the results that we can obtain on the basis of these methods. What do we learn, at the end of the day, when we apply these methods?

My main goal for this lecture series is to give you a general overview of what Diachronic Construction Grammar tries to accomplish. That overview is offered from my personal perspective, which is necessarily subjective. Depending on who you ask, you might hear very different points of view on certain aspects of what I will discuss. As always in science, I would encourage you to adopt whatever you find useful from my discussion and engage critically with whatever you find unconvincing or worthy of criticism. In the ten lectures I will draw on ideas about Diachronic Construction Grammar and language more generally, that I have tried to express in three different books that I have written.

My book *Germanic Future Constructions* (Hilpert 2008) reflects the way I think about constructions and their associations with lexical items, as well as how these patterns of associations may shift over time. In the lectures, that idea will inform the discussion in several places. In *Constructional Change in English* (Hilpert 2013), I tried to develop a general account of how language change can be understood from a constructional point of view. I also discussed a variety of different corpus-based methods for the analysis of constructional change.

The third book, which has just come out in a second edition, is *Construction Grammar and Its Application to English* (Hilpert 2014/2019). That book is my attempt to summarize in general terms what Construction Grammar is all about and what sets it apart from other linguistic theories. While the ten lectures will take these works as a backdrop, I also need to point out that my thinking on these issues has not developed in a vacuum. There are a number of colleagues who I have been lucky enough to collaborate with and whose ideas will therefore inform the discussion. I will make reference to research that I have been doing together with Stefan Th. Gries, with Florent Perek, with David Correia Saavedra, and with Susanne Flach. My views on Diachronic Construction Grammar have also been strongly influenced by other people that you know, Adele Goldberg, Elizabeth Traugott, Graeme Trousdale, and Holger Diessel. In these lectures, I will sometimes point out issues where I think we disagree, but in principle our views are very similar, so there is a lot more that we have in common than what we see differently. But even so, I think it is always useful to talk about the issues where you actually disagree.

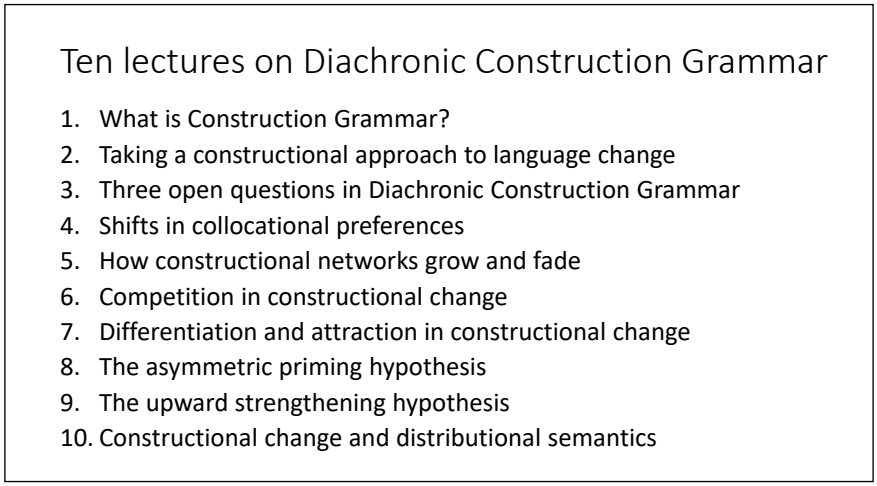


FIGURE 1

Coming now to the structure of the lecture series, one thing I would like to say is that I wanted these lectures to be accessible to all of you, regardless of your background. I assume some familiarity with basic notions in Cognitive Linguistics, but not much more. This means that some of you will undoubtedly recognize ideas that are quite familiar to you, but I promise to contextualize them in such a way that hopefully, I will make you see them in a new way. We'll start today with two lectures that will provide the theoretical basis for everything else that I have to say.

In this first lecture, which simply has the title *What is Construction Grammar?*, I will try to clarify my assumptions and define my theoretical terms. The lecture for this afternoon is entitled *Taking a construction approach to language change*. In that lecture, we will move into the subject of language change. Once these ideas are in place, the third lecture, on *Three open questions in Diachronic Construction Grammar*, will take us to a different area of debate, where current research has not yet found a consensus. The first three lectures here can be seen as the theoretical groundwork for the rest of the lectures.

Lectures 4 to 7 will exemplify different ways of dealing with the open questions and with the debates in Diachronic Construction Grammar. Lecture 4 will focus on the relation between constructions and lexical items, and more specifically, shifts in the collocational preferences of constructions, and what we can learn from those shifts.

Lecture 5 takes up the idea of a constructional network, and it will discuss how constructional networks may develop over time. How do these networks grow? How do they branch out and how do they fade away in the end?

Lecture 6 addresses the notion of competition in constructional change. Some constructions are mutual alternatives, so that they can be seen as competing with one another. Much research in sociolinguistics actually draws on the metaphor of competition between linguistic forms, so I will explore what this idea implies for constructional theories of constructional change.

Lecture 7 will look at two different trajectories of constructional change, namely differentiation, constructions becoming more different, and attraction, constructions becoming more similar to one another. I will look at the reasons for differentiation and attraction and how developments of this kind can be studied.

Lecture 8 and 9 will bring us back into the realm of linguistic theory. The big question that they both try to address is why languages change in the way they do. Lecture 8 addresses what I call the asymmetric priming hypothesis, which tries to explain the unidirectionality of semantic change in grammaticalization. Lecture 9, on the upward strengthening hypothesis, addresses the emergence and entrenchment of grammatical markers from a constructional perspective.

In the final lecture, I will discuss recent connections between work in Diachronic Construction Grammar and corpus-linguistic work on what is called distributional semantics. I will close with a discussion of how theory and methodology can come together in useful ways.

With this broad overview in place, let me come to the first lecture – What is Construction Grammar? A very short answer to that question is that Construction Grammar is a theory of what speakers know when they know

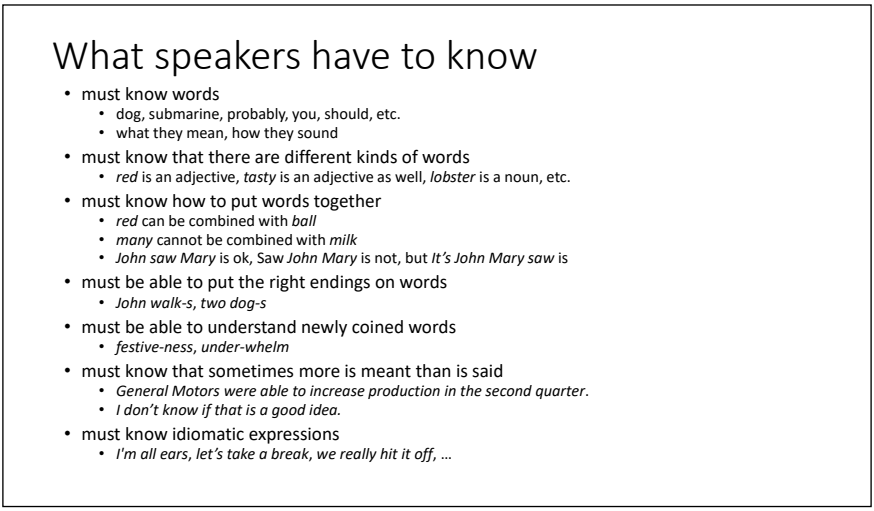


FIGURE 2

a language, that is, when they know how to produce and process language. In other words, Construction Grammar has the same goals as any other cognitive theory of language. What is it that speakers have to know? There actually is not much disagreement. Anyone studying language would have to come to terms with certain things that speakers simply have to know in order to engage in a conversation.

Here's a list of things that speakers of any language have to know in order to talk. I have given you examples from English here, but the main ideas would be the same for really any language that you study. Speakers have to know words, what they mean and how they sound. They must know that there are different kinds of words. For English, there are different word classes like adjectives, nouns and verbs, and speakers know that these word classes behave differently. Speakers must know how to put words together, how to form phrases and sentences out of the lexical words that they know. Speakers must be able to put the right endings on words. If your language is morphological, if it has lots of inflections, then you have to know how to put words together from these smaller materials. Speakers must be able to understand newly coined words. There are derivational morphological word formation processes that enable speakers to produce new words, and hearers have to be able to understand what these new words mean. Speakers must know that sometimes more is meant than what is said. For example, if I tell you *I do not know if that is a good idea*, I do not express just my ignorance on a certain point, but rather I tell you that this is not a good idea. That meaning is not right there in the words, but that is something that you infer. Lastly, speakers must know idiomatic expressions, combinations of words that have non-compositional meanings that convey a kind of meaning that cannot be inferred by understanding the parts of the expression.

In this laundry list of things that speakers have to know, you recognize some traditional domains of linguistic research. How do speakers know what words can be put together? That is the domain of syntax. How speakers are able to produce new words? That falls into the domain of linguistic morphology. Understanding that more is meant than what is said, that is what we usually study in pragmatics. Many linguistic theories view these points as distinct, as falling into different areas of linguistic knowledge, different modules even of linguistic knowledge. Construction Grammar is different in this regard because Construction Grammar posits that all types of linguistic knowledge can be seen as being of the same type, as being cut from the same cloth.

What speakers have to know, from the perspective of Construction Grammar, can be expressed much more concisely than this list, namely that speakers must know constructions. In other words, all the items from our

long list of things that speakers have to know, lexical items, word classes, syntactic patterns, and so on and so forth, can and should be re-conceptualized as knowledge of constructions. Constructions are defined as form-meaning pairings – symbolic units that pair linguistic form with conceptual meaning. This is the underlying basic proposal that Construction Grammar makes. To properly appreciate how this proposal works, what it implies and where it maybe has its limits, I want to flesh it out and discuss ten basic ideas that I view as fundamental for Construction Grammar. If you've understood those ideas, then you're really in a good position to assess what the constructional enterprise is all about.

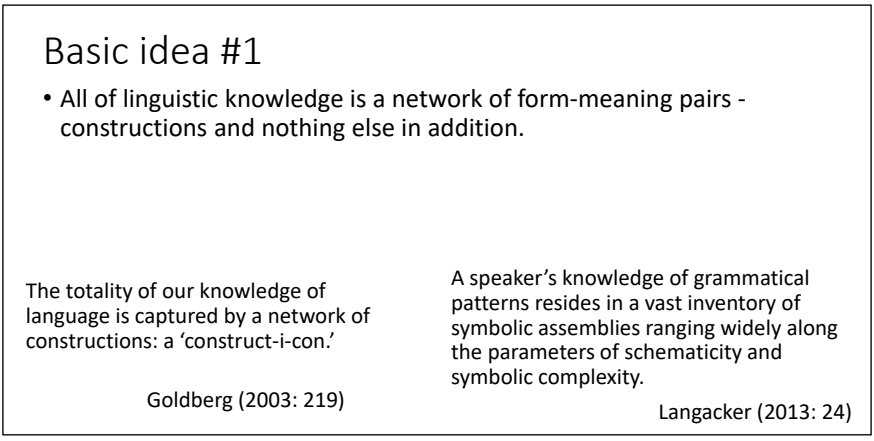


FIGURE 3

Let's jump right in with basic idea #1. All of linguistic knowledge is a network of form-meaning pairs – constructions and nothing else in addition. This idea re-captures what I have been saying so far, especially the last bit there, '*nothing else in addition*', that is sometimes contested. There are colleagues who may be generally sympathetic towards Construction Grammar, but who would feel skeptical about it. Can you really capture everything about linguistic knowledge with form-meaning pairs? Don't you need some kind of abstract syntax? Don't you need some overarching pragmatic principles? What about phonemes? The list goes on, but I would like to stress that this idea is to be taken very literally. Our goal as Construction Grammarians is to explain everything in language through constructions and nothing else.

To give some emphasis to this, I have given you two quotes here, one from Adele Goldberg and the other from Ronald Langacker, both state exactly this idea. Adele Goldberg (2003: 219) states that '*the totality of our knowledge of language is captured by a network of constructions: a 'construct-i-con'*.' Langacker

(2013: 24) formulates it as follows: '*A speaker's knowledge of grammatical patterns resides in a vast inventory of symbolic assemblies ranging widely across the parameters of schematicity and symbolic complexity*.' This idea forms the very basis of everything else I am going to say.

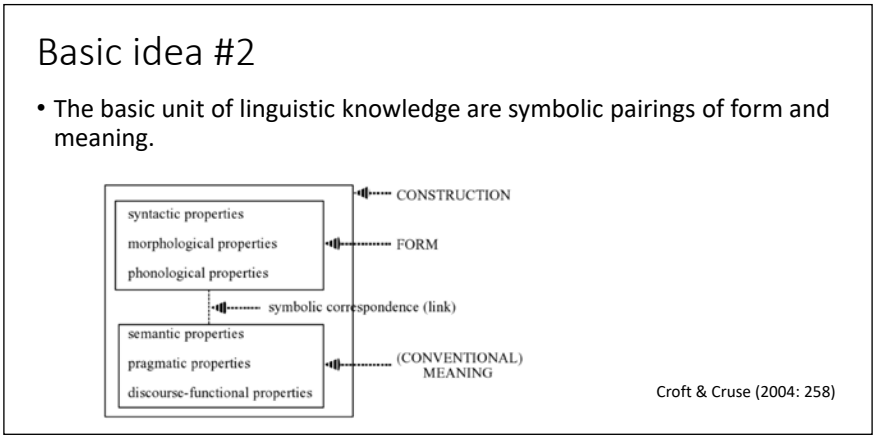


FIGURE 4

The basic unit of linguistic knowledge are symbolic pairings of form and meaning. This idea captures how we define what a construction is. Constructions, the basic units of linguistic knowledge, are defined as pairings of form and meaning. Form, as you see in the diagram on the slide, taken from the work of Croft and Cruse (2004: 258), comprises phonological structure, morphological structures and syntactic structure. Meaning includes semantic, pragmatic and discourse-functional meaning. So there are different shades of form and meaning, and these are linked and connected through the symbolic link, which is an association that is typically arbitrary and established through convention. The diagram on this slide represents a broad consensus in the field, although I will have more to say about it as we go along.

For now, let's carry on, here is basic idea #3. That is the notion that '*constructions vary in terms of their degrees of complexity and schematicity*.' You've seen this already in the quote by Ronald Langacker that I gave you earlier, here we just flesh it out a little more.

This is a diagram that I have adapted from Langacker's work. Langacker (2005: 108) conveys the idea that constructions vary along two axes, the axis of complexity, that is the *x*-axis here, and the axis of schematicity, that is the *y*-axis. The *x*-axis represents a continuum from mono-morphemic simple constructions to more complex patterns that have several different parts. The *y*-axis represents another continuum that ranges from very specific constructions to

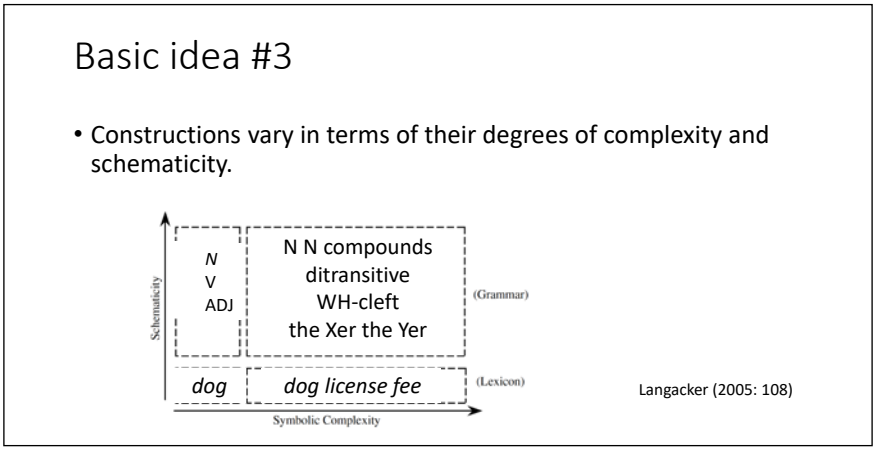


FIGURE 5

more schematic and abstract constructions. This means that on the *y*-axis we have a continuum between lexis at the bottom and grammar at the top, lexis being more specific in meaning, and grammar more schematic.

Let me illustrate this diagram a little further with concrete examples. A construction that is low in complexity and low in schematicity would be a mono-morphemic word such as *dog*. Constructions that are low in complexity but high in schematicity are schematic word classes such as nouns, verbs, or adjectives, which according to basic ideas #1 and #2 we would think of as constructions. They're simple but very general, abstract and schematic constructions. If we move on to constructions that are low in schematicity but high in complexity, we get compound words such as for example, *dog license fee*, which are internally complex. They have structure. But as far as their meaning goes, they are highly specific. That leads to the fourth possible combination, constructions that are both complex and schematic. Here we have grammatical constructions in the traditional sense, units like noun-noun compounds, the ditransitive construction, cleft sentences, constructions like the comparative correlative construction, *the Xer the Yer*, and so on and so forth. One idea that I will come back to several times in this lecture series is how we can tell where exactly in this coordinate system of complexity and schematicity we should locate a specific construction that we are talking about and that we are studying.

Constructions are idiosyncratic, that is, they are to some extent unpredictable. What this means is that as a learner of a language, you cannot deduce how they work from first principles, but rather you have to learn and memorize them. They have characteristics that you cannot figure out or deduce logically, even when you have a lot of knowledge about the language already.

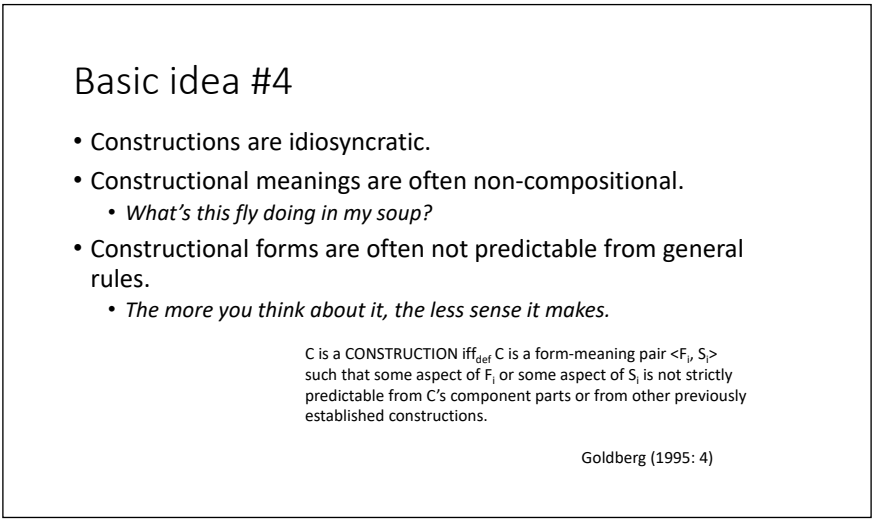


FIGURE 6

Constructional idiosyncrasies have figured prominently in the work of Chuck Fillmore, which is foundational for Construction Grammar. Idiosyncrasies may pertain to both linguistic meaning and linguistic form. With regard to meaning, they give rise to the phenomenon of non-compositionality. For example, in an utterance like '*What's this fly doing in my soup?*', we understand that the speaker is not just asking a question but actually making a complaint. This meaning component is not expressed by the individual words, but rather it emerges from the holistic properties of the construction. Idiosyncrasies with regard to form mean that we have constructions whose form is not predictable from general morphosyntactic rules. This is the case, for example, in the comparative correlative construction: *The more you think about it, the less sense it makes*. Syntactically, this sentence does not look like any other construction of English that we might have come across. This is something that we need to learn and memorize as second language learners of English.

The basic idea of constructional idiosyncrasies is so central for Construction Grammar that it has made its way into one of the most influential definitions of what constructions are, namely the definition proposed by Adele Goldberg in her book on argument structure constructions (1995: 4):

> C is a construction if and only if that construction is a form-meaning pair, such that some aspect of the form or some aspect of the meaning is not strictly predictable from the constructions' component parts or from other previously established constructions.

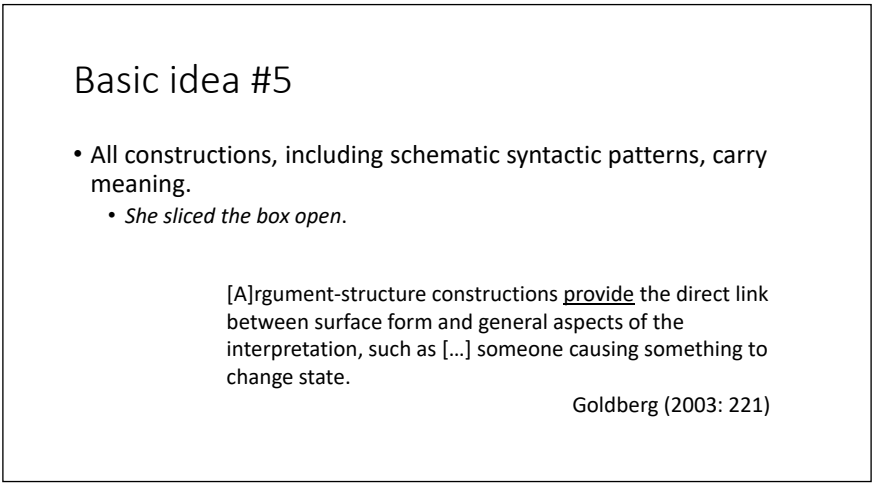


FIGURE 7

One question that you might ask here is why construction grammarians make such a big deal out of idiosyncrasies. If you look at authentic natural language use, you'll actually find that idiosyncrasies are ubiquitous. You find them everywhere. You run into them a lot more than traditional accounts of language structure would have us believe, and that is why it is such an important idea.

I am coming to basic idea #5, which is the idea that all constructions, including schematic syntactic patterns, carry meaning. Many working linguists will tell you that during their years of study, there have been two or three studies that have opened their eyes towards a certain aspect of language and that made a lasting impact on the way they came to think about language. For me, one such text was Adele Goldberg's work on argument structure constructions. That work makes the general point that syntactic patterns like the ditransitive construction or the resultative construction are not just formal syntactic patterns. Instead, they are symbolic units that carry meaning. In a sentence like *She sliced the box open*, the structure of that sentence conveys the meaning that as the result of her slicing, the box opened. Today, it seems very obvious to me that the syntactic patterns can convey these ideas. But before reading Goldberg, I thought that syntax is about arranging words into phrases and that meaning chiefly resides in the words, not in the patterns. In the words of Goldberg herself, *argument structure constructions provide the direct link between surface form and general aspects of the interpretation, such as someone causing something to change state* (2003: 221). So for me, this idea substantially changed how I viewed language.

Basic idea #6 is the so-called 'principle of coercion'. When there is a conflict between lexical meaning and the meaning of grammatical constructions,

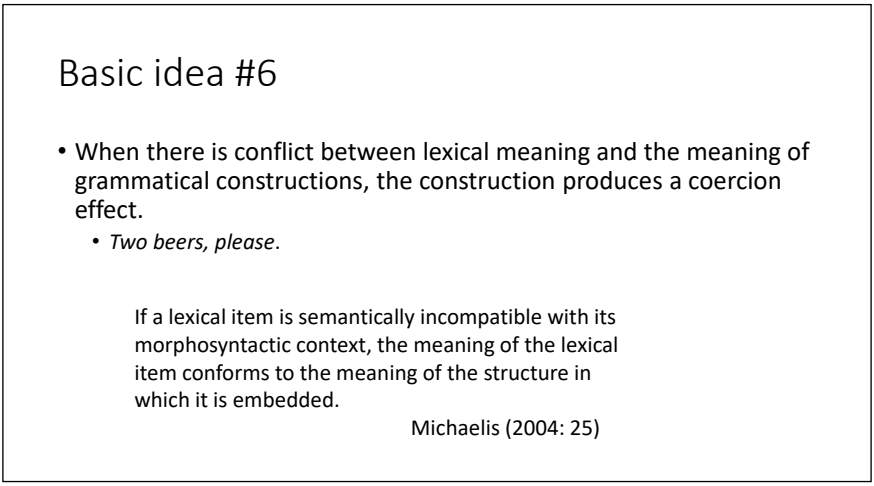


FIGURE 8

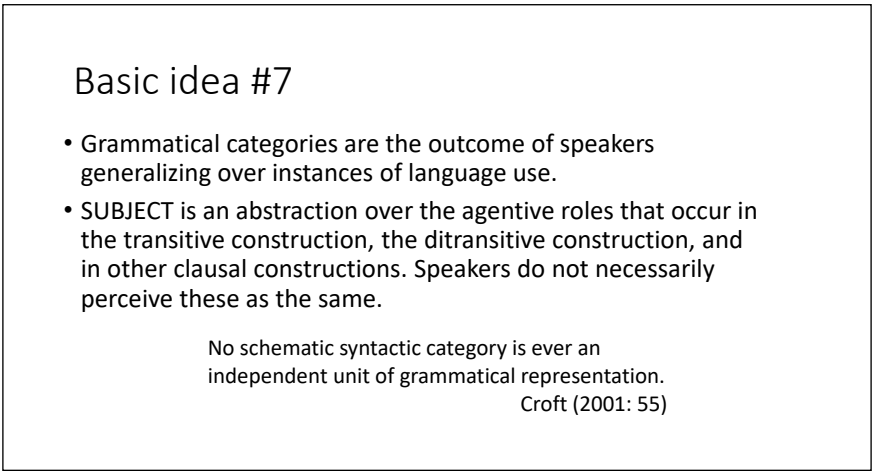


FIGURE 9

the construction produces a coercion effect. Let me illustrate this. In English there are nouns such as *beer*, which function as so-called mass nouns. They denote substances that are not easily counted. When I put them in a context where they are treated as countable, I do something to their meaning. Take an utterance such as *Two beers, please*. Instead of treating *beer* like the mass noun that it is, I convey that I would like two units of beer, that is, two glasses of beer or two bottles of beer. This is the principle of coercion by construction, which was formulated by Laura Michaelis as follows (2004: 25): *If a lexical item is semantically incompatible with its morphosyntactic context* (for example *beer*, which as a mass noun is incompatible with the context of the plural, so that an

uncountable is made countable through the plural), *the meaning of the lexical item conforms to the meaning of the structure in which it is embedded*. In other words, the construction wins. If there is conflict between a lexical item and its constructional context, then the construction wins out.

Basic idea #7 is a notion that is now generally accepted in usage-based linguistics, but that I found quite far-reaching when I first encountered it in the work by Bill Croft, in his book (2001) on Radical Construction Grammar. What he argued specifically was that *grammatical categories are the outcome of speakers generalizing over instances of language use*. That is, a grammatical category such as subject, for instance, is not a grammatical primitive or a unit that is basic, but rather it is the opposite. It is an emergent phenomenon. It is a generalization or an abstraction over the agentive roles that occur in the transitive construction (*John kicked the ball*), the ditransitive construction (*John gave Mary the book*), or other clausal constructions (*John promised to pick me up*). Speakers do not necessarily perceive these as exactly the same, but they perceive these roles as similar enough to instantiate a broad category of subject in English. Across languages these categories are not the same. The same holds for other notions that we as linguists are perhaps used to seeing as very basic. We have been trained to work with categories such as nouns, verbs, cases like dative and accusative, subordinate clauses, and so on. All of these high level generalizations are really emergent phenomena. They're not grammatical primitives, but rather they are the outcome of your experience with many tokens of language use.

Croft (2001: 55) formulates this idea like this: *No schematic syntactic category is ever an independent unit of grammatical representation*. Every category is the outcome of speakers hearing many instances of language use and drawing a generalization from that experience. This is important, since it means that we have to ask ourselves this: When is a phenomenon that we are studying a construction? Can we assume that speakers have drawn a generalization from the input that they have had? What is the evidence for this? We will come back to this question a couple of times.

Basic idea #8 is an idea that is near and dear to my heart. Let me explain what it is about. The idea is that constructional meaning is reflected in associations between syntactic patterns and lexical elements. The sentence *John gave Mary the book* instantiates what we call the ditransitive construction, which has as its basic meaning the idea of a transfer. It therefore comes as no surprise that the verb *give* is the one verb that is most strongly attracted to that construction, and that is most strongly associated with that construction. A similar point can be made for the *way*-construction, *He elbowed his way through the crowd*. This construction conveys the idea of movement along a path that is difficult and laborious. In that construction we tend to find verbs such as the denominal

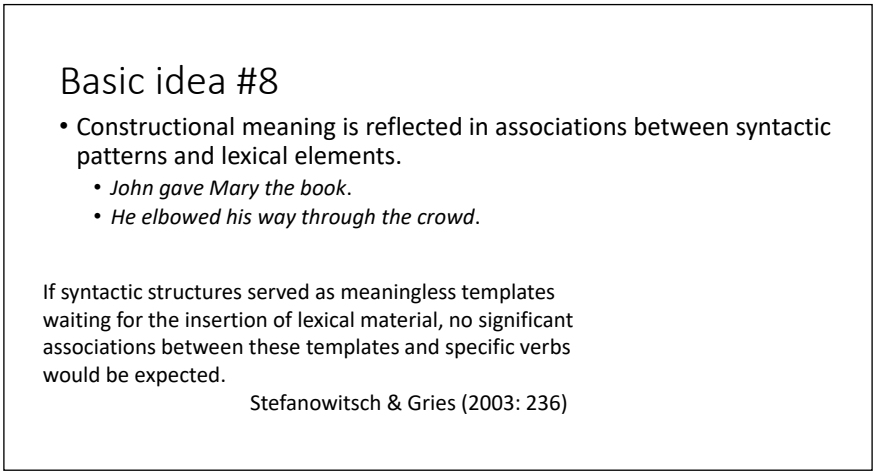


FIGURE 10

verb *elbow*. It means that you create your path, pushing other people with your elbows, for example when you enter the subway. This conveys exactly the kind of difficult movement that the construction commonly expresses. If you like, you can see this as a kind of harmony in meaning between the meaning of a construction and lexical items that occur in that construction. You could see this as another piece of evidence that syntax is in fact meaningful. This slide shows a quote by Stefanowitsch and Gries (2003: 236): *If syntactic structures served as meaningless templates waiting for the insertion of lexical material, no significant associations between these templates and specific verbs would be expected*. If syntax were really just a set of rules of putting words together, then why do we see these harmonious patterns of specific verbs being attracted to specific syntactic contexts?

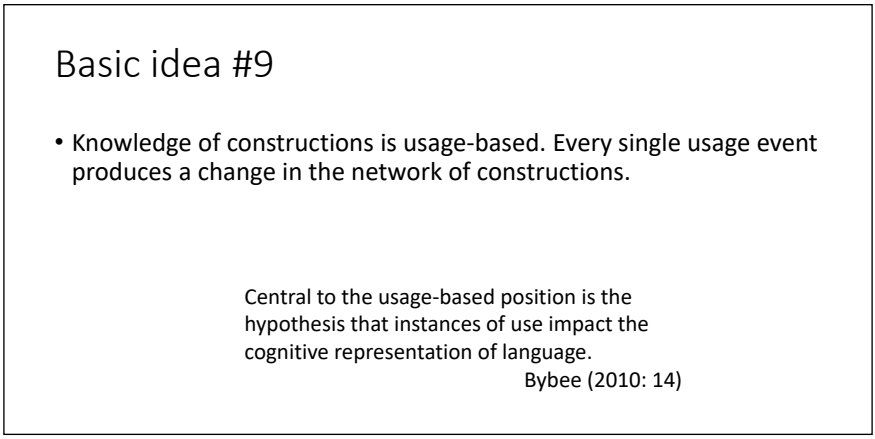


FIGURE 11

Basic idea #9 is that knowledge of constructions is usage-based. Joan Bybee, one of the principal architects of usage-based linguistics, formulates it in this way (2010: 14): *Central to the usage-based position is the hypothesis that instances of use impact the cognitive representation of language*. This means that every instance of language use leaves a little imprint on our knowledge of language. This may seem strange at first because it seems to imply that we actually remember everything we've ever heard, every conversation that we've ever been in. I do not know about you, but I forget things all the time, including my keys and my boxed lunch. I have three children and sometimes when I talk to them, I have to go through all three names before I finally get to the right one. I like to think that I am still a normal human being. Normal human beings sometimes forget things. Now with language, most language that we encounter instantiates words and patterns that we have heard lots and lots of times. Hearing the same structure and the same words once again will not change our representations a whole lot. By contrast, hearing structures that deviate from what you've heard before will actually force you to adjust your linguistic representations, just a little bit. Perhaps when you're listening to me, you may need to get used to my accent, the way I pronounce my words, and the melody of my phrases. This is day one, so you're still adjusting and you're still trying to figure out my vowels and other aspects of the way I say things. But you'll see that over day two, day three and day four, you will gradually settle into the way of processing my speech. You will find it easier to follow. That means that your cognitive representations of language have indeed changed just a little bit, just by listening to me. That is the core idea of usage-based linguistics, and that brings me to basic idea#10, which is the bedrock of any usage-based understanding of language, and which I find best expressed in the work of Michael Tomasello.

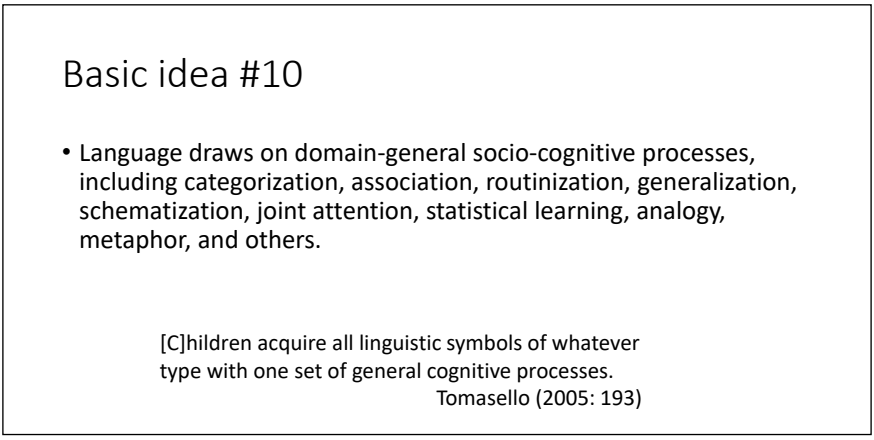


FIGURE 12

The idea is that language draws on domain-general socio-cognitive processes, including categorization, association, routinization, generalization, schematization, joint attention, statistical learning, analogy, metaphor, and others. The list goes further, but these are the main processes. Whereas, for example, generative linguists assume that we come into this world with language-specific knowledge, a universal grammar, cognitive linguists and construction grammarians work on the assumption that a specific combination of general cognitive and social skills is actually enough of a basis for language learning. This combination of skills explains why humans have language and why other animal species do not. No one disputes that humans and other primates differ in this regard. We have language, but other species do not. We need an explanation for that, and the usage-based explanation is that our configuration of socio-cognitive skills is a different one. The way Tomasello (2005: 193) puts it is that [*C*]*hildren acquire all linguistic symbols of whatever type with one set of general cognitive processes*.

These ten ideas are points that I am prepared to defend against any criticism that could be leveled against them. I assume them as a foundation of everything else that I will have to say in this lecture series. Seeing as these points are rather general, I would like us to move on with questions that go into some more detail with regard to constructions, what they are and how we can tell whether a linguistic form is actually a construction.

It is perhaps easy to identify the *way*-construction or the ditransitive construction, but beyond that, how can constructions be identified? One question that my students ask all the time is how do I know if something is a construction? Is everything a construction? I want to go over four strategies that you can apply when you're asking yourself whether you are dealing with a construction.

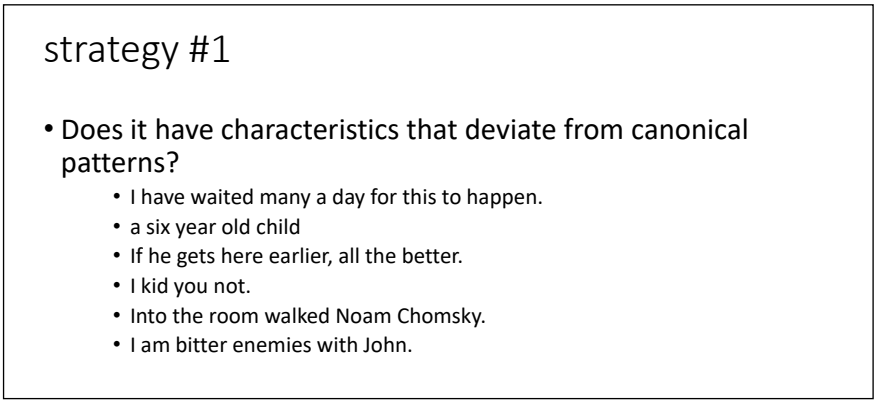


FIGURE 13

The first strategy relates to basic idea #4, the idea that constructions are idiosyncratic, that there is something unpredictable about them. When you're looking at a linguistic structure, do you see characteristics that deviate in some way from canonical patterns of the language that you are studying? Can you figure it out from other regularities that you know, or is there something that you would have to learn and memorize that is specific to this pattern? This slide lists a few examples that show formal idiosyncrasies of this type.

For example, in the utterance *I have waited many a day for this to happen*, the phrase *many a day* will look odd to many of you, since it shows an unusual sequence of the quantifier *many*, the singular indefinite article *a*, and a singular noun. Normally, we expect to see *many* with a noun that is in the plural, as in *many days*. Here it is *many a day*, which is clearly different from canonical English syntax. If you find any deviation of this kind, you know that you're looking at a pattern with some irregularity, which fits our definition of a construction.

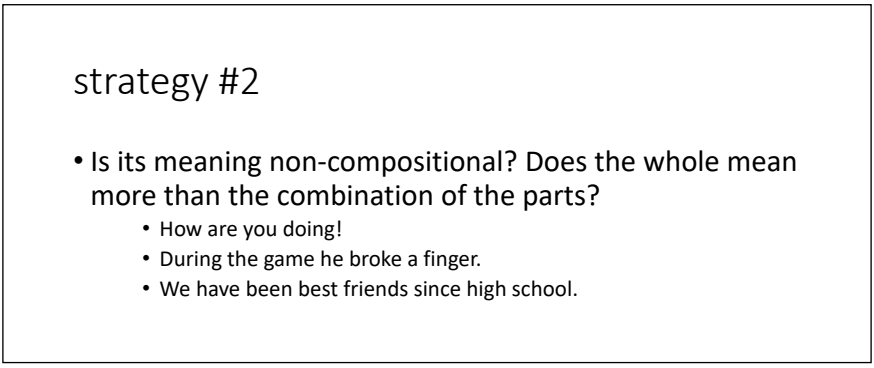


FIGURE 14

The second related strategy would be to look for non-compositional meanings. If we have a particular example of language use, we can ask ourselves whether its meaning is non-compositional. Does the whole somehow mean more than the combination of the parts? Let's just take a quick look at the first example. The utterance *'How are you doing!'* may look like a question, but it also has the non-compositional meaning of a greeting formula, and that is a matter of convention. That is something that you cannot figure out on the basis of the words alone. It is something that you have to learn on the basis of contextual information. Whenever the meaning of the whole is more than the meaning of the parts, then you know that you're dealing with a construction.

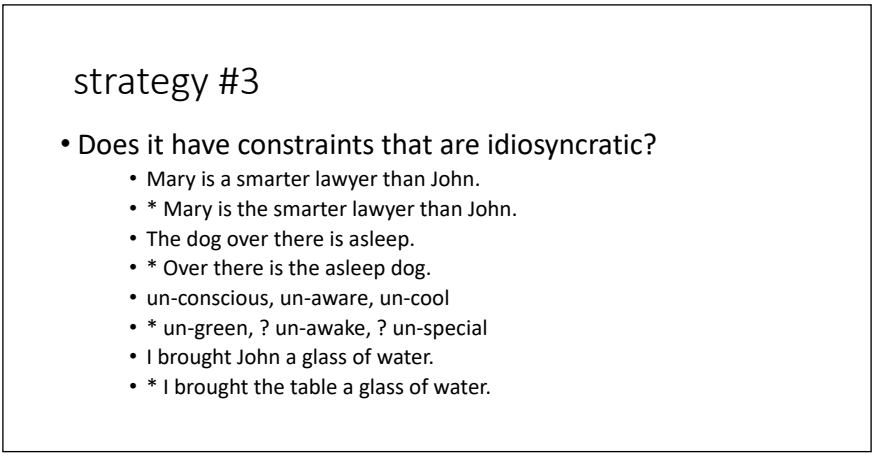


FIGURE 15

For strategy #3, you need to ask the following question. Does the pattern that I am looking at have constraints that are idiosyncratic? This strategy, I am ready to admit, is a little tricky, specifically for second language learners of English, because it requires you to manipulate the pattern in several ways, and make judgments about what can and cannot be said. As a linguistic methodology, introspective grammaticality judgments have serious problems, but I think that there are contexts in which they show us something.

Let us take, for example, the utterance *The dog over there is asleep*. Speakers of English use adjectives like *asleep* in what is called a predicative construction, in which the adjective follows a form of the verb *to be*. If you try to use *asleep* in a different syntactic position, as an attributive adjective, as in **the asleep dog* or **the asleep baby*, speakers of English will actually give you strange looks, because that is not the way you use this kind of adjective. *Asleep* does not work in that way. You can't put it before a noun. You have to say *the baby is asleep* rather than **the asleep baby*. Restrictions of this kind mean that we are dealing with a construction. There is an idiosyncrasy about *asleep* and other related adjectives that has to be learned. The other examples on the slide illustrate the same point with other structures, but the general point in all cases is that whenever a manipulation of a grammatical structure yields an expression that sounds odd or unacceptable, that means that you're looking at a construction.

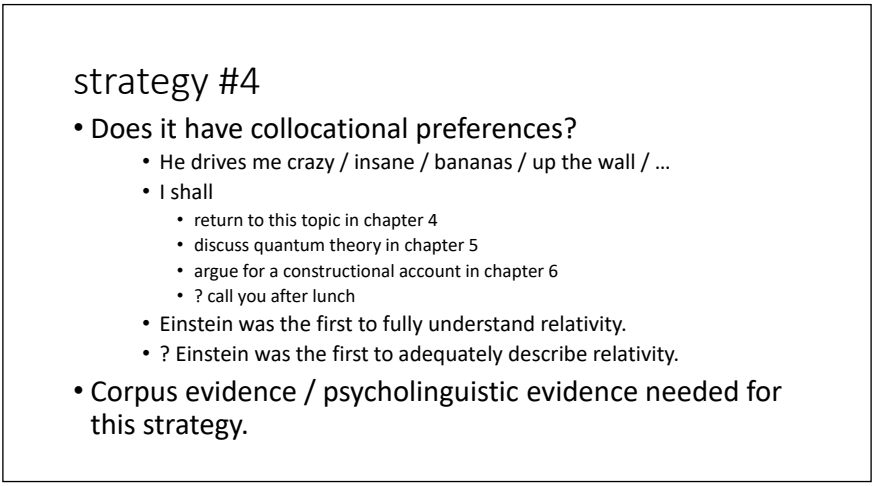


FIGURE 16

Strategy #4 brings us back to the basic idea that constructions have collocational preferences and that syntactic patterns tend to be associated with specific lexical items. Whenever we have a syntactic pattern that is not just indiscriminate with regard to the words that it occurs with, but that shows particular associations, then we can say that we found a construction. The question to ask is whether a linguistic form has collocational preferences and what these preferences are. Strong collocational preferences are evident in the case of idioms, such as the *drive someone crazy* construction. There are a handful of elements that can appear in the final predicative slot of the construction. You can *drive someone crazy*, you can *drive them insane*, you *can drive them bananas*, *mad*, or *up the wall*, but that is already more or less the whole spectrum. There are many adjectives or other expressions that do not work in the *drive crazy* construction. You cannot **drive someone happy* or **drive someone sane*, so there are limitations, and these limitations instantiate constraints that the construction has.

There are other constructions that are a lot more open with regard to their collocates, but that show nonetheless a recognizable profile of collocational preferences. For example, the English auxiliary *shall* is strongly associated with lexical verbs such as return, as in *I shall return to this topic*, or discuss, as in *I shall discuss this in chapter five*. It is much less associated with other lexical verbs. An utterance such as *I shall call you after lunch* is possible. But it is less idiomatic than *I shall argue* or *I shall discuss*. The last example on this slide illustrates the so-called split infinitive construction in English, *Einstein was the first to fully understand relativity*. *To fully understand* is a split infinitive that consists of the infinitive marker *to*, the verb *understand* and the adverb *fully*,

which so to speak splits the two in half. This construction works well with certain adverbs, but not so well with others. *To fully understand* sounds good to most speakers, but **to adequately describe* does not. It is not ungrammatical, but it is clearly worse. In any case, the bottom line would be that the presence of collocational preferences is a hint that a construction is present. You can try to argue for collocational preferences on the basis of grammaticality judgments, as I have done here, but really the more suitable evidence for this would come from corpus studies or from psycholinguistic experiments. I would also like to add that strategy #3 and strategy #4 boil down to the same idea: They show preferences and restrictions on a continuum from hard constraints to probabilistic biases.

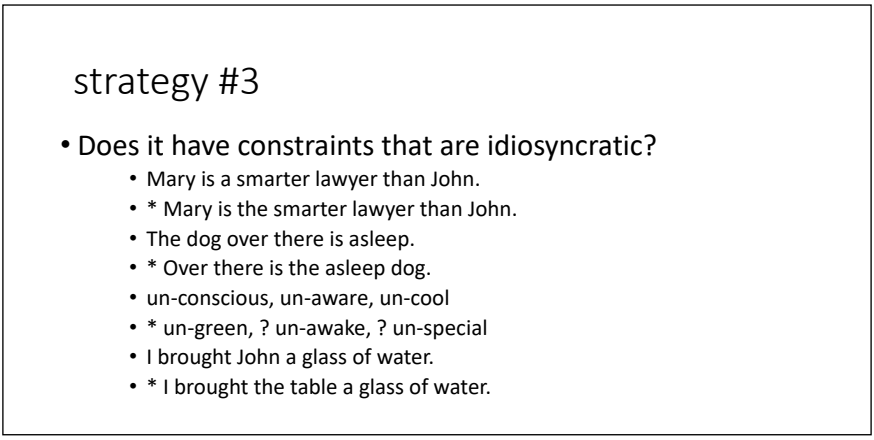


FIGURE 17

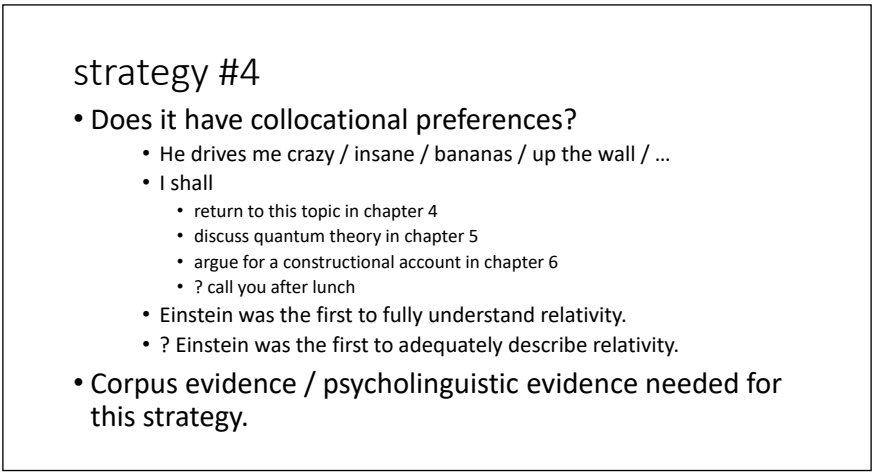


FIGURE 18

Strategy #3, which prompts us to ask whether a construction exhibits idiosyncracies, shows hard constraints, manipulations that you absolutely cannot do with a construction.

Strategy #4 relates to probabilistic preferences. This relates to patterns of co-occurrence that are preferred and dispreferred. Some patterns are highly entrenched, others are possible, but do not work quite as well as others.

Putting it all together then, a linguistic form is a construction if it deviates from canonical patterns, if it shows non-compositional meanings, if it has idiosyncratic constraints, and if it has collocational preferences. I know that there are members of this audience for whom everything I have said so far is well known and perhaps even self-evident, and I thank you for bearing with me up to this point. I would now like to leave the well-trodden paths and discuss ideas that are somewhat more controversial.

Specifically, I would like to talk about five controversies, which represent issues that construction grammarians do not agree on with each other. These controversies, as I see them, are current construction sites of the field, where the architects do not really agree if they want to build a bridge, or maybe rather a tunnel, or perhaps both.

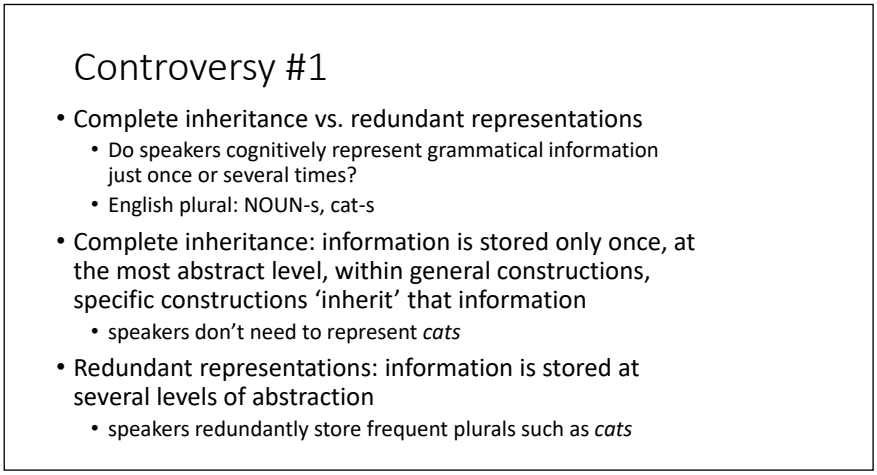


FIGURE 19

The first controversy is concerned with relations between constructions in the constructional network, that is, the idea that is commonly known as inheritance. Everyone agrees that constructions are connected, but what exactly are the consequences are of these connections? That is a matter of debate, and the principal conflict here is between two views, which we can label "the complete inheritance view" and "the view of redundant representations", respectively.

The underlying question is whether speakers cognitively represent grammatical information just once, or rather several times.

Let me give you an example. You and I can understand the word *cats*, *cat* in the plural form. Is that because we have memorized the word *cats*, or is it because we know the plural construction that tells us to form *cats* by adding an *-s* to the singular form *cat*? Is it one or the other, or is it perhaps even both? The view from complete inheritance is that anything that you can describe with a generalization does not have to be separately stored and memorized. Complete inheritance means that information is stored only once, namely at the most abstract level within a general construction. Then specific constructions can inherit that information. They can look it up at a higher level of abstraction in the constructional network. By virtue of that, speakers actually do not need to represent *cats*. You can see that is a very economical and very elegant way of storing information. You can encode lots of information with relatively little storage, but at the same time it raises questions.

The opposing view, the view of redundant representations, is held by researchers like Ewa Dąbrowska (2017), who has actually argued forcefully for it in this very room at the forum. She argues that information is stored at several levels of abstraction. Even though we do not technically need to memorize the word form *cats*, because it is such a frequent word, we actually cannot avoid remembering it, and we end up with a redundant representation. On that account, information is stored at several levels of abstraction, so speakers redundantly store frequent plurals with their lexical items in addition to a general plural construction.

If you ask me where I stand on the issue, I appreciate Charles Fillmore and his work with all my heart. But here I would side with Ewa Dąbrowska's view, which embodies the perspective of current usage-based linguistics.

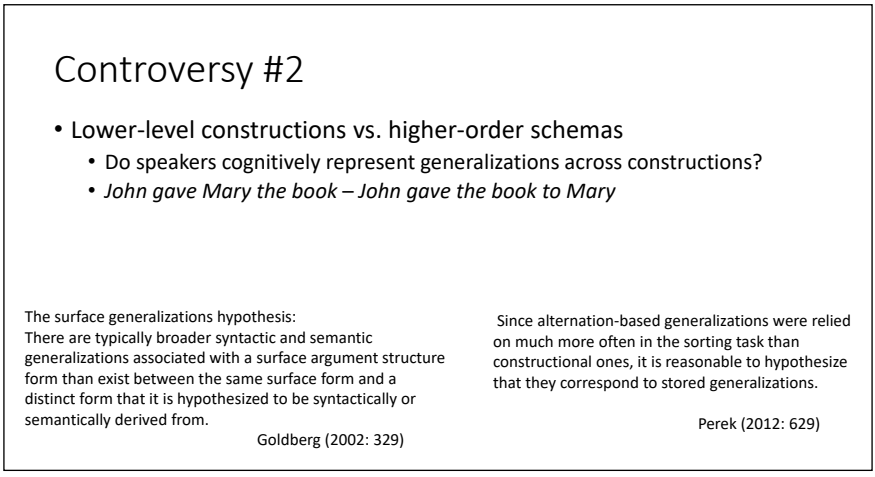


FIGURE 20

Controversy #2 brings us back to the continuum of complexity and schematicity that I talked about earlier. How abstract are the generalizations that speakers draw? If you remember the coordinate system that Langacker proposes to us, where exactly do we place a given construction? Is more information represented at low levels of generalization, or are there highly abstract schemas, perhaps even schemas of schemas, or meta-generalizations? Do speakers cognitively represent generalizations across constructions? That is under debate. For example, is there a generalization across the ditransitive construction, *John gave Mary the book*, and the prepositional dative construction, *John gave the book to Mary*? You could argue that both express a transfer, so they have things in common, and as human beings, we categorize items that have features in common. Why not? There could be a generalization of this kind. Adele Goldberg has been an advocate of the idea that low-level generalizations are very important. She has formulated the so-called "surface generalizations hypothesis" (Goldberg 2002), which goes as follows. The surface generalizations hypothesis states that

> there are typically broader syntactic and semantic generalizations associated with a surface argument structure form than exist between the same surface form and a distinct form that it is hypothesized to be syntactically or semantically derived from.
>
> GOLDBERG 2002: 329

Even though the ditransitive and the prepositional dative have features in common, what they have in common is less substantial than what each individual construction has in terms of its individual characteristics. But this does not necessarily mean that speakers do not draw any abstract generalizations. Florent Perek has done empirical work on this, finding that speakers actually use higher-order schemas when they reason about language. When you give them a categorization task, they will draw on generalizations that reach across several constructions. Perek conducted a sorting task and concluded the following (2012: 629): *Since alternation-based generalizations were relied on much more often in the sorting task than constructional ones, it is reasonable to hypothesize that they correspond to stored generalizations*. This indicates that speakers store alternations of constructions as meta-generalizations. Adele Goldberg and Florent Perek, as some of you might know, are by now co-authors of a series of studies, so their views are not entirely incompatible. I will come back to this issue later in this lecture.

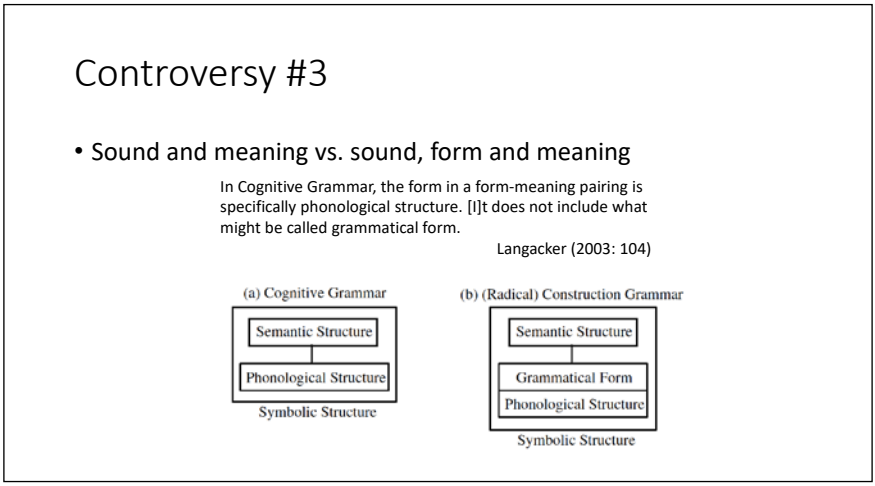


FIGURE 21

Controversy #3 plays out between Ronald Langacker on the one hand and William Croft on the other, two founders of cognitive linguistics. The issue is how we understand constructions as pairings of form and meaning. Specifically the question is, is form just sound, or is form sound and morphosyntax? Langacker is crystal clear on this (2003: 104): *In Cognitive Grammar, the form in a form-meaning pairing is specifically phonological structure. [I]t does not include what might be called grammatical form*. The diagram on this slide shows that Croft's Radical Construction Grammar subsumes grammatical, morphosyntactic form under the form side of constructions. Langacker himself only assumes a link between semantic structure and phonological structure. Cognitive Grammar attempts to reduce all linguistic structure to concepts and sounds. It is an ambitious reductionist enterprise. For William Croft, grammatical form is the substance that linguists are working with: words, suffixes and syntactic patterns. It is hard to part with that working material. Can morphosyntax really be reduced to sound? What do we do with part-of-speech constructions like nouns or notions like subject? Can we find sounds that correspond to these categories? Where does Langacker's linking of concepts and sound leave notions such as linear sequence? I will simply leave you with these questions and move on to controversy #4, the question whether morphemes are constructions or whether they are parts of constructions.

Adele Goldberg has produced several overviews of different construction types, including morphemes like English affixes *pre-* or *-ing*. The table on this slide shows different examples of constructions varying in size and complexity, and morphemes are part of that. This view is not generally shared.

Controversy #4

- Morphemes as constructions vs. morphemes as part of constructions

Goldberg (2006: 5)

TABLE 1.1. Examples of constructions, varying in size and complexity

Morpheme	e.g. *pre-, -ing*
Word	e.g. *avocado, anaconda, and*
Complex word	e.g. *daredevil, shoo-in*
Complex word (partially filled)	e.g. [N-s] (for regular plurals)
Idiom (filled)	e.g. *going great guns, give the Devil his due*
Idiom (partially filled)	e.g. *jog* <someone's> *memory, send* <someone> *to the cleaners*
Covariational Conditional	The Xer the Yer (e.g. *the more you think about it, the less you understand*)
Ditransitive (double object)	Subj V Obj1 Obj2 (e.g. *he gave her a fish taco; he baked her a muffin*)
Passive	Subj aux VPpp (PP_{by}) (e.g. *the armadillo was hit by a car*)

Booij (2013: 256)

$[V_{TR_i}\text{-able}]_{A_j} \leftrightarrow [[\text{CAN BE SEM}_i\text{-ed}]_{PROPERTY}]_j$

Morphemes are parts of constructions, but not constructions themselves.

FIGURE 22

Geert Booij has worked out a constructional account of morphology and states that morphemes are parts of constructions, but not constructions themselves. If we have, for example, a construction that has the suffix *-able* in English, that pattern maps onto a semantic structure, but the affix by itself does not. It can't be used by itself. Both views have advantages. If we say that morphemes are constructions, we acknowledge that they are symbolic pairings of form and meaning, and that we can maintain the idea that knowledge of language is knowledge of constructions and nothing else. If we say that morphemes are parts of constructions, we recognize that they actually need a linguistic context to be produced, a host that they can attach to, and that their meaning comes about in that specific context, but not in others. I will have more to say about morphological constructions in later lectures, so we will come back to this.

Controversy #5 is a clash of two views on frequency. Are associations between constructions and lexical elements measured best by raw frequencies or by a collocational measure? Anatol Stefanowitsch and Stefan Th. Gries have developed collostructional analysis as a way of finding construction-specific patterns of lexical preferences. That is, when we count the frequency of a lexical item in the context of a specific construction, we also need to take into account how often that lexical item is used elsewhere, outside of the construction. That tells us whether it is actually occurring frequently in the construction because it is attracted to that construction, or simply because it is very frequent everywhere else as well. Joan Bybee sees things differently.

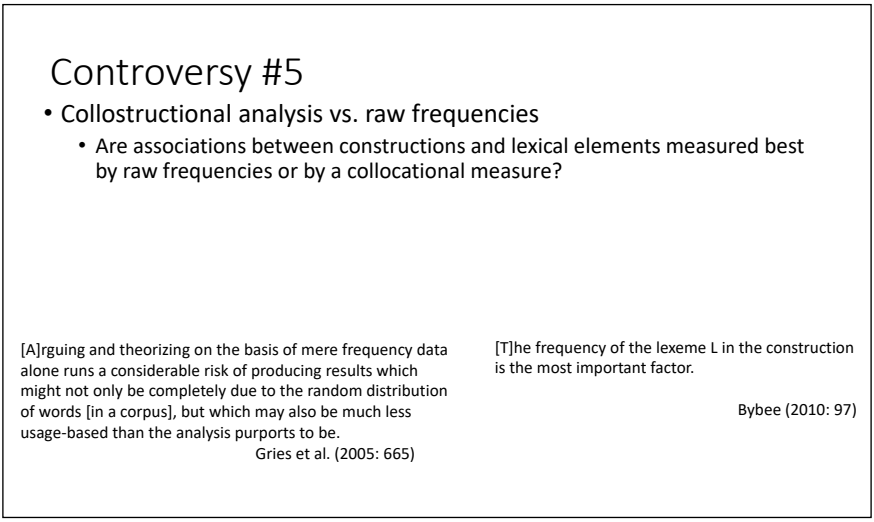


FIGURE 23

The position that Stefanowitsch and Gries hold can be summarized as follows (2005: 665):

> [A]rguing and theorizing on the basis of mere frequency data alone runs a considerable risk of producing results which might not only be completely due to the random distribution of words [in a corpus], but which may also be much less usage-based than the analysis purports to be.

Bybee, on the other hand, defends the use of raw frequencies. She (2010: 97) states that *the frequency of the lexeme L in the construction*, and here she means raw frequency, *is the most important factor*. Now, with regard to this controversy, my own position lines up with Stefanowitsch and Gries. As you will hear in the next lectures, I have been working extensively with collostructional methods. I believe that the evidence that you get from those methods actually speaks for itself.

Summing up, the five controversies that I have discussed concern the conflict between complete inheritance versus redundant representations, the importance of low-level constructions versus higher-order schemas, and the idea of constructions being forms that are paired with sound, or meanings that are paired with sound and grammatical structures. Then we have the controversy of morphemes as constructions versus morphemes as parts of constructions, and lastly, collostructional analysis versus raw frequencies. Up to this point, we have covered the basic notions of Construction Grammar, we have

seen some strategies that allow us to define constructions, and we have learned about some controversies.

Let me finish this first lecture with an outlook on some new directions in Construction Grammar. Earlier this year, Adele Goldberg (2019) published a book with the curious title *Explain Me This*, which explores two theoretical notions, namely, coverage and statistical preemption. I will explain what both of these are. The central question of the book is what Goldberg calls the *explain-me-this* puzzle. **Explain me this* is an ungrammatical sentence of English. It should be *Explain this to me*. *Explain* famously does not work in the ditransitive construction, and speakers of English know this. The question is, how do they know this? They haven't been told so by their parents. They haven't read it in a book. They somehow came to understand that it is not possible to use *explain* in this way. How do you learn not to say things? That is a veritable scientific puzzle that Goldberg tries to solve. She asks, how is it possible that speakers of a language accept certain utterances that they have never heard before as completely idiomatic, while they reject other utterances as simply impossible and ungrammatical? Consider these two examples. *Vernon tweeted to say she does not like us* is a sentence that you may have never seen before, but that speakers judge to be possible. **She considered to say something* is also something that you may never have heard before, but here speakers will insist that this does not sound right. How do these different judgments come about? The phenomenon, I should add, is not limited to English. If you do not have intuitions on these two sentences, do not worry. In any language, there are new ways of saying things that are easily possible. There are some unusual word combinations that are fine, and then there are other combinations that speakers will reject as not possible. Goldberg tries to explain why this is so.

She argues that there are two central factors. The first is coverage, which involves the mutual similarity between different instantiations of a construction. The second is statistical preemption, which relates to the idea of competition between constructions. Let me say a bit more about coverage first. Coverage can be broken down into the following ideas. For any given construction, there is a degree of mutual similarity between different instantiations of a construction. Speakers have highly detailed linguistic memories. That is the basic idea of usage-based linguistics that I presented to you. For any construction that we use, we keep a record of the examples that we hear, how similar they are to each other and where any new example that we find fits in. That applies across all levels of linguistic structures. That applies to you adjusting to my vowels. It applies to speakers hearing a new and original instance of the ditransitive construction. When we hear a new token of language use, we integrate that into our old memories. If the new token is just like everything

we've ever heard before, our overall representations are not changed. But when this new instance is different, then it subtly shifts our representations to a new space. Our linguistic categories and their representations are continually formed and updated, so that our knowledge of language continually changes over time. That means that any construction that we are talking about has a certain quality of coverage. Let me make this more concrete.

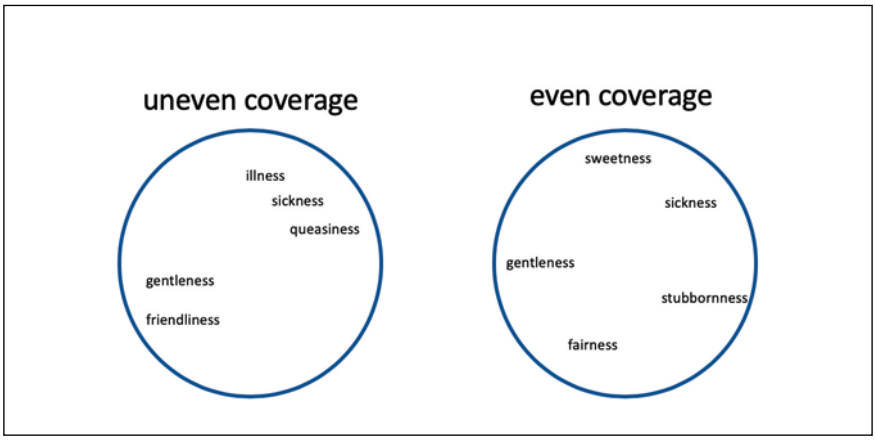


FIGURE 24

Here we have two hypothetical scenarios for which I am using the English derivational suffix *-ness* as an example. You can think of this as two speakers and their respective experience with the English *-ness* construction. In the first case, the construction has exactly five different types, and these five fall into two clearly defined groups. We have *gentleness* and *friendliness*, which are both about someone being an agreeable person. Then the other three, *illness*, *sickness*, and *queasiness*, all relate to not feeling well in different ways. This is what Goldberg would call "uneven coverage". The distribution of types is not very homogeneous, instead there are clusters.

In the second case, we have a speaker who has a very different experience with the *-ness* construction. We have the same construction with the same number of types, five different words, but those types all convey very different ideas. We have *gentleness*, *sweetness*, *sickness*, *stubbornness* and *fairness*, which encode very different meanings. This would illustrate what Goldberg calls 'even coverage'. All words are approximately the same in terms of their similarity and distance from each other.

Now imagine what happens if these two speakers that we have here encounter a new type of the construction, for example, the word *carelessness*. In the first case, *carelessness* is situated right between these two clusters that the

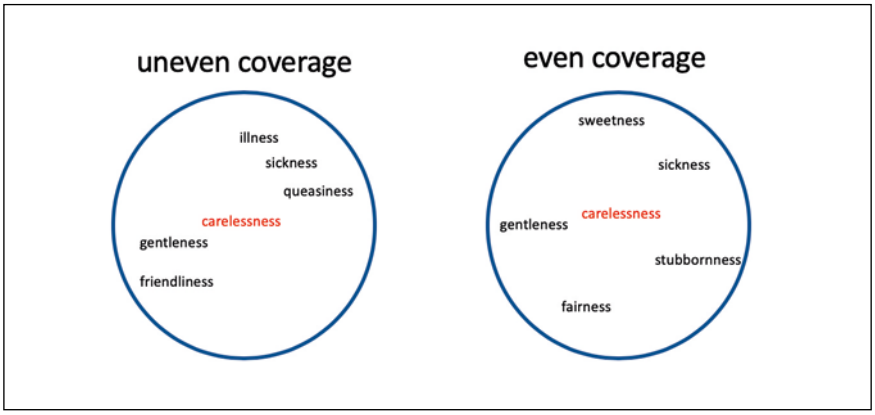


FIGURE 25

speaker has come to represent. *Carelessness* requires the speaker to update and change their knowledge of the construction in substantial ways. They realize that there are instances of the construction between the established clusters. Instead of the two well-defined clusters, there is now a more continuous semantic spectrum.

In the second case, the distance of every type to every other type is still more or less the same. Nothing much changes, just one more type has been added to a semantic spectrum that is already semantically diverse. The implication is that constructions with even coverage allow new types much more easily than constructions with uneven coverage.

Going back to the earlier state of affairs, which speaker do you think will be more inclined to form new types, to come up with new words that have this *-ness* suffix? The one on the left will probably produce types that relate to these two clusters, but not new types that would be situated between them, in the middle of the semantic space. That is what is meant by the term coverage.

The second notion is captured by the term statistical preemption. Statistical preemption relates to the idea of higher-order schemas that I talked about earlier in connection with the work of Florent Perek. The idea is that speakers form generalizations over sets of functionally similar constructions, like the ditransitive (*John give Mary the book*) and the prepositional dative (*John gave the book to Mary*). Speakers realize that these two constructions are similar, and they keep track of the frequencies of lexical elements that occur in them. They take note of asymmetries that they view as striking, as conspicuous. This can actually explain how speakers learn not to say certain things. To make this idea more concrete, let me show you an example of how Boyd and Goldberg (2011) investigated this.

On the following slides, you will see moving pictures, and I just want you to say silently in your mind what happens.

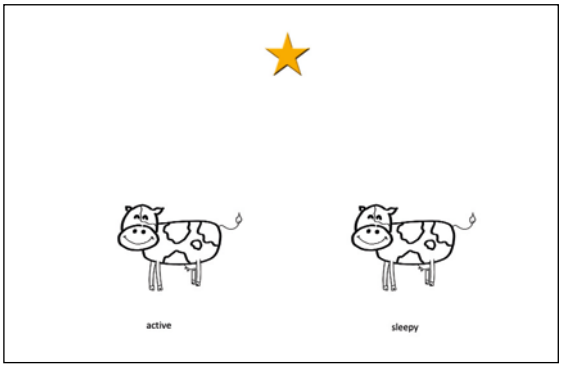


FIGURE 26

On this slide you see two cows. There is an *active* cow and a *sleepy* cow. Participants in this experiment would see something like this [The left cow moves to the star], and they would have to describe the depicted event by saying *The active cow moves to the star*.

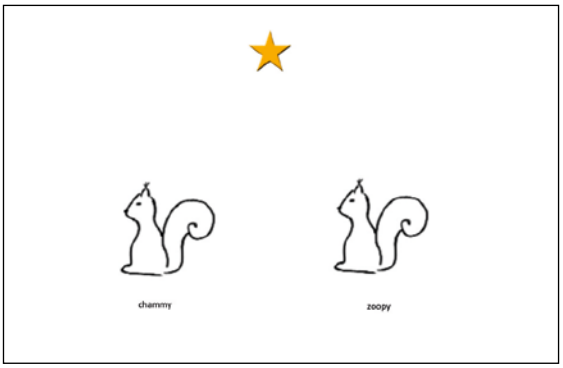


FIGURE 27

This slide shows two squirrels. These words below them are descriptions, not names. They are novel adjectives that are meant to describe these squirrels. You show participants this kind of scenario [The right squirrel moves to the star], and they would say *The zoopy squirrel moves to the star*.

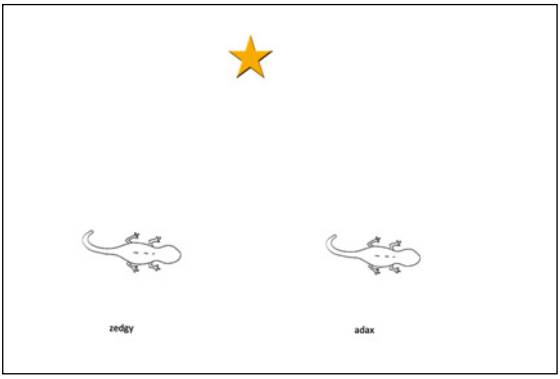


FIGURE 28

Here is another example of the experimental stimuli. The slide shows two lizards, and one of them moves to the star [The right lizard moves to the star]. Here speakers might say *The adax lizard moves to the star* or *The lizard that is adax moves to the star*.

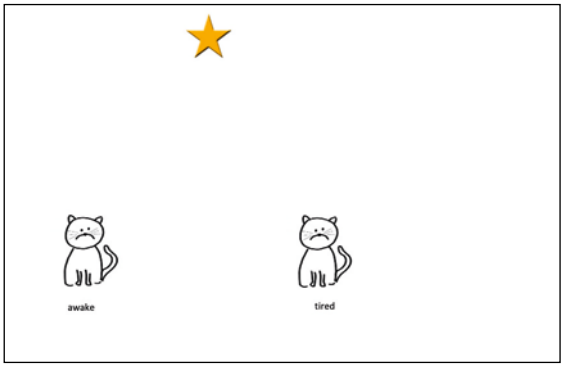


FIGURE 29

Consider this final example. This slide shows two kittens. When the left of them moves to the star, speakers would say *The kitten that is awake moves to the star*.

The adjectives that Boyd and Goldberg (2011) used in this experiment have certain characteristics. The stimulus with the kittens features one of the adjectives that I discussed earlier, so-called *a*-adjectives like *asleep* or *awake* or *alive*. They can only be used predicatively in English. It is only possible to say *The kitten that is awake moves to the star*, but you cannot say **The awake kitten moves to the star*.

The most interesting type of stimulus in Boyd and Goldberg's (2011) experiment is illustrated by the one with the two lizards. Do the participants say *The adax lizard moves to the star*, or do they recognise *adax* as one of these

a-adjectives that have to be treated in a special way? Do they prefer to say *The lizard that is adax move to the star*? If they do, that means that they have formed a generalization on the basis of other *a*-adjectives that they have heard, and they project the constraints of that *a*-adjective construction to new adjectives that they haven't heard so far. As a consequence, they would avoid **the adax lizard*, because they assume it to be ungrammatical. Think about that. They have intuitions about ungrammatical uses of a word that they have never heard before. They have learned not to say certain things, despite the fact that they do not have any active evidence apart from the distributional knowledge of what they've heard so far.

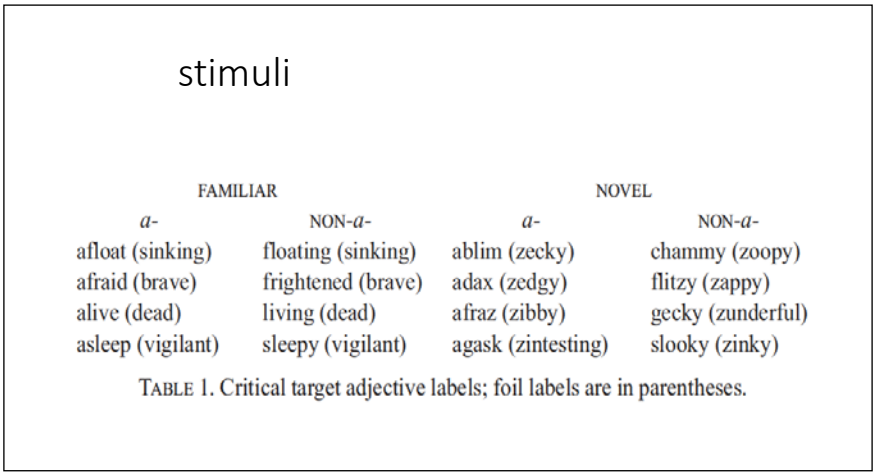


FIGURE 30

Here are the different types of stimuli that Goldberg and Boyd (2011: 66) used. The experiment cleverly cross-cuts ordinary non-*a* adjectives like *sleepy* with made-up artificial adjectives such as *chammy* or *zoopy*. The question is whether speakers treat a made-up adjective like *adax* like ordinary adjectives such as *sleepy*. Do they say *The lizard that is adax moves to the star*, or do they say *The adax lizard moves to the star*? They haven't heard the word *adax* before, so they might assume that it works like any other adjective, but the crucial conclusion is that they do not.

The main result of the study can be seen in the graph on the slide here (Boyd and Goldberg 2011: 69). The important piece of information is the height of the bars in this chart. The higher the bar, the more speakers in the experiments actually chose a description that involved the attributive adjective like *the active cow*. If the bars are low, that means that the speaker rather chose a relative clause construction like *the cow that is active*.

results

- main effect of adjective type
  - a-adjectives typically used with relative clauses
- main effect of familiarity
  - novel adjectives used more often attributively
- interaction of adjective type and familiarity
  - unfamiliar a-adjectives less often used with relative clauses than familiar a-adjectives

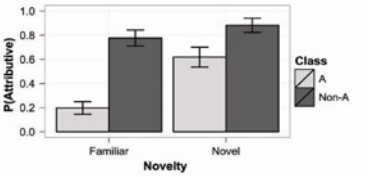


FIGURE 31

What you can see is that there is a major difference between non-*a* adjectives in dark gray, which have a large proportion of attributive uses like *the active cow*, and *a*-adjectives in light gray like *awake* or *asleep*. In the familiar condition, very few of these are used before the noun. There are some people who say *the awake kitten*, and but not many. The crucial category is shown by the second light gray bar. These are the adjectives like *adax*. You see that speakers are influenced by the presence of an initial *a*. They say *the adax lizard* less often than could be hypothesized. To them **the adax lizard* does not sound quite right, and this shows that they have generalized from *awake* and *asleep* to a new word, namely, *adax*. They have learned not to say a certain thing.

In the context of statistical preemption, there is one further controversy that I would like to mention, which concerns the way in which statistical preemption is supposed to work. I would like to discuss two competing accounts, one by Adele Goldberg herself and one proposed by Anatol Stefanowitsch. Stefanowitsch uses the term 'negative entrenchment' for his account. What is the difference between those two? Goldberg's (2019) position is that statistical preemption works in such a way that speakers reject the creative use of a construction when there is an alternative that they know about. When speakers know a conventionalized and alternative expression, they won't get creative. Speakers know that *the child that is afraid* works fine. They have never heard the phrase **the afraid child*, which would be a conceivable alternative, and as a consequence they shy away from using *a*-adjectives before the noun. Likewise, the speakers have experienced the verb *want* with a *to*-infinitive complementation

pattern lots of times, as in *She wanted to say something*. By contrast, *want* with an *-ing* type complement, as in **She wanted saying something*, is never encountered, despite the fact that it would be a possible alternative. The asymmetry between these alternatives is what leads speakers to disprefer **She wanted saying something*. For Goldberg, it is crucial that there is competition between two alternatives constructions that mean approximately the same thing.

Stefanowitsch (2011) argues a point that is subtly different. For him, learning not to say a certain thing is not necessarily due to competing alternatives. Rather, negative entrenchment for Stefanowitsch works in such a way that speakers reject the creative use of a construction when they have heard that construction frequently in other contexts, but never before in the creative one. *Say* famously does not work in the ditransitive construction. According to Stefanowitsch, that is because *say* is a very frequent verb. It occurs in many constructions, but the speaker has never heard it in the ditransitive.

With regard to this controversy, I am actually happy to let you know that I am currently involved in experimental work together with Adele Goldberg where we try to test the merits of both points of view. I do not think they are mutually exclusive. I am convinced that statistical preemption works in a way that Goldberg proposes, but it remains to be seen if negative entrenchment also works.

The last idea I'd like to present this morning is called constructional contamination. What is it? It is

> an effect whereby a subset of instances of a target construction is affected in its realization by a contaminating construction, because of a coincidental resemblance between the superficial strings of instances of the target construction and a number of instances in the contaminating construction.
>
> PIJPOPS AND VAN DE VELDE 2016: 543

Let me unpack that definition and explain how constructional contamination works in practice. Constructional contamination can affect a target construction such as the English passive. Here's an example sentence of the passive, *the disease was sexually transmitted*. We have a participle and an adverb in this particular sequence. The adverb comes first, the participle follows. This target construction may be influenced by a contaminating construction that is superficially similar. In the case of the passive, a potentially contaminating construction is a noun phrase construction, where we see the exact same order of adverb and participle, first *sexually*, then *transmitted*. The sequence

hypothesis: high frequencies in the NP >> preference for ADV-PPART

adverb	participle	passive (ADV-PPART)	passive (PPART-ADV)	complex modifier NP
well	known	**1594**	4	110
best	known	**957**	7	212
also	found	**608**	12	169
widely	used	**501**	55	11
randomly	assigned	**444**	33	39
often	called	**407**	1	93
also	included	**283**	2	205
sexually	abused	**201**	5	74
privately	owned	151	6	**290**
publicly	traded	60	2	**180**
well	established	**376**	3	31
highly	regarded	122	2	**265**
dimly	lit	64	1	**297**
randomly	selected	**308**	38	13
clearly	defined	105	6	**241**
democratically	elected	42	12	**284**
hard	hit	91	**245**	1
specifically	designed	**187**	**136**	8
better	prepared	**331**	1	1
often	seen	**310**	3	2

Table 2: Twenty frequent adverb-participle collocations

FIGURE 32

of adverb and participle appears in two different syntactic contexts. One may influence the other. What is crucial is that the target construction, the passive, allows variation. Speakers of English can say *The disease was sexually transmitted* or *The disease was transmitted sexually*. Both are grammatically possible, but in the contaminating construction, there is no variation, only the adverb initial order is possible.

A relevant question to ask is whether frequent usage of *sexually transmitted* in the noun phrase construction can lead to a relative preference of that order, adverb first and then participle, in the passive. This can be tested empirically.

In work that I have done together with Susanne Flach, we have examined corpus data to check whether high frequencies of an adverb participle combination in the noun phrase construction correlate with the preference for adverb initial order in the passive. Our results indicate that combinations that occur frequently in the noun phrase will have a contaminating effect on the passive. For example, the combination *privately owned* is a combination that is very frequent in the noun phrase construction, and when we compare the frequencies in the passive, we see that there is a strong asymmetry. *Privately owned* is much more frequent in the passive than *owned privately*. Speakers frequently say *The company is privately owned*, but the alternative *The company is owned privately* is a lot less frequent. You can explain that in terms of this combination of the noun phrase construction, which contaminates the use of the passive construction.

To conclude, data from the English passive offers support for the idea of constructional contamination, which means that frequent collocations in one

construction can influence syntactic variation in another syntactically homonymous construction. This means that in the speaker's knowledge, syntactic schemas are connected, and superficial structural similarities are enough for speakers to form and entertain connections between constructions.

I am coming to an end. This morning, I have given you ten basic ideas of construction grammar, which I think sum up the enterprise. I have given you a number of strategies that allow you to identify constructions. I have talked about five controversies where architects of construction grammar are currently debating how we should think of certain notions. Finally, I have outlined a couple of new developments. Researchers in Construction Grammar are detecting the limits of constructional productivity with the notions of coverage and statistical preemption, and they are detecting patterns of association through evidence of constructional contamination. There is one new development that will form the backbone of every lecture that follows from now on, namely, the application of constructional approaches to the study of language change, which is what we do in Diachronic Construction Grammar.

Starting with Lecture 2, I will thus focus on language change. I look forward to seeing you this afternoon. Thank you very much for your attention.

LECTURE 2

Taking a Constructional Approach to Language Change

Welcome back to Ten Lectures on Diachronic Construction Grammar. In the last lecture, I have presented Construction Grammar as a cognitive approach to linguistic knowledge. In this lecture, I will try to show how these ideas can be applied to the study of language change. As the title of the lecture suggests, I will be taking a constructional approach to language change.

The central question that I want to begin to answer in this lecture is why and how a constructional approach to language change might differ from other ways of studying how language develops over time. What is special about Diachronic Construction Grammar? Why should you be interested in it? Does it allow us to see and understand things that we would not be able to understand otherwise? What are the advantages that a constructional approach can bring us? I want to address these questions by going back to the ten basic ideas of Construction Grammar that I started out with in the previous lecture. The basic ideas that are foundational for Construction Grammar by implication also form the basis for any approach that takes Construction Grammar into the diachronic domain. Idea #1 is that all of linguistic knowledge is a network of form-meaning pairs. For diachrony, this implies that we will be thinking about language change as change in that network of constructions. What happens to the network of constructions when one construction or several constructions change?

Idea #2 is that the basic unit of linguistic knowledge are symbolic form-meaning pairings. That means that speakers today know constructions that are different from the constructions that were used by speakers of earlier generations. Our knowledge of form-meaning pairings is different from the knowledge of speakers that lived generations ago.

Idea #3 is the observation that constructions vary in terms of complexity and schematicity. You remember the quote by Langacker that I gave you in that

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13690978.

context. Sociolinguists have been telling us for a long time that variation and change are really two sides of the same coin. One question we can ask with regard to diachrony is how constructions change in terms of complexity and in terms of schematicity. Here we are really entering the realm of questions that have been asked in grammaticalization studies. How does complexity build up in languages? How do highly schematic constructions and syntactic patterns develop? I will talk about these questions.

Idea #4 relates to the fact that constructions are often idiosyncratic, unpredictable and riddled with exceptions. How does this unpredictability come about? What about the opposite, that is, the tendency for irregular forms to become regularized through analogy? On average, do constructions become more regular or more irregular as time goes on?

Idea #5 is the claim that all constructions carry meaning. For diachrony, this raises the interesting question of how syntactic patterns like the ditransitive construction acquire their meaning historically and what happens to their meaning over time.

Idea #6 is the principle of coercion, which states that constructional meaning wins out over lexical meaning. From a diachronic perspective, it would be very interesting to investigate when and how a construction starts to bring about coercion effects. When, for example, did it become possible in the English language to turn mass nouns into accountable units via the plural construction? You remember the example *Two beers, please*.

Idea #7 states that grammatical categories are the outcome of speakers generalizing over instances of language use. That has strong implications for diachrony as well. Concrete instances of language use, taken from diachronic corpus data, should be able to reflect the emergence and the development of grammatical categories. I have engaged with this idea in my work, and we'll see a number of examples in later lectures.

The same is true for idea #8, the notion that constructional meaning is reflected in associations between syntactic patterns and lexical elements. By implication, diachronic shifts in such patterns of association should be indicative of semantic change, and this idea has been central to my own research.

Idea #9 is that knowledge of constructions is usage-based. This idea has the implication that as new instances of a construction are produced, the representation of that construction changes as well, leading to potentially further uses that then let speakers repeat the cycle. This brings us to the role of diachronic corpus data and the analysis of such data, which will be an important part of the next lectures.

Finally, idea #10 is that language draws on domain-general social cognitive processes, like categorization or joint attention. If we take that point seriously,

it means that all language change is due to the cognitive and social pressures that are at work in the here and now. The question is, how do these pressures bring about long-term changes, such as the grammaticalization of constructions? That is a puzzle that I want you to appreciate. There are forces acting on language in the here and now and they have to be the locus of language change. But these changes accumulate to developments that are much longer than the life of a single speaker. How is that possible? How does that work?

The basic ideas of Construction Grammar already provide a rich foundation for the constructional study of language change that makes us see new things and ask new questions. I will come back to these ideas one by one during the next lectures.

Let me start, however, by saying something general about constructions and language change. I have been lucky to be part of a vibrant community of researchers who have been interested in Diachronic Construction Grammar. Over the past ten or so years, more and more work has adopted a constructional perspective on language change. These studies focus on form-meaning pairings and their diachronic developments. The alignment between historical linguistics on the one hand and Constructional Grammar on the other has become increasingly popular, and even though I welcome that development, it has always puzzled me a little bit. Construction Grammar represents a very different tradition than historical linguistics. It adopts a synchronic perspective, it takes a cognitive, mentalist stance, and that is not the case for many historical approaches. Work in language change always goes beyond a steady state. It does not just describe a speaker's knowledge at a certain point in time. Historical linguistics also goes beyond the confines of a single human mind. If we want to talk about regularities and how languages change over decades, centuries, perhaps even millennia, then we are making generalizations that go beyond what happens in any single speaker or any single human mind. Even though I have always been convinced that Diachronic Construction Grammar is a fascinating approach that has its justification, I have found it surprising that it turned out to be as popular as it has. What is so attractive about it?

I can of course only speculate, but the best answer that I can give brings us back to basic idea #4 that I outlined earlier, the observation that constructions are typically unpredictable and idiosyncratic. Nobody, I think, is more aware of the unruliness of language than historical linguists.

Analysts of language change are very much aware that historical developments often are unpredictable and idiosyncratic. In fact, there is a very nice quote by Paul Hopper and Elizabeth Traugott (2003: 131) illustrating that very point. In their textbook on grammaticalization, they state that

> There is nothing deterministic about grammaticalization and unidirectionality. Changes do not have to occur. They do not have to go to completion. In other words, they do not have to move all the way along a cline.

With regard to idiosyncrasies, Hopper and Traugott were construction grammarians all along.

The constructional way of thinking about language resonates with ideas that historical linguists had already entertained, specifically if they were working within a broadly functional framework. For them, it was more or less self-evident that language change is the sum of many constructions changing individually, often in unpredictable ways. This contrasts with structuralist and generative perspectives, in which language change is seen as catastrophic and systemic, so that one change triggers another until the entire system revolves and changes. I refer to the work of Bloomfield (1933) and Lightfoot (1979, 1999) in this context.

In contrast to these views, one might actually adopt the opinion that every construction has its own history, which in this strong form is probably not true. It is tempting to think along these lines, but it is perhaps not quite what the constructional view implies. Remember that knowledge of language is conceived of as a network of symbolic units. Due to interconnections in the constructional network, changes in one construction can be thought to bring about changes in related constructions, but how exactly that works is something to be figured out. A more nuanced view of this could be captured by a slogan that I have tried to popularize. What I said was "Grammatical change is not a zero-sum game" (Hilpert 2013: 4). When you pinch the system on one end, it does not always extend at the other end, or vice versa. Linguistic systems are fluid and have some tolerance. Changing one part might have consequences, but it is not a zero-sum game. You have to maintain a kind of balance, a system where everything holds itself in place as the structuralist notion has it. Change on the constructional view is not always systemic. One construction's success does not have to come at the price of another's demise, but changes typically relate to one another, and changing constructions influence one another.

Before I work out in more details what a constructional theory of language change looks like, I would like to point out a few issues that such a constructional theory does not cover. Let me tell you what it is not. There are several well-known phenomena in language change that do not, in my view, lend themselves particularly well to a constructional analysis, principally because they systematically affect many or even all constructions of a language. The prime example of such a phenomenon would be regular sound change that affects all words in the language. If you have a change where all long /e:/ vowels

turn into long /i:/ vowels across all the words of a language, you could describe that across all individual constructions that are affected, but in doing so you would miss the larger generalization that is there.

Another example would be a massive loss of morphology due to language contact. If two languages are coming into contact and large numbers of learners acquire these languages, some of their morphological complexity is going to disappear. That is not something that would be specific to any one or two constructions, but rather, this happens across the board. Another example in English is there has been a syntactic change from head-final to head-initial across different phrase types. Old English used to be head-final in verb phrases and in auxiliary phrases. In Present-Day English, these phrases are head-initial. This development is best accounted for by a generalization that affects more than just one construction.

Another example is sociolinguistic change in response to extralinguistic developments, as for example dialect levelling in areas with increased speaker mobility. What happens to the local variety once speakers are very mobile? These phenomena, I would argue, capture generalizations that hold across many different constructions. We could apply a constructional approach, but we would miss more insightful, broader generalizations.

Diachronic Construction Grammar focuses on the developmental trajectories of individual constructions where this is useful. This of course has been the focus of another theoretical approach to language change, namely grammaticalization theory. I would like to say a few words about that approach.

What is different between Diachronic Construction Grammar and grammaticalization theory? These two frameworks have a lot in common, and they are adopted by overlapping communities of researchers. Nonetheless, I find it useful to consider for a moment how the two frameworks differ from each other and what their respective aims are, because they are not quite identical.

I take it that many of you in this room are broadly familiar with grammaticalization as a theory of how closed-class elements come into being. For my purposes, I adopt the definition of grammaticalization that has been proposed by Paul Hopper and Elizabeth Traugott (2003), who formulate it as

> the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions.

My understanding of grammaticalization further owes a lot to Christian Lehmann's work. Lehmann (2015: 15) conceives of grammaticalization as a progressive development towards ever more compact linguistic structures.

Structures that are only loosely connected in discourse become more tightly integrated through syntacticization. Syntactic structures have a tendency to fuse together through morphologization. Morphological structures blend into one another to form synthetic structures. Eventually, parts of these structures may reduce to zero. The general appeal of grammaticalization theory, as I see it, is motivated by two factors. On the one hand, grammaticalization theory states broad empirical generalizations that account for lots of cases across many different languages. On the other hand, it makes testable predictions for data that we may come across in the future. This is already a point that sets grammaticalization theory apart from Diachronic Construction Grammar, which at this point has not been able to generate a system of testable hypotheses in quite the same way.

I have said that I view Diachronic Construction Grammar and grammaticalization theory as largely overlapping, but as not completely coextensive. There are reasons to say that grammaticalization has a narrower scope than Diachronic Construction Grammar. Specifically, there are patterns of lexicalization and lexical-semantic change that we would subsume under Diachronic Construction Grammar, but that are outside the scope of grammaticalization.

For example, there are some processes that never happen in grammaticalization, but that are common in lexical-semantic change, i.e. semantic narrowing. The English word *meat* used to mean "food in general". In Present-Day English it has narrowed down to mean "animal flesh". There is further the phenomenon of amelioration in semantic change. The English adjective *nice* meant "foolish", now it means something like "pleasant". It has acquired a more positive meaning. That is not the kind of meaning change that you see in grammaticalization.

In many definitions of grammaticalization, there are differences, but many definitions exclude word order changes, which would of course instantiate change in the constructional network. One example of this concerns the loss of English V2 constructions, another concerns changes in argument structure, specifically the diachronic increase of transitive structures in English. All of these examples suggest that the linguistic changes that grammaticalization focuses on form a subset of those that Diachronic Construction Grammar is concerned with.

However, that is not the whole story. You can also make the opposite case, arguing that some aspects of grammaticalization go beyond changes in individual form-meaning pairings, so that grammaticalization could be said to have a wider scope than Diachronic Construction Grammar. Let me give you two examples to illustrate this.

One example comes from the work of Christian Lehmann (2015), and it pertains to what he calls paradigmatization, which is the tendency of

grammaticalizing constructions to form paradigms or to integrate into already existing paradigms. That, if you like, is a generalization about developments that affect groups of constructions. Grammatical domains like case, person, number or tense tend to recruit a small group of closed-class elements into their service, and then these elements tend to express semantic oppositions, and they tend to converge in terms of their morphosyntactic behavior. This happens in similar ways across different grammatical domains that have different formal expressions across different languages. In other words, to say that we frequently observe paradigmatization in language change is to express a meta-generalization about how groups of constructions tend to change over time. It is broader than analyzing the developmental trajectory of a single construction or group of constructions.

There is another example that comes from a very different theoretical background. Ian Roberts and Anna Roussou (2003) have developed a generative approach to grammaticalization, which is in many ways opposed to Lehmann's work. One aspect in which it runs counter to the Lehmannian view is that Roberts and Roussou view syntactic scope increase as definitional for grammaticalization. What they say is that grammaticalization involves syntactic reanalysis that assigns the grammaticalized form to a higher node in the syntactic structure.

An example for this would be the grammaticalization of lexical verbs into auxiliary verbs. When lexical verbs become auxiliaries, they are assigned to an operator position that sits up a little bit higher in the syntactic tree. Now you do not have to agree with any particular theoretical model of syntax to appreciate the generalization that is at stake. Across several different grammatical domains, across different construction types, we observe scope increase, and that will be a formal generalization that reaches across individual constructions and that expresses a more general property of grammaticalization. Both the example of paradigmatization and the example of scope increase capture generalizations across many different constructions.

All of these differences suggest that grammaticalization theory and Diachronic Construction Grammar are not quite the same, but it remains a given that grammaticalization theory has been gravitating towards a constructional perspective over recent years. For example, when we take Hopper and Traugott's definition of grammaticalization that I mentioned earlier, we see that there is an interesting difference between the 1993 edition of their textbook and the version that came out ten years later, in 2003. The 1993 version embodies what we could call the "item-based view". There the definition states that "*grammaticalization is the subset of linguistic changes through which a linguistic item in certain uses becomes a grammatical item, or through which*

a grammatical item becomes more grammatical". This is the item-based view. The 2003 version has been updated to reflect the status of constructions: "*the change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions*".

The shift towards a constructional view has also been commented on by Joan Bybee (2003), who has the following to say about it:

> The recent literature on grammaticalization seems to agree that it is not enough to define grammaticalization as the process by which a lexical item becomes a grammatical morpheme, but rather it is important to say that this process occurs in the context of a particular construction.

To sum this up, what is reflected in the changing definitions of Hopper and Traugott and in the quote by Bybee is that there is an increasing focus on changes that affect the syntagmatic axis of language during grammaticalization. This explains in part why grammaticalization theory has been aligning with Construction Grammar.

There are, however, further differences that I find to be considerable. One of them being the fact that grammaticalization theory makes testable predictions. I have been mentioning that fact. One of these predictions concerns what's called "unidirectionality". That is the idea that changes proceed in a very constrained way that is irreversible. Let me give you a non-linguistic example for unidirectionality. This morning at breakfast I had a bowl with yogurt with a little bit of jam on top. If I stir the yogurt with a spoon, it will mix with the jam until I have a fairly homogeneous mixture. Let us say that I have been moving the spoon towards the right. If I take the spoon and turn it back to the left three times, I won't get my jam back. The mixing process is unidirectional, and grammaticalization theory holds that many processes of language change are actually like mixing yogurt with jam. They go into one direction, but not in the opposite one. Generalizations like the hypothesis of unidirectionality are one example where grammaticalization theory has a wider scope than Diachronic Construction Grammar, because the hypothesis applies to a broad range of constructions, not just a single one.

Diachronic Construction Grammar is concerned with many changes that are in fact bidirectional. For example, English gives its speakers the possibility to use verbs as nouns and nouns as verbs. There is the verb *run* and "*I can go for a run*". There is a noun *butter* and "*I can butter a slice of bread*". Another example would be analogical change. Frequently analogical change turns an irregular form into a regular one. The verb *weep* in English forms the past tense with an irregular form *wept*, but you will find it used in a regularised way where

speakers opt for *weeped*. Importantly, analogical change does not always target the regular form that has the highest type frequency. Sometimes the target is a smaller class with a few highly salient members. The verb *sneak* is regular, but speakers started to use an irregular past tense form *snuck*, using an analogy with verb forms such as *strike* and *struck*, *stick* and *stuck* as salient members of this irregular category. The third example is that in lexical semantic change, semantic narrowing coexists with semantic broadening. Semantic narrowing is illustrated by the example of *meat*. The converse process is widening. The English noun *dog* referred to a specific breed of dog, now it refers to the entire species. In grammaticalization, we do not regularly see semantic narrowing, as items usually extend towards broader, more abstract meanings.

By contrast, in grammaticalization, developments are supposed to go in one direction only. This can be illustrated, for example, with the development of affixes that turn from independent words into structures that are dependent on a host structure. The English regular past tense, written as *-ed*, derives from a formally independent verb form with the meaning "did". Another example, the adverbial suffix *-ly* in *friendly*, derives from an independent word meaning "body" that acquired the meaning of similarity and which ultimately turned into the suffix that we are using today.

The hypothesis of unidirectionality states that independent elements lose in formal and semantic substance and turn into dependent elements, not the other way around. That is why it is interesting to pay close attention to cases that seems to go against the overall tendency. There is an example that you are perhaps aware of, namely the use of the English suffix *-ish* as an independent word. If I want to say that some activity took me about two hours, I can say that "*It took me two hours ish*", meaning about two hours. Grammaticalization scholars would see this as an anomaly. It is not supposed to happen, but every now and again it does happen. Despite these counterexamples, grammaticalization theory incorporates the unidirectionality hypothesis as a way of making predictions about unseen data.

In the framework of Lehmann, unidirectionality does not only apply to the development of suffixes out of independent words, but it actually reaches across a set of related properties of language. Lehmann identifies six separate unidirectional processes that I will briefly present.

Erosion means that as forms grammaticalize, they lose in substance and they become shorter. Condensation means that grammaticalizing forms shrink with respect to their syntactic scope. The suffix *-ly* used to be an independent word, but now it just forms part of an adjective. The process of paradigmatization captures that as forms grammaticalize, they integrate themselves into a group of grammatical forms with similar properties.

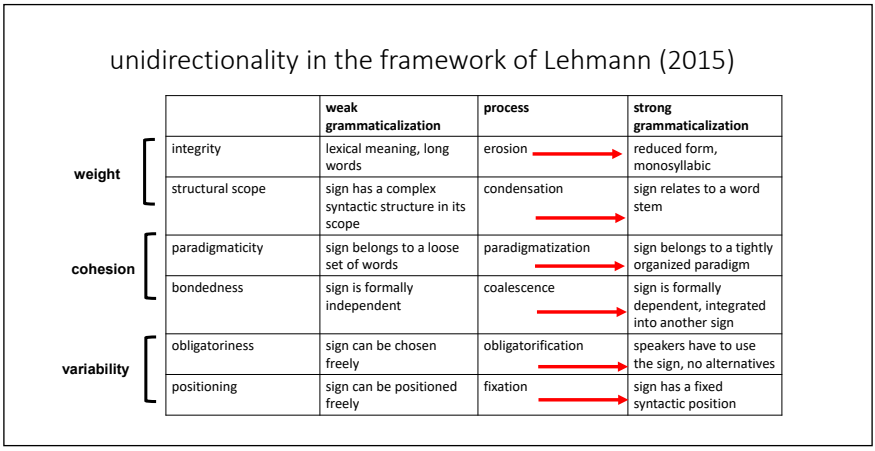


FIGURE 1

Coalescence refers to the increasing dependence of a grammaticalizing form to a host structure.

Obligatorification means that strongly grammaticalized signs have to be used as a matter of convention. For example, in languages that have articles that are definite and indefinite, the speaker has to pick one, depending on the context. The speaker no longer has the freedom to include the article or leave it out.

Finally, fixation means that as a form grammaticalizes, speakers become increasingly constrained with regard to the position in the utterance where a sign can be used.

Why am I going through all of this? My general point is that grammaticalization theory proposes this elaborate system of interlocking continua, which are tied to very specific and strong empirical predictions. We expect language change to proceed along these lines, but not in the opposite direction. Diachronic Construction Grammar, by contrast, has up to this point not been able to generate a similar set of hypotheses that could be put to the test in a systematic fashion.

To bring my juxtaposition of the two frameworks to a close, what I want you to take away is that I see the two as closely related enterprises that show substantial overlap, but that also each have characteristics that are respectively their own.

With all of this in mind, let me now outline the project of Diachronic Construction Grammar in more positive terms. I will take as my starting point the basic idea that linguistic knowledge is to be conceived of as a large structure network of form-meaning pairings. I would like to advance the view that

Diachronic Construction Grammar is the study of changes happening in that network. Knowledge of language, from the view of Construction Grammar, is a network of constructions. Language change, from the view of construction grammar, would be change that happens in that network.

In the next part of this lecture, I want to go over four aspects of that kind of change. First, how new constructions emerge or disappear. Second, how existing constructions change in form and meaning. Third, how links in the network emerge or disappear. Fourth, how existing links in the network becomes stronger or weaker. I am going to start with the emergence of new constructions, which is undoubtedly what has captured most of the attention of researchers working in Diachronic Construction Grammar.

Elizabeth Traugott and Graeme Trousdale (2013) have created a technical term that captures the emergence of constructions. The term is "constructionalization", and it is defined in the following way:

> constructionalization is the creation of form_{new}-$\text{meaning}_{\text{new}}$ (combinations of) signs. It forms new type nodes, which have new syntax or morphology and new coded meaning, in the linguistic network of a population of speakers.

This means that a new symbolic unit is coming into being, but there is one fairly important addition, and that would be that "*formal changes alone, and meaning changes alone cannot constitute constructionalization*". We cannot make new symbols by adding just one part of their structure. It has to be both form and meaning. To help us understand this concept a little better, let me try to break down the definition into its component parts.

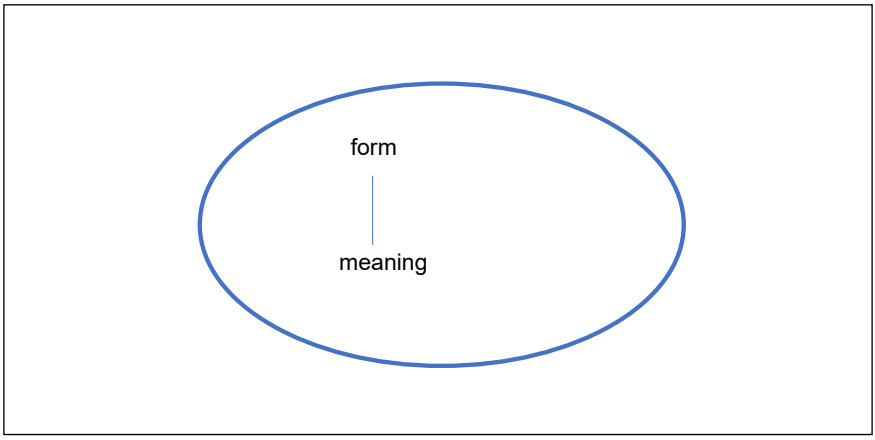


FIGURE 2

We start with a construction, a pair of form and meaning.

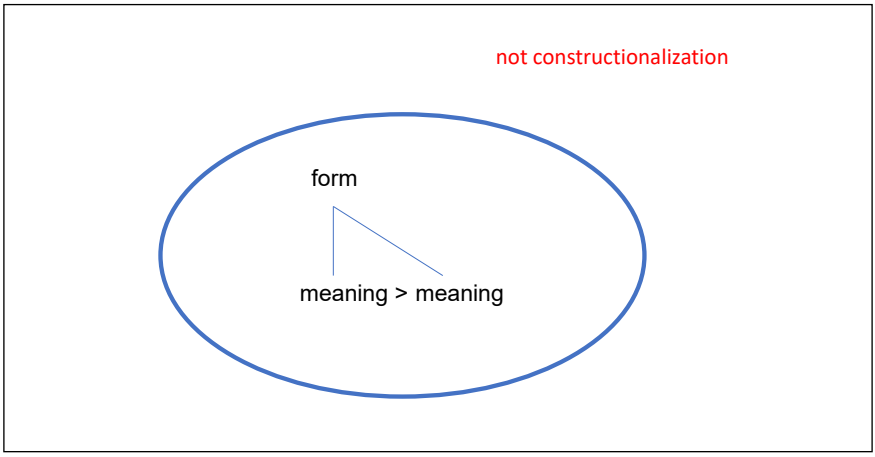


FIGURE 3

In the first step, the meaning may become extended to a second meaning, still associated with the same old form. For example, a lexical item may be extended to cover new semantic territory. An adjective such as *sweet* no longer refers to just to taste, but also to an emotional quality. When I say '*That was so sweet of you*', we'd have a new meaning of *sweet* attached to the same form. According to Traugott and Trousdale, that is not constructionalization. That is a semantic change.

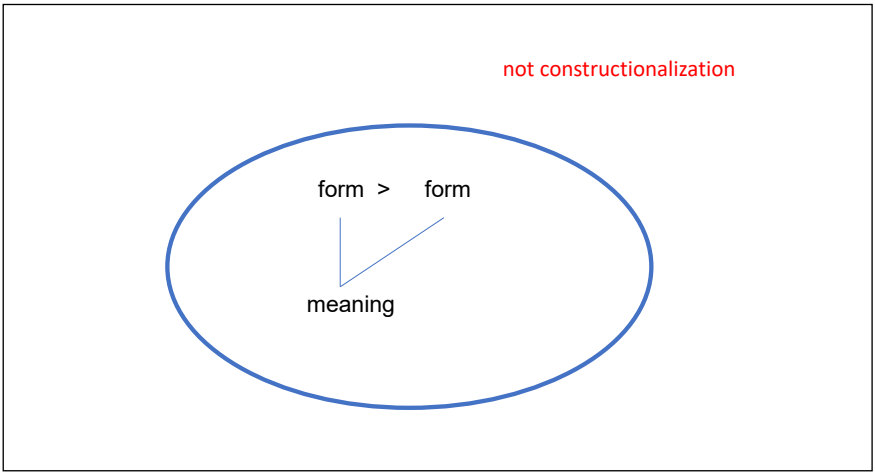


FIGURE 4

Conversely, let's say that we have a form-meaning pair, one form, one meaning, and then a new variant of the form develops. It could be a shortened pronunciation. The English word *family* /'fæmili/ is quite often shortened down

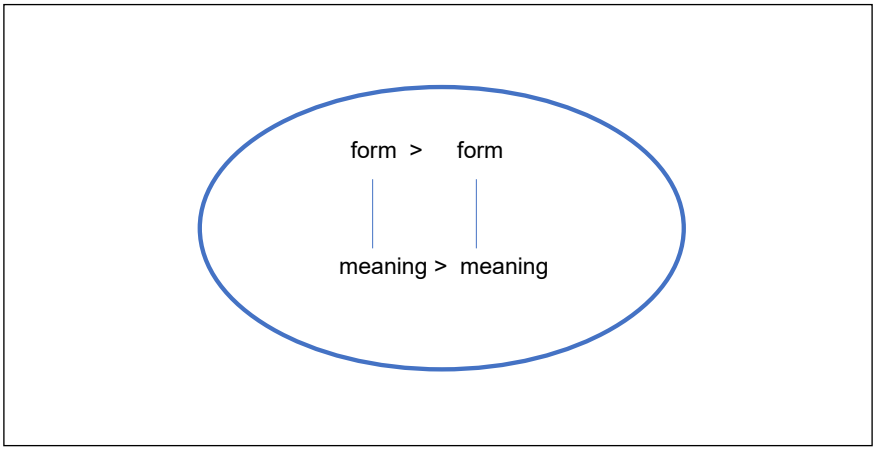


FIGURE 5

to *family* /ˈfæmli/. The first variant has three syllables, the second one has only two syllables. According to Traugott and Trousdale (2013), that would not be constructionalization.

Now suppose that we have a form-meaning pair and that pair develops in such a way that at some point a new form develops, and simultaneously, there is also a new meaning that develops. Eventually, speakers come to perceive and use the pairing of the second form and the second meaning as a form-meaning pair of its own, separate from the first form-meaning pair.

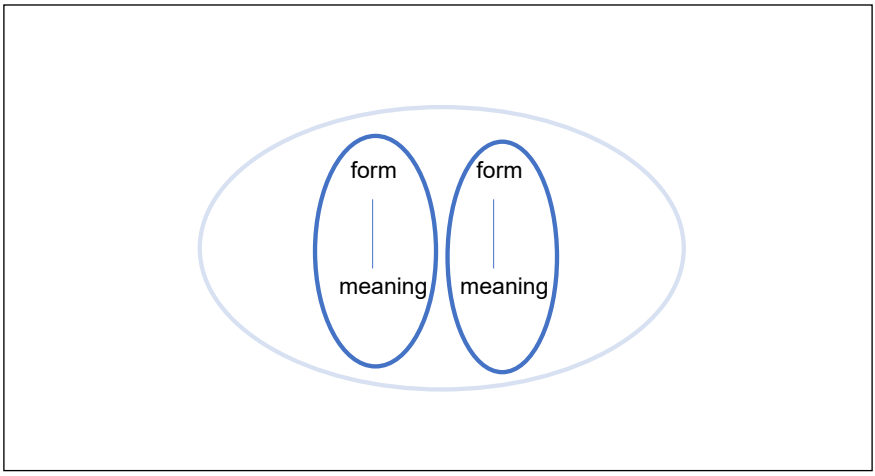


FIGURE 6

That case represents what Traugott and Trousdale (2013) would call constructionalization, the emergence of a new form-meaning pair in the network of constructions that constitutes a speaker's knowledge of the language.

There have been proposals to the effect that this process, constructionalization, should be equated with grammaticalization. Dirk Noël (2007) was one of the earliest researchers to talk about Diachronic Construction Grammar as a theoretical framework in its own right. What he says is this:

> In Construction Grammar constructions are by definition grammatical, so that the historical emergence of constructions amounts to becoming part of the grammar, and what better term to denote this than grammaticalization.

I am afraid that I disagree with this proposal, since I find that it rests on a somewhat unfortunate interpretation of the term "grammar". What is true is that Construction Grammar as a theory attempts to model speaker's knowledge of the language in total. In that sense, grammar could be a term that stands for everything that a speaker knows, i.e. all of a speaker's knowledge. Still, I do not think it is helpful to say that in Construction Grammar, constructions are by definition grammatical. Not all constructions are grammatical. That term should be reserved for constructions that are advanced on the clines that grammaticalization research has worked out. Articles are grammatical because they are highly dependent on a host structure and because their use is obligatory. Relative clauses are grammatical because they are syntactically complex and convey a very schematic kind of meaning. Lexical words like *bicycle* or *lecture* or *bottle* are constructions, but they are not grammatical constructions. They're lexical constructions.

In their 2013 book on constructionalization, Traugott and Trousdale make a distinction between two different types of constructionalization, one which they call lexical or contentful constructionalization, and another type that they call grammatical or procedural constructionalization. Lexical constructionalization refers to the coinage of new lexical items, such as *photobomb*, *twitterverse* or *Brexit*. Lexical constructionalization typically starts with the instantaneous creation of a new form which is then gradually propagated in the speech community and which conventionalizes through usage over time.

By contrast, grammatical constructionalization concerns the emergence of new grammatical constructions. In English, examples would include the emergence of the passive with the verb *get*, as in "*It is ok as long as you do not get caught*", or the double-*is* construction, "*The problem is is we are out of money*", or what's been called contrastive reduplication, "*Does he like me or does he like-like*

me?" Grammatical constructionalization bears all the features of grammaticalization that I have described earlier. The constructions involve grammatical dependencies. They encode abstract meanings. They arrange themselves into paradigms.

Traugott and Trousdale (2013) identify three aspects of constructionalization that they view as central, and that allow them to distinguish between lexical constructionalization and grammatical constructionalization. These three aspects concern the compositionality of a construction, the schematicity of a construction and the productivity of a construction. For example, when a new passive construction such as the *get*-passive undergoes grammatical constructionalization, there is a decrease in compositionality, so the verb *get* no longer means just "get". At the same time, there is an increase in schematicity. The overall construction is not just *get*, but it is rather *get* plus a slot for a verb in the infinitive. There is an increase in productivity, meaning that as time goes on, we find more and more lexical verbs that enter the past participle slot, that are used with *get* in order to form an instance of the *get*-passive. For grammatical or procedural constructionalization, we have increases in schematicity and productivity and a decrease in compositionality.

For lexical/contentful constructionalization, we have a different profile. Specifically, we have decreases for all three aspects, i.e. for compositionality, schematicity and productivity. Let's take the example of *Brexit*. *Brexit* is a blend from *Britain* and *exit*, but it encodes a very specific meaning of "Britain's exit from the European Union". There is no strict compositionality, and there is no schematicity, as the meaning of *Brexit* is quite specific.

Traugott and Trousdale (2013) use these notions as a way of capturing the broad difference between grammar and lexis, which is a distinction that I think is important, even if it is not crisp and categorical, but rather non-discrete and gradual. There are elements that are clearly lexical, like *dog* or *friendly*, which are contentful and which have specific meaning. There are grammatical elements such as determiners, pronouns, auxiliaries or the ditransitive construction, which clearly convey procedural meaning and which are discursively secondary.

Then, there are lots of in-between cases, for example, newly grammaticalized auxiliaries such as the verb *help* in "*help solve the problem*", and other examples that are not quite grammaticalized, but not quite lexical either. On this view, grammatical constructions convey a very specific type of meaning that can be called procedural. In the words of Traugott and Trousdale (2013), procedural meaning can be defined as follows: "*A grammatical sign cues how the speaker conceptualizes relationships between referents within the clause*". This captures notions such as subject and object and their functions within

a clause-level predicate construction. I have added a second definition here by Holger Diessel (2019), who defines the term in a slightly different way: "*Grammatical constructions provide processing instructions that guide listeners' semantic interpretation of lexical expressions*". This means that how our interpretation of a lexical item depends on the grammatical context. This actually brings us back to the principle of coercion that I mentioned earlier this morning. When I say "*Three beers, please*", the plural construction guides the listeners' interpretation of the word *beer*. Putting it all together, we can circumscribe procedural meaning as meaning that corresponds to questions like *Who did what to whom?*, *When did it happen?*, *How sure are we that it happened?*, and *What part of the event are we talking about?*

In all of this, you recognize bits and pieces of basic idea #3, Langacker's (2005) observation that constructions vary in terms of their degrees of complexity and schematicity. Procedural meaning is notably more schematic than the meanings that are associated with lexical material. Grammatical constructionalization in the sense of Traugott and Trousdale (2013) is concerned with the emergence of constructions in more abstract areas of linguistic structure that accommodate category schemas and constructional schemas. While this gives us a basic understanding of grammatical constructionalization, there is one more distinction that I need to introduce in this context, and that distinction concerns two different types of grammaticalization.

So far, I have talked about grammaticalization in terms that presented it as a tendency towards increasingly compact linguistic structures from discourse to syntax to morphology, and eventually to zero. This is commonly called the view of grammaticalization as reduction. Grammaticalizing forms lose their autonomy, their complexity, their syntactic freedom and their phonetic substance. This works very well for phenomena such as the creation of morphological affixes out of formerly independent words, or for the reduction of the *be going to* construction into *gonna*.

Grammaticalization as reduction is essentially the view of grammaticalization that is presented by Christian Lehmann, with its six unidirectional processes that lead to increasingly compact and compressed structures. All of that works very well for the structures it has been intended to deal with. But there are other phenomena that we might want to call grammaticalization, but that do not fit into this view.

There is a second type of grammaticalization running counter to the first one, and that type can be described in terms of expansion. The gist of the matter is that not all grammaticalizing constructions become more fixed and integrated, lose in semantic substance, and decrease in syntactic scope. Some constructions show the exact opposite behavior. There are three phenomena

that I briefly want to talk about, namely host-class expansion, the increase of syntactic scope, and semantic expansion.

The term host-class expansion signifies that a grammaticalizing construction over time increases its range of hosts, that is, the range of elements that co-occur with the construction. Let me start with the English *way*-construction, as illustrated by *He made his way through the room* or *He elbowed his way out of the subway*. This construction has historically come to be used with an ever-greater range of verbs. At first, it used to be restricted to verbs that relate to the laborious creation of a path. Now you can *cheat your way into law school* and *sing your way into the charts*. The construction become more open to different kinds of verbal predicates.

Another example are noun-participle compounds like *doctor-recommended*, *child-tested* or *chocolate-covered*. If we look at that kind of construction diachronically, we find that the number of participles occurring in that construction has been on the increase. Host-class expansion can also be observed on the syntactic level. The example that I can give you here are *it*-clefts in English, which used to be restricted to examples in which the focus phrase was a noun phrase: *It was the butler who killed them*. In present-day English, we have a number of other elements that can occur as the focus phrase, for example, prepositional phrases like *It is in May that she's coming* or *ing*-clauses, as in *It is eating broccoli that I just can't bring myself to do*.

Increase of syntactic scope is what we observe when a grammaticalizing construction comes to modify increasingly larger syntactic units. Lehmann predicts that the exact opposite should happen. Grammaticalizing units should decrease progressively in their syntactic scope, but we see the opposite with discourse markers that are based on adverbs, as for instance the word *actually*. *Actually* can be an adverb. I can ask *Is this measure actually necessary*, and in that sentence, *actually* has an adjective phrase in its scope. In a sentence like *They actually wanted to talk to you*, *actually* has a verb phrase in its scope. In its use as a discourse marker or sentence adverbial, it has an entire utterance in its scope. This would be the case for examples such as *Actually, this does not seem like a good idea*. *Actually* has progressively increased the size of the syntactic contexts over which it has scope.

The same goes for the clause connector *as long as*, which used to be just a modifier for a noun phrase, as in *We will do this for as long as a year*. It is expanded into contexts where it has scope over a clause: *As long as you keep it frozen, it will stay edible*.

The last phenomenon that I want to talk about with regard to grammaticalization as expansion is semantic expansion. Over the course of time, grammaticalizing constructions come to be used with an ever-greater range of

meanings. This applies to, for instance, the development of grammatical auxiliaries. Let's take the English modal auxiliary *may*, which is used in deontic and epistemic meanings. The example *You may now kiss the bride* expresses permission and thus deontic modality. *That may have been a mistake* expresses a logical possibility. *May* has expanded semantically over time.

The same goes for the example of *as long as* that I mentioned a minute ago. Originally, this refers to a time span. The example *as long as you keep it frozen* refers to the time during which you keep something frozen. When I say *As long as you have the money, you can come in*, I do not refer to a period of time, but rather I refer to the condition that you have the money. The shift from temporal to conditional meaning instantiates semantic expansion.

When we consider how Traugott and Trousdale define grammatical constructionalization, it seems that their view aligns closely with the view of grammaticalization as expansion, rather than Lehmann's view of grammaticalization as reduction. Let us look at a few concrete examples of grammatical constructionalization and their developments.

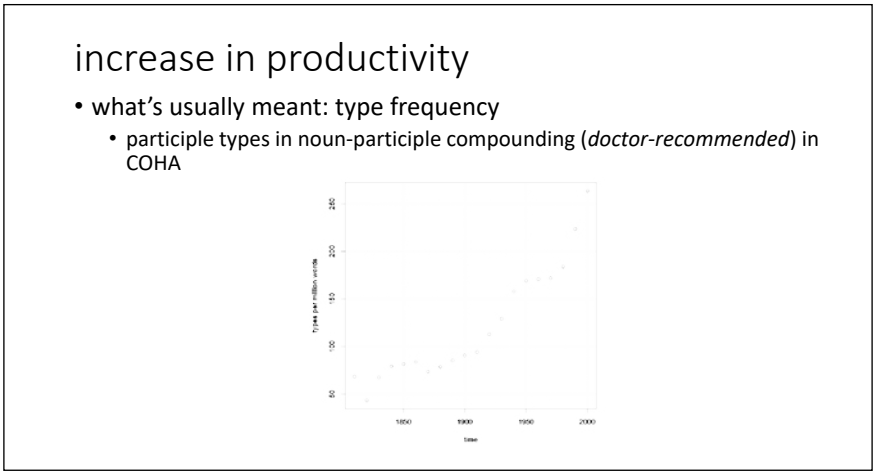


FIGURE 7

Let's start with the increase in productivity that happens during grammatical constructionalization. When Traugott and Trousdale (2013) discuss increases in productivity, they refer to increases in the type frequency of a construction. How many different lexical items are found in usage with a given construction? On this slide, you can see a graph with an increasing curve over time. That curve represents the growing number of participle types that are found in the English noun-participle compounding construction, as for example *doctor-recommended* or *kid-tested*. As time goes on, more and more different

participle types are included in instances of this construction, and this would represent an increase in productivity.

When Traugott and Trousdale (2013) discuss increases in schematicity, what they have in mind is that a construction acquires a meaning that is increasingly abstract. In the grammaticalization literature, this process goes by the name of semantic bleaching. There are many examples of this. For example, the English *be going to* future construction no longer necessarily encodes motion. We find it being used with inanimate subjects in utterances such as *Inflation is going to be a problem*. The example does neither convey intention nor movement. The construction has become more schematic in its meaning. Also, the example of *as long as* applies here, as it no longer just encodes time, but also a condition, as in *as long as you have the money*.

Decreases in compositionality mean that the meaning of constructions becomes less and less transparent. In other words, the idiosyncrasies or unpredictable characteristics of the construction are on the rise. The development starts with broadly compositional meanings. For example, in the expression *a bit of*, the example *He gave me a bit of bread* refers to a piece of bread that corresponds to something that you can bite off, a small chunk. When I say *I need a bit of sleep*, that is a short period, which is not exactly the same thing as a bite-sized object. When I say *That is a bit of a secret*, is that a limited part of a secret? Is that only secret-like in some ways? You see how the compositionality of the expression *a bit of* reduces over time and gives way to a more holistic meaning.

The same applies to the English *have*-perfect, which combines the verb *have* and a past participle. Early uses of the construction are used to express actual possession. The example that you see often used in this context is *I have the enemy bound*, which denotes that the enemy has been won over and is in the state of being tied up. When I say *I have read the book*, I still presumably have that book in my possession somewhere. I may have given it away, but it was in my possession at some point. But when I say *I have slept well*, is that period of sleep in my possession? Was it in my possession when I was actually asleep? That is debatable. What is clear, however, is that the compositionality of the *have*-perfect has over time become less compositional.

Grammatical constructionalization, according to Traugott and Trousdale (2013), is a process that involves an increase in productivity, an increase in schematicity, and a decrease in compositionality. Constructionalization would be the moment when all of these developments come together, and a new node appears in the construct-i-con. This node has to instantiate a new form-meaning pair, such that both the form and the meaning are recognized as new by speakers of the speech community. This brings up the question of how we should think about changes that happen to an existing form-meaning pair. For

instance, what about semantic expansion without formal change or phonological reduction without semantic change?

Dirk Noël (2007) pointed this out a long while ago in a quote that may have prompted Traugott and Trousdale to coin the term constructionalization:

> What Diachronic Construction Grammar has so far failed to do, however, is draw an explicit distinction between the initial formation of a construction, that is a primary association of a meaning with a particular (morpho)syntactic configuration, and the possible subsequent change of a construction into a more grammatical one.

In the grammaticalization literature, there is a distinction between primary grammaticalization on the one hand, which corresponds to Noël's formulation of the initial formation of a construction, and secondary grammaticalization, which corresponds to the subsequent change of construction into a more grammatical one. Traugott and Trousdale's constructionalization covers the initial creation, but what about subsequent changes?

This is something that I have been thinking about. I have been using the term "constructional change" in order to capture all the processes that can affect existing constructions, using the following definition:

> Constructional change selectively seizes a conventionalized form-meaning pair of language, altering it in terms of its form, its function, any aspect of its frequency, its distribution in the linguistic community, or any combination of these.

This definition is intentionally very broad, since it is meant to engage with all aspects of constructions. The most important aspect of the definition is actually the very first part, namely that constructional change is selective about what it affects. It selectively seizes a conventionalized form-meaning pair. This means that constructional change is not a system-wide change, or a change that affects multiple constructions at the same time. It is really a very local kind of change, and the types of change that may affect single form-meaning pairs. They are, however, of a more general nature, and you're all very familiar with them.

There are changes in form, such as phonological reduction of *I am going to* to *I am gonna* and further to even more reduced forms.

Changes in form also concern the obligatorification of a particular part of a construction. Here I come back to the English *way*-construction, which has historically come to include an obligatory path or goal constituent. That wasn't

the case all along. Early on, we find examples like "*The legions speed their headlong way*", without a specification of where they are actually going. As time goes on, the relative frequency of examples with a goal or path constituent, they steadily increase in relative frequency until we approximate a hundred percent.

Changes in form also subsume what I talked about in terms of host-class expansion. Here again is the example of the *it*-clefts who expand from noun phrases to prepositional phrases to adverbial phrases and *ing*-clauses, and you find even other constituent types. Changes in form is one type of change that can be subsumed under constructional change.

The same is true of changes in meaning, which I do not need to exemplify in much detail. Let me just come back to the adverb *actually*, which can be used as either an adverbial stating factuality, *He actually handed in his thesis last week*, or as a discourse marker in examples such as *Actually, he handed in his thesis last week*.

Changes in meaning are ubiquitous in grammaticalization, from lexical meaning to more abstract grammatical meanings. Since I have mentioned the term, let me give you an example of secondary grammaticalization. One way in which secondary grammaticalization can manifest itself is semantic change in grammatical elements such as the sentence connector *since*. Originally an expression of a temporal relation, *since* expands semantically to express a the meaning of causality, as in *Since I have a German passport, I do not need a visa for Poland*. That example does not make a statement about a temporal sequence of events. I have had this passport all my life, I am not referring to situation of before and after.

As for changes in frequency, what you find being discussed most often in the literature are changes in text frequency. Some forms fall out of fashion, and speakers use them less and less. Other forms come into a fashion and increase in frequency. That is one aspect, but there are other types of frequencies that are also worthy of investigation. I have mentioned the increase in type frequency of the noun-participle construction in English, which is illustrated by forms such as *work-related*. This construction type has enjoyed a tremendous success in terms of increasing type frequency.

Changes in frequency further subsume changes in the relative frequency of constructional variants, that is, competing or alternative constructions, such as the *s*-genitive and the *of*-genitive in English. Historically, it can be shown that *s*-genitives have changed in their relative frequency profile with regard to possessors that are inanimate. One hallmark of the *s*-genitives is that they typically feature an animate possessor, as in *John's friends*. However, in present-day English, something like *yesterday's events* is a possible way of expressing that

something happened yesterday. That wasn't always possible, nowadays it is. Another example involves the ditransitive construction and the animacy of its recipient role. In the example *Let's give the turkey five more minutes*, the turkey won't receive those five minutes, as it is not an animate recipient. It didn't use to be the case that ditransitive could incorporate recipients that are in fact inanimate.

As for the last type of frequency change, when we investigate constructions and how they change, a crucial aspect is change in the social distribution of a construction. Who is using a particular construction? Does a construction spread out from a small group of speakers to a broader range of age brackets and a wider geographical distribution? The infamous *be like* quotative in English, *I was like, what's that all about*, started as a construction that only a particular subset of speakers would use, but it spread to larger and larger communities. A term of address like *dude* instantiates gendered language use that is typical for male to male speech. It spread to female speakers, who prefer it in same-gender conversations.

Typically, constructional change alters multiple aspects of a construction at the same time. What I have said so far addresses two concerns of Diachronic Construction Grammar, namely how constructions emerge and how they change. Knowledge of language, according to Construction Grammar, is a network of constructions with form-meaning pairings that are mutually connected in various ways. What I have not covered up to this point, but what will be a major focus of what I have to say in later lectures is how we can think about connections between constructions.

In the last couple of minutes for this lecture, let me talk about connectivity change. A lot can change in the network when new links emerge or old links disappear. I would like to discuss two examples. First, we can think about meaning extensions as connectivity change. In meaning extension, such as the meaning of *hopefully* or the meaning of *actually* that I have been talking about, an existing form is linked to a new meaning, which may already exist, albeit linked to a different lexical item.

Second, another type of connectivity change would involve a newly emerging construction which would be linked to a construction that is functionally equivalent. Speakers identify the new construction as an alternative, that is, as a possible competitor, or a possible alternative, to an established construction. I have mentioned the *get*-passive. When this construction came into being, it would have been connected in speaker's minds to the already existing *be*-passive. Speakers noticed it as a new way of expressing the same idea. This is something we could call "synonym linkage". When a new element comes into a language, it is connected to existing elements that express related ideas.

existing links become stronger or weaker

connectivity change

- change in mutual association strength

FIGURE 8

existing links become stronger or weaker

connectivity change

- change in mutual association strength
- I dislike ...
  - ... to do the dishes (*to*-infinitive)

  - ... doing the dishes (*ing*-clause)

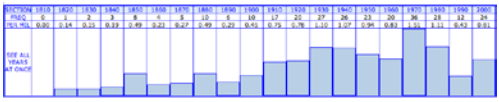


FIGURE 9

Emerging or disappearing links give us interesting material to work with, but that is not the whole story. Much more typically perhaps, what happens is that existing links become stronger or weaker. What you see on this slide here is a display of words that frequently co-occur with the adjective *gay* in the Corpus of Historical American English. The darker the shade of the cell in the table, the more frequent the collocation pattern. You see that some collocations

were highly entrenched in the 19th century and a very different set of collocations are entrenched in the late 20th century. Early collocations were expressions such as *gay colors*, a *gay laugh* and a *gay party*, those are joyful events. Today we talk about *gay men*, *gay rights* and *the gay community*. This reflects the semantic change that the adjective *gay* has undergone. As *gay* changed semantically, its connections changed as well, to the point that there are now two form-meaning pairings.

This not only happens in semantics but also in syntax. This slide shows another data example from the Corpus of Historical American English. This time it is the complement-taking behavior of the verb *dislike*. In Current American English, *dislike* primarily takes *ing*-clauses as complements, as in *I dislike doing the dishes*. This wasn't always the case. As late as the early 20th century, speakers used *dislike* with *to*-infinitives, as in *I dislike to do the dishes*. In Present-day English, this connection between *dislike* and the *to*-infinitive has all but disappeared from the grammar. An existing link has become much weaker, and you can see in the smooth decline of frequencies that this must have happened gradually.

Another type of connectivity change or changing strength in links would be concerned with form-meaning links in polysemous constructions. Polysemous forms are linked with several interrelated meanings, and these links vary in strength. For example, the English verb *to miss* is connected to several meanings. It can have the meaning of "longing", as in *I miss my grandma* and that meaning occurs with a certain frequency. Another meaning is "to fail to reach in time", as in *I missed my train*. There is also the meaning of "not being in possession of" as in *He was missing a front tooth*. The words in the context of *miss* disambiguate those meanings. But the links between *miss* itself and these different meanings vary in strength, and that can be operationalized in terms of how frequent a verb is used in a certain sense.

I am coming to a close here. What I have talked about are mainly the issues of constructionalization, i.e. how new nodes emerge in the constructional network and the opposition of lexical constructionalization and grammatical constructionalization as Traugott and Trousdale conceive of it. I have talked about constructional change, the behavior that existing nodes in the network can exhibit with regard to their form-meaning frequency or distribution in the community of speakers. I have drawn your attention to connectivity changes, the emergence or disappearance of links in the constructional network and what this means for things like meaning extension or other types of semantic development. Lastly, I have talked about changes in connection strength, the associative links between two constructions that can become stronger or weaker as time goes on. With that, I would like to come to a close and thank you for your attention.

LECTURE 3

Three Open Questions in Diachronic Construction Grammar

Good morning, everyone. Welcome to Lecture 3 in this series of Ten Lectures on Diachronic Construction Grammar. In the last two lectures, I have been reviewing the theoretical foundations of Construction Grammar and how a constructional approach can be applied to the study of language change. We've seen that Diachronic Construction Grammar overlaps substantially with research on grammaticalization, but we've also seen that there are a number of distinguishing features that reflect different goals and different assumptions of the two respective approaches.

I have mentioned the fact that Diachronic Construction Grammar is a relatively young research enterprise that is currently gaining in popularity, but that at this point is also not fully matured. There are several issues that are left to be worked out in detail. In this lecture, I want to continue with that theme, so it has the title "Three Open Questions in Diachronic Construction Grammar". What I hope to do is to outline three issues that are currently unresolved and that we need to come to terms with. As I said yesterday in Lecture 2, Diachronic Construction Grammar is the study of changes in the constructional network. I tried to outline four major processes that can inform the study of such changes, namely, constructionalization, which denotes the emergence of new nodes in the constructional network, and constructional change, which captures all the changes that happen to existing nodes in the network with regard to the form of constructions, their meaning, their frequency or their distribution in terms of who in the community of speakers uses these constructions. Then I have discussed connectivity changes, how new links emerge or how links disappear in the constructional network, and I have talked about changes in connection strength. I have mentioned construction-internal links between form and meaning that can become stronger or weaker, for example, in the

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691014

case of polysemous items. I have also discussed associative links between two different constructions, which can become stronger or weaker. I will say more about these types of connectivity changes today. All four of these types of change can guide research into language change, but in this lecture, I want to take the opportunity to think about some of my personal doubts and issues that I think need further consideration.

As a starting point, take the issue that Construction Grammar and historical diachronic linguistics are somewhat strange bedfellows. Construction Grammar, at its basis, is a cognitive approach that focuses on the question of how language is mentally represented. This is the main goal that Construction Grammar inherits from earlier approaches, including Chomskyan generative linguistics. What do speakers know when they know a language? How can we describe as accurately as possible the cognitive system that allows a human being to acquire and use language?

Diachronic linguistics, on the other hand, is the study of how forms and meanings of a language change over time. This can be studied with a cognitive perspective, but it does not have to be. Many linguists study language change without adopting a cognitive perspective. There are a number of important differences in the respective outlooks of Construction Grammar on the one hand, and diachronic linguistics on the other, including the fact that Construction Grammar tries to model linguistic knowledge in the individual, while diachronic linguistics is by necessity about processes going on in the population. The approaches thus target different levels of granularity.

Construction Grammar is a cognitive enterprise that tries to find out how language is represented in the mind, which is not necessarily the case in diachronic linguistics, which may really content itself with statement about language as it has been passed down to us in historical documents. Importantly, Construction Grammar focuses on process that happen on what we could call a human time scale. When I am talking to you, uttering a sentence takes me a few seconds, and you process that sentence word by word, understanding what each word means and what kind of word it is. That happens at the time scale of milliseconds. Conversations can take minutes or sometimes hours. Learning a language can take a few years. How your language changes over a lifetime takes a couple of decades, but that is essentially the range of it. Those are the human time scales that Construction Grammar would be concerned with. By contrast, the processes that are addressed in diachronic linguistics are processes that happen over much longer time intervals that often exceed the lifetime of a speaker, which is an important difference. I am not saying that Construction Grammar and diachronic linguistics are contradicting each other or cannot be unified, but contrasts like these make it actually quite surprising that the combination of the two enjoys such popularity.

In this lecture, I would like to raise three questions that might help us understand how exactly the two approaches fit together and what issues still need to be resolved. First, what is the object of study in Diachronic Construction Grammar? Second, I will come back to the issue of constructionalization and ask a seemingly innocent question: When is a new construction a new construction? The answer is not as trivial as you might think. Thirdly, what knowledge is represented by the nodes in the network and what knowledge resides in the connections between them? I have said yesterday that so far Construction Grammar focuses rather strongly on the nodes. Much attention has been devoted to the constructions, the form-meaning pairings. It remains to be worked out in detail what the relations between constructions contribute to the overall picture.

Let me start with the first question. What is the object of study in Diachronic Construction Grammar? At first glance, the answer seems to be rather trivial. The objects of study are constructions and how they change over time. However, constructions and their development over time can be understood in very different ways. Do we refer to constructions as mental representations of language, as cognitive generalizations that may change over time? Or do we refer to constructions as linguistic forms that we can observe in historical documents? It is my impression that quite often the difference between the two is not explicitly addressed. The primary objects of study, for the most part, seem to be the forms of language that we find in the physical record.

Dirk Noël (2007: 178) has pointed out that "*many functionalists and cognitivists have long been working with a pre-theoretical constructional notion*". That is, the term construction has been used because it is a very convenient label, but it has been unclear what the term actually means. Is it a mental generalization or is it a linguistic form? The term construction works very well as a label for morphosyntactic structures that avoids theoretical preconceptions and assumptions. It can be conveniently used in order to talk about linguistic forms, their meanings, and their historical developments, even when considerations such as speakers' mental representations of language are really not at stake, and you do not want to make any claim about these representations changing or developing in any way. In other words, Diachronic Construction Grammar is a framework that can be adopted as a descriptive tool for the study of language change, and any reference to the psychological reality of form-meaning pairings of constructions can be left implicit.

The idea of a Diachronic Construction Grammar that is completely agnostic towards cognition may strike you as problematic. It could be suggested that Diachronic Construction Grammar should in fact embrace both the mind and the physical record, and that it should set itself the task of figuring out how one relates to the other.

Making inferences about the linguistic knowledge of speakers from many generations ago is difficult, but we can nonetheless point to encouraging examples where this has been attempted. I would like to draw your attention to work done by Peter Petré and his group at the university of Antwerp, which conducts the *Mind-bending Grammars* project that pursues this goal. The project work has created large corpora of individual writers, which allows the precise analysis of changes in the language use of those writers over time.

Other relevant research that has been done includes works by Christoph Wolk and colleagues (2013) who have been investigating the diachronic development of the English dative alternation, the choice between the transitive construction *I wrote my sister a long letter* and the alternative prepositional dative construction *I wrote a long letter to my sister*. Based on Present-day corpus data and on psycholinguistic experimentation, quite a lot is known about how the dative alternation works.

In particular, we know that there are certain determining factors that bias speakers either towards one variant or the other. One factor that influences the choice is verb type. Do we have a verb such as *write*? Or do we have a verb such as *give* or *send*? These verbs have different preferences for either the ditransitive or the prepositional dative construction. Also, the pronominality of the subject influences speakers' choices. Do we have a form such as *I* or do we have a long nominal form like *the secretary of the president*? The givenness of the recipient plays a role as well. In the example *I wrote my sister a long letter*, does *my sister* represent given information? If the recipient is expressed by an indefinite noun phrase, such as *a colleague of mine*, the grammatical form tells us that the recipient is new information that has not yet been introduced to the discourse. Another conditioning factor is the length of the theme that is being transferred. The noun phrase *a long letter* has just three words. *A very long letter with lots of explanations in it* would be a heavy noun phrase with ten words. The longer the theme, the more likely it becomes that speakers choose the ditransitive over the prepositional dative. Finally, the animacy of the recipient is another conditioning factor. Speakers are more likely to use the prepositional dative with recipients that are not animate, for example in the sentence *I brought food to the table*.

All of these factors can be seen to influence speakers' behavior in present-day language use. In experiments under laboratory-controlled conditions, speakers are sensitive to these factors. Corpus studies that are based on synchronic corpus data show that these factors influence speakers' choices. What Wolk and colleagues (2013) wanted to find out is whether the same factors could also be seen at work in historical corpus data. If they were, then that would mean that we can actually investigate the competence of earlier generations of speakers.

Wolk et al. (2013: 383) start with the observation that for Present-day English, corpus data and psycholinguistic data line up rather neatly:

> The likelihood of finding a particular linguistic variant in a particular context in a corpus can be shown to correspond to the intuitions that speakers have about the acceptability of that particular variant, given the same context.

That is the first conceptual step. What we find in corpora reflects rather tightly the cognitive processes that we measure in psycholinguistic experiments. That alignment allows us to use historical corpus data and extrapolate from the historical corpus data to the cognition of the speakers who were alive at that time. Provided that we set the variables of length, givenness, and so on, to the right values, speakers can decide very accurately in a forced-choice task whether a ditransitive or prepositional dative is more appropriate in that situation, and this is evidence that the multivariate profile of a construction in corpus data relates to speakers' knowledge of language. Presumably then, that allows us to extrapolate from synchronic findings to historical corpora to investigate the grammatical knowledge of speakers in the past.

Wolk and colleagues (2013: 384) claim this explicitly: "*Our work ultimately aims to illuminate aspects of the linguistic knowledge that writers in the Late Modern English period must have had, and how this knowledge has evolved over time*." Without a question, the goal of reconstructing the knowledge of speakers of earlier generations is a highly ambitious goal. There are surely limits with regard to the time depth for which we can hope to achieve that goal. For Early Modern English, we still have a sizable amount of data. We have a fairly good idea of the social conditions under which these texts were created. That can give us some confidence in actually pursuing this goal, but the further we go into the past, the more difficult the task becomes.

This brings up another question, namely, should Diachronic Construction Grammar adopt what George Lakoff has called the "cognitive commitment"? The cognitive commitment would be the idea that we want our cognitive linguistic research to be in line with what is known about human cognition in general. We try to model what is going on in speakers' minds. If we want to do that, should we limit our investigations to phenomena that we can plausibly hope to investigate in terms of cognition? This I think is a question that calls for a compromise. We have to recognize that for many phenomena in historical linguistics, any claims about cognition are hazardous. There is too little linguistic data. There is too little that is known about the social context of language use. Still, we might want to talk about phenomena of linguistic

changes that happened in the past in terms of constructions, even when we are not quite sure whether we can come to reliable insights about cognition. Without that kind of research, the Construction Grammar enterprise would lose a lot of important insights, and I discussed some of them yesterday in the context of grammaticalization. What I would suggest is that researchers in Diachronic Construction Grammar should aim to be transparent about their assumptions. Not everybody has to try to reveal what earlier speakers would have known, but I think everybody should be transparent about what it is that they are trying to do.

With that cautious conclusion, let me come to the second open question. When is a new construction a new construction?

In Lecture #2 yesterday, I have reviewed Traugott and Trousdale's (2013: 22) definition of constructionalization, i.e. the creation of new form-meaning pairs in the network of constructions. I also mentioned Traugott and Trousdale's (2013: 22) point that formal changes alone and meaning changes alone do not constitute constructionalization. You'll recall it is not enough for a form-meaning pair to develop a new meaning. It is also not enough for a form-meaning pairing to develop a new formal variant. That is also not constructionalization. Rather, it has to be a new form and a new meaning, so that we have two increasingly independent form-meaning pairings which then instantiate constructionalization. Let us look at this process in some more detail with a concrete example.

As an example, let's take the English verb *confirm*, which is a transitive verb that takes a direct object. The form of the construction is thus *confirm* and a noun phrase that is the direct object. The meaning of that construction is that you verify that something is true. *Confirm* can appear in utterances such as *That letter confirmed my worst fear* or *He confirmed the rumor*.

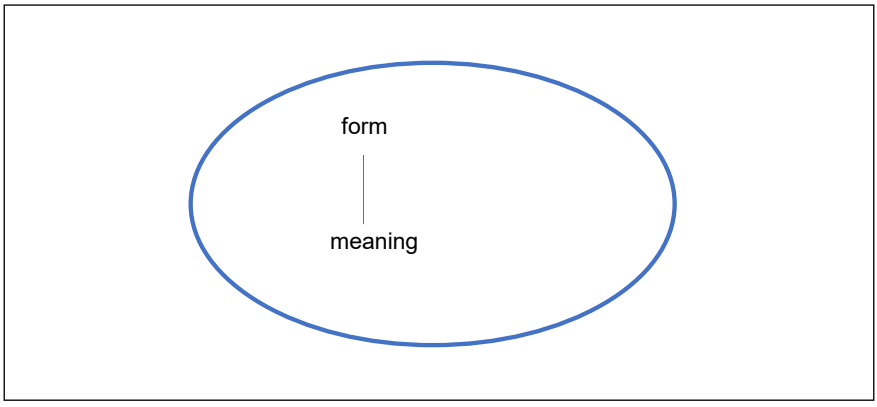


FIGURE 1

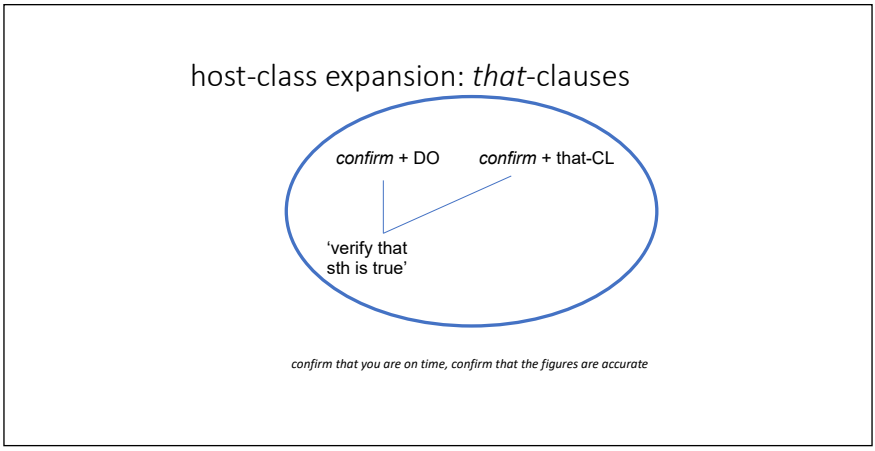


FIGURE 2

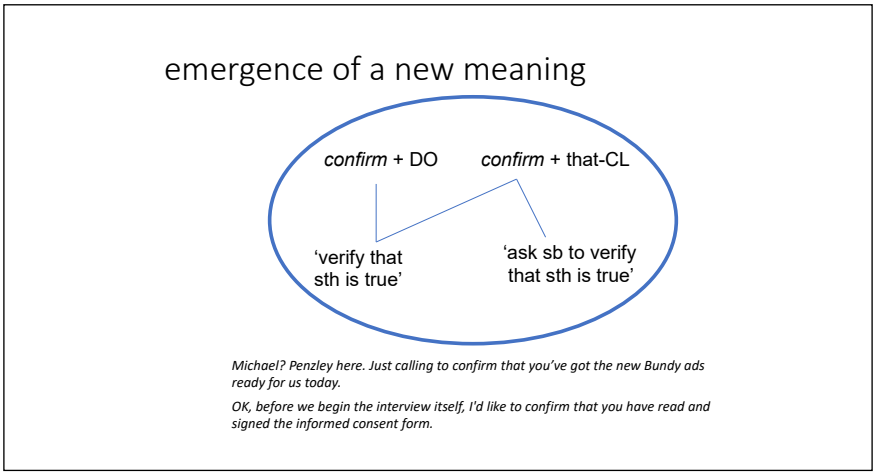


FIGURE 3

Over time, the English verb *confirm* has undergone a change in form, so that it underwent what I called "host-class expansion" yesterday. It began to be used with not only direct objects, but also with the *that*-clauses as complements. You could find utterances such as *You have to confirm that the figures are accurate*. That illustrates host-class expansion. *That*-clauses with *confirm* are already well attested in the nineteenth century, and they have steadily increased in frequency over the past two hundred years. Today, *confirm* and *that*-clauses are tightly associated. You recognize that we have a new form, which means that constructional change has occurred.

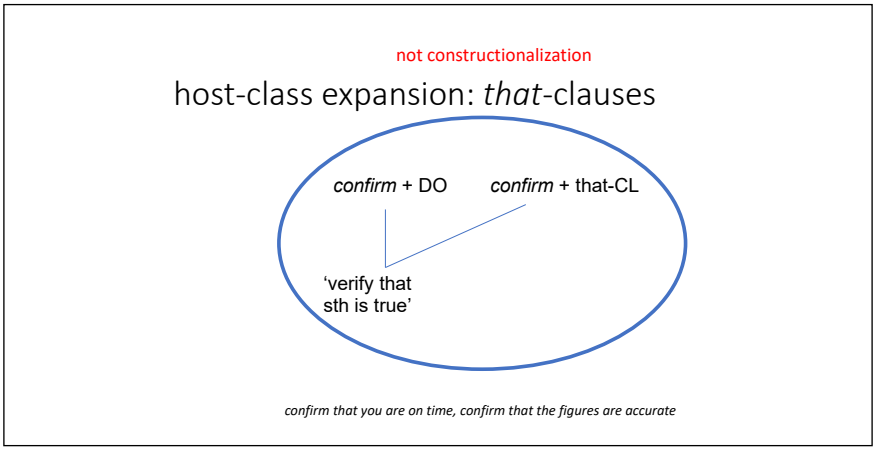


FIGURE 4

More recently, a new meaning of *confirm* has emerged, and this meaning is associated exclusively with the uses that include the *that*-clause. The new meaning is attached to the new form, but not to the old form, and it is metonymically related to the old meaning. It adds what you could call an intersubjective component. Instead of the meaning "verify that something is true", the new meaning is "ask someone to verify that something is true".

Let me discuss the corpus examples on this slide. Someone is calling up another person and asks, *Michael? Penzley here. Just calling to confirm that you've got the new Bundy ads ready for us today*. At that point, the caller does not know whether or not the new ads are ready. They are asking the other person

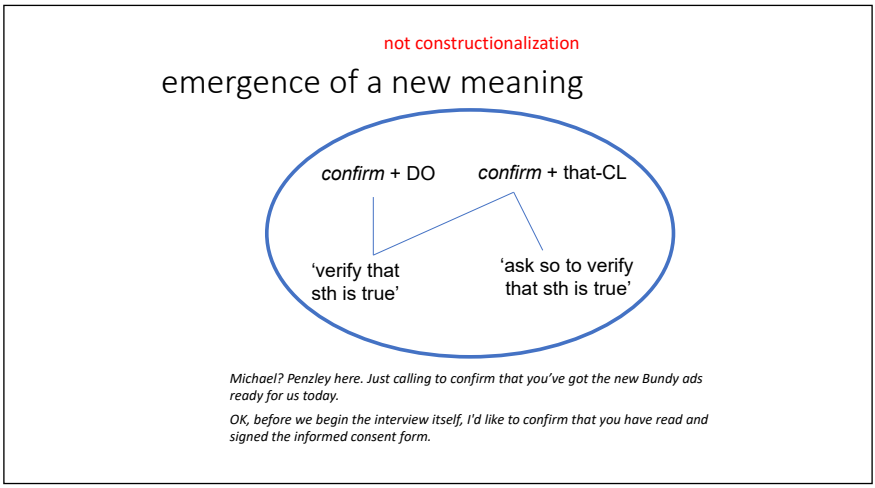


FIGURE 5

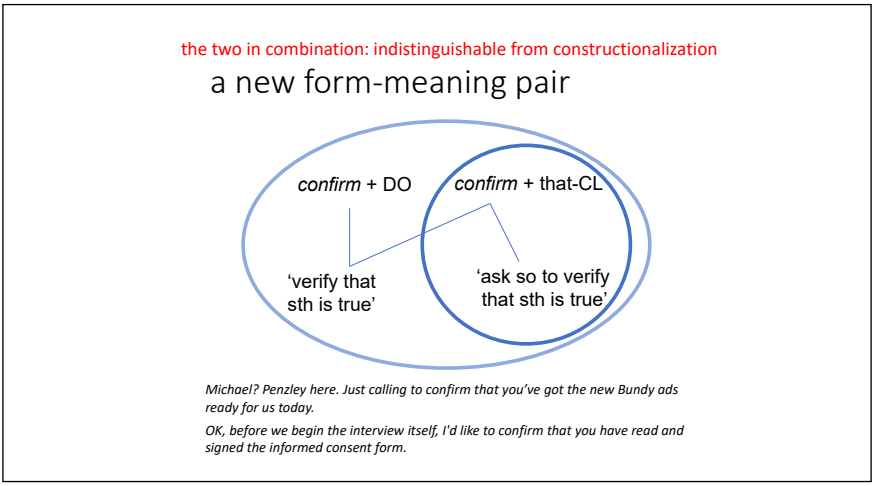


FIGURE 6

for that information. *Confirm* means "I make sure that you tell me" rather than "I tell you that". That is a new meaning. By now there is a new form and a new meaning, which would make this a case of constructionalization.

If we review the steps, the first change is the emergence of a new form through host-class expansion. That by itself is not constructionalization.

Based on that new form, a new meaning emerges, but taken on its own, that is also not constructionalization, but just a constructional change that leads to a new meaning. This leads us to a paradoxical situation in which all the formal requirements for constructionalization are met, while the individual steps do not constitute constructionalization.

To review: two constructional changes in combination yield a new form-meaning pair, and as such, this development is indistinguishable from what Traugott and Trousdale (2013) call constructionalization. My point is the following. Even though the definition of constructionalization seems very clear on paper, in practice when you look at concrete examples in corpus data, and you see how a construction develops, you may sometimes be looking at a sequence of constructional changes which in the end conspire, so as to look like constructionalization. This point has been made in similar form by Börjars et al. (2015) in a review of Traugott and Trousdale (2013). They have argued that what counts as constructionalization crucially depends on the previous steps of constructional change that are taken into consideration. In other words, constructionalization is a relative notion. It is not something that is objectively there in language change. Rather, it depends on the perspective of the observer. Let me try to explain this.

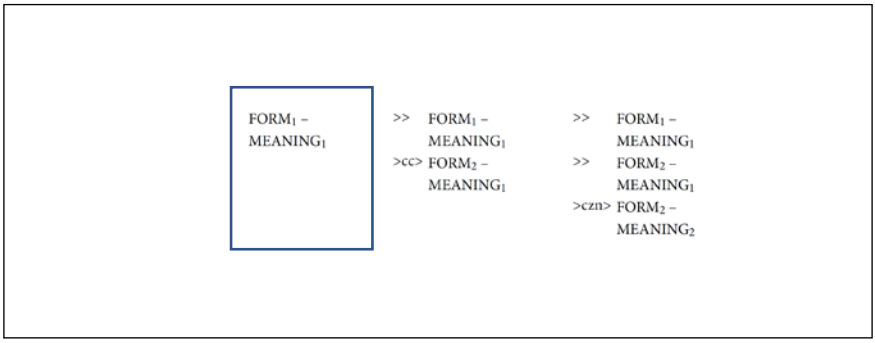


FIGURE 7

The starting point of the development is a form-meaning pair that can be expressed as $FORM_1$, $MEANING_1$. If that pair develops a new form, that constitutes constructional change, as we have seen with *confirm* and its host-class expansion. As a result of that process, the original $FORM_1$-$MEANING_1$ pair is now accompanied by a second one, in which $FORM_2$ is associated with $MEANING_1$. Everyone would agree that this is constructional change, but then the development continues, resulting in a second meaning that is attached to $MEANING_2$. That would be *confirm* with *that*-clauses with the new meaning. Applying Traugott and Trousdale's definition, we could say that the pair of $FORM_2$ and $MEANING_2$ qualifies as a case of constructionalization. A first form-meaning pair, $FORM_1$, $MEANING_1$, gives rise to a second one, namely $FORM_2$, $MEANING_2$. There is however a complication.

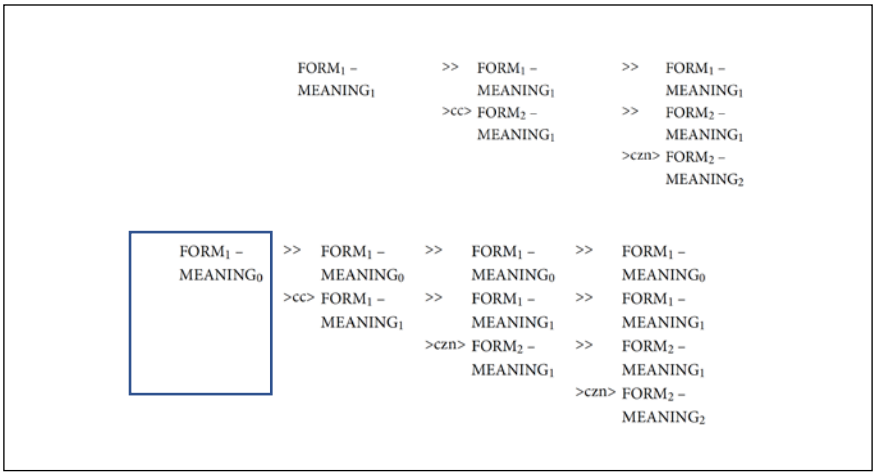


FIGURE 8

Let us suppose that we work with the very same data, but we trace the construction's history further back. Let us further suppose that at some earlier stage, the construction had the same form, but it had a different meaning. That is what is called MEANING_0 on this slide. In that case, the original starting point, FORM_1-MEANING_1, is reached through one meaning change, which represents one step of constructional change. Once the new form develops from that, we have FORM_2-MEANING_1, counting back from FORM_1-MEANING_0. There is now a new form and a new meaning, which means that already at this stage, we would be justified to call the development a case of constructionalization. The phenomenon is exactly the same, it is just that the analysis has gone further back in time. We have assumed a different perspective. This means that the term constructionalization is relative to the developmental starting point that is chosen by the analyst. What is counted as FORM_1 and MEANING_1 influences the result of what we are entitled to see as a new construction.

When is a new construction a new construction? I hope to have shown that it is not so trivial to decide. I do not mean to suggest that constructionalization is not a term that you should use. I find that constructionalization is a useful label for the emergence of new form-meaning pairs in the constructional network, but I think we should be aware that in practice, distinguishing between constructionalization and conspiring constructional changes may be impossible, or at least it may depend on our point of view. More constructively perhaps, I would like to say that there are other distinctions than "constructionalization vs. constructional change" that may ultimately be more useful for the study of change in the constructional network.

For that, I would like to propose a matrix of possible changes. Here you see a table with two cross-cutting dimensions. In the columns, we have changes that affect form, that affect meaning and that affect connections between constructions. In the rows, we have the phenomenon of emergence, strengthening, weakening, and disappearance. Every cell in this table illustrates a type of change that can happen in the constructional network.

Starting with the emergence of forms, this would be Traugott and Trousdale's constructionalization, i.e. new forms with new meanings appear, for example lexical elements like *selfie* or *Brexit* or grammatical constructions such as the *get*-passive.

The emergence of new meanings is not necessarily the emergence of new constructions, although the two may go hand in hand. During my lifetime, I was lucky enough that a new concept emerged, namely *wireless internet access*. That is a new concept. It is being expressed and lexicalized in different ways, but it is a new idea that came into the world.

	form	meaning	connection
emergence	new forms appear: *selfie* *Brexit*	new concepts appear: 'wireless internet access'	new connections are formed: *gay* forms a connection with 'homosexual'.
strengthening	forms become more frequent: *like* as a discourse marker	the concept 'gluten intolerance' increases in popularity and use	connections gain in strength: *fantastic* with the meaning 'wonderful'.
weakening	forms become less frequent: *whom* as a relative pronoun	concepts become less frequent: 'sale of indulgences'	connections fade: The verb *dislike* used less with *to*-infinitives.
disappearance	forms disappear: the word *affuage* is no longer used	concepts disappear: 'the right to gather firewood in a forest'	the English ditransitive is no longer used with *forbid*

FIGURE 9

New connections can be formed between forms and meanings. Yesterday I have showed you the collocates of *gay* which reflect the new meaning "homosexual" that has become associated with that word at some point.

What about the strengthening of forms? This would reflect the case where forms become more frequent. The discourse marker *like* would be one example of a form being used more extensively.

The same also happens with meanings. Certain concepts can become more popular over time. This is illustrated by the concept of "being intolerant against gluten". *Gluten intolerance* is a concept that has increased in popularity and use.

Connections can be strengthened. The word *fantastic* in English, referred to "things that are unbelievable or mythical". Nowadays, there is a much stronger association of *fantastic* with the meaning of "wonderful". Originally, in order for something to be fantastic, it had to be a unicorn or a sorcerer. Today, toilet paper can be fantastic if it is really good toilet paper.

Moving on to the weakening row, forms can become less frequent. The English relative pronoun *whom* is falling out of use, becoming less frequent.

Concepts can become less frequent as well. In the Medieval Christian church, there was a concept of *sale of indulgences*. If you had committed a sin, you would simply pay a sum of money as a form of redemption. We can still understand the concept, but it is no longer very common.

Connections can fade. Yesterday I presented frequency trends of the verb *dislike* with two different complementation patterns. *Dislike* is used less and

less with *to*-infinitives, which means that the connection between the verb *dislike* and the *to*-infinitive construction is becoming weaker and weaker.

Forms, meanings, and connections can disappear. The form that you see here, *affuage*, is obsolete. Let me show you the meaning that once was associated with it.

It is "the right to cut firewood in a forest", a privilege that you have as a peasant. We are no longer in the habit of collecting firewood. It is therefore not too surprising that this concept has disappeared from usage.

The last cell in the table concerns the disappearance of connections. I could have listed the connection of *dislike* with the *to*-infinitive here, which is nonexistent for many speakers of English. There is however another example that I want to present, namely the connection between the English ditransitive construction and verbs such as *forbid*. It is no longer possible to use *forbid* in the ditransitive construction.

My basic point here is that constructionalization, which concerns the emergence of form and meaning, represents a small subset of all the changes that go on in the constructional network. There is a world of other changes left to be explored. Regularities in those changes remain to be discovered, which I think is a worthwhile project.

With that, I would like to come to my third question. What knowledge is represented by the nodes, and what knowledge is represented by the connections between them? Let me draw your attention to what I would like to call the "fat node problem". I said yesterday that Diachronic Construction Grammar is a very young enterprise that maintains a strong focus on idiosyncratic form-meaning parings. It is fair to say that up to now, the focus has really been on the nodes in the network, rather than the connections, and I consider that to be a problem. Current models of the constructional network store nearly all the information in the nodes, while only very little information resides in the connections. Let us examine what the purpose of these connections is. Two types of connection have been recognized as central. First, there are symbolic links that connect form and meaning. Second, the network is structured by inheritance links, which are categorizing relationships that obtain between concrete constructions and the more abstract patterns they instantiate. Let me show you this.

This slide shows a small snippet of a constructional network (Croft and Cruse 2004: 264), and the network shows how idioms such as *kick the bucket* and *kick the habit* inherit information from a more general construction, namely transitive *kick*, which is shown here as "subject *kick* object". Transitive *kick* inherits aspect of the transitive construction, "subject transitive verb object", and the transitive construction in turn inherits aspect from the clausal construction,

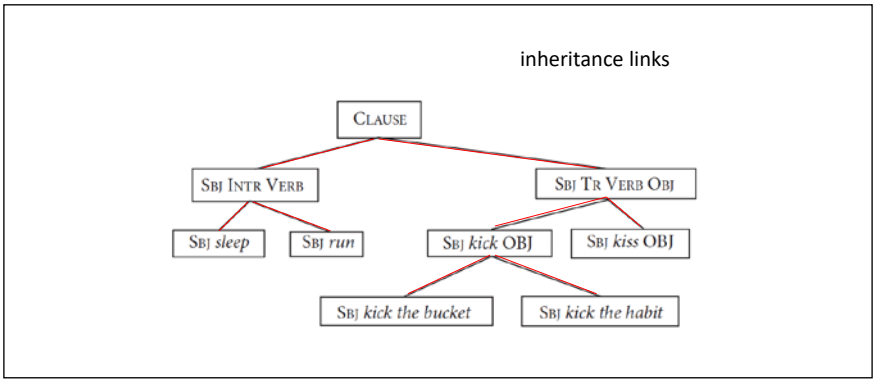


FIGURE 10

which includes not only transitive but also intransitive verbs, ditransitive verbs, complement-taking verbs, and so on and so forth. The inheritance links are what I have marked up in red, categorizing relationships between constructions of varying degrees of schematicity.

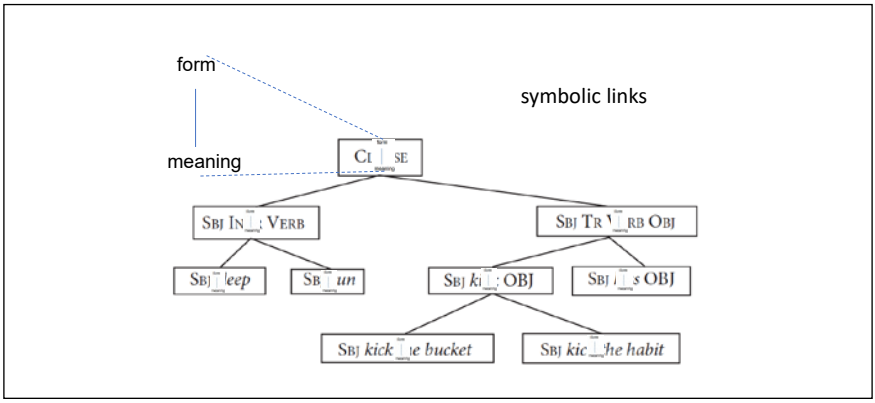


FIGURE 11

Besides inheritance links, there are also symbolic links within each construction in the network. I have tried to visualize this on this slide. Within each construction in the network, each node has within itself a form and a meaning, and the two are connected through a symbolic link. There is a symbolic link in the clause construction. There is a symbolic link between form and meaning in the transitive construction. There is one in transitive *kick*. There is one in *kick the bucket* and *kick the habit*, and so on and so forth.

The fat node problem is the following. Each construction in this network has information inscribed at the form side and at the meaning side. These are

these properties of constructions that we as construction grammarians try to figure out and ascribe to the respective constructions. Sometimes you see this information formalized in attribute value matrices, so-called AVMs. In work by Kay and Fillmore (1999) for instance, you will find AVMs with information on case, agreement, and whether a structure is a maximal projection or not. All of the complexities of non-compositional meanings of a construction, all of the formal constraints and the idiosyncrasies, all of that is typically presented as information that is stored within the nodes. Yesterday I discussed redundant representations. If information is stored redundantly in the network, that means the lower nodes here, *kick the bucket* and *kick the habit*, are just as fat, if not even fatter than the upper ones, because all the information that is stored up at the clause level construction is represented again one more time as *kick the bucket* or *kick the habit*. As the nodes get more concrete, they represent more and more specific meaning in addition to the general meanings, and so they get fatter and fatter.

Now you might ask why there should be anything wrong with that. Isn't that how we are supposed to be talking and thinking about constructions? Yes and no. Not everybody thinks about constructions in this way, and not everybody thinks that it is a good way to think about constructions.

Here's a quote by Dick (Richard) Hudson (2015: 692), whose own theoretical framework is word grammar:

> I believe that language is, indeed, a network, and that this network is, indeed, a structure. Many other readers may protest that they too see language as a network; after all, cognitive linguists envisage 'an elaborate network comprising any number of conventional units linked by categorizing relationships' Langacker (2000: 12) or 'a network of constructions which captures our grammatical knowledge of language in toto, i.e. it is constructions all the way down' (Goldberg 2006: 18). But notice that in these cases the complex units which the network connects have their own internal structure which is not part of the network. […] [N]etwork theory goes further by claiming that 'it is networks all the way down'.

Instead of constructions all the way down, it is networks all the way down.

What Dick Hudson has in mind is something like a neural network in which the nodes are a lot less complex. In a neural network, the nodes receive activation, and if that activation passes a certain threshold, they pass on activation, they fire, but they do not store information on case or agreement or any syntactic idiosyncrasies. These are slim nodes. They can only do one thing. You might say they are less sophisticated in comparison to the Construction

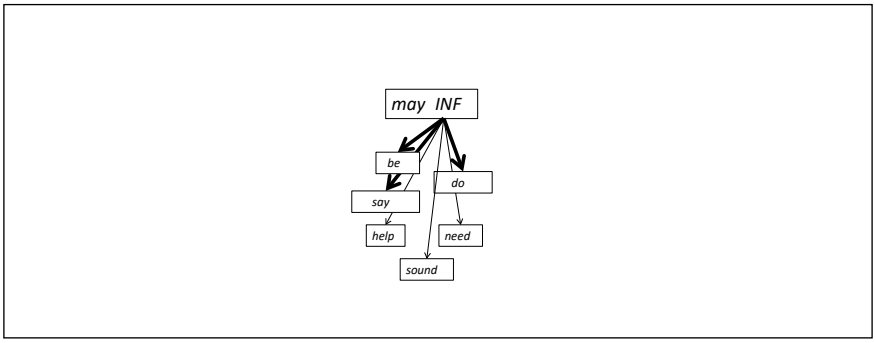


FIGURE 12

Grammar nodes, but undoubtedly, this is a lot more elegant. The information on syntactic constraints in a model like this is not stored in the nodes. It is stored in the way the nodes are connected. A lot more information is stored in the actual network arrangement rather than in the nodes themselves.

According to Hudson, inscribing information into the nodes is a kind of cheating. We present our model of language as a network when instead we take great liberties to describe all kinds of information in the way we want directly at the level of the construction. How can we make a Diachronic Construction Grammar more 'networky' and less 'nody'? That would be what I want to discuss in the rest of my time this morning.

To give you a quick example of how we could think of constructions in a way that comes closer to the idea of a network, let me say a few words about the English modal auxiliaries and how they have changed in recent times. I take a view of language in which modal auxiliaries, such as English *may*, which you see here, form an associative network with other linguistic units, and in particular, the lexical verbs that they take as infinitive complements. I think of the modal auxiliary *may* as a construction that has an open slot for a verb in the infinitive, and then there are links, associative links to lexical verbs that can fill that infinitive slot. Here we have the thick arrow to the verb *be*, meaning that *may* frequently co-occurs with *be*.

An assumption that I make is that when a speaker knows how to use a modal auxiliary, they do not only know its general morphosyntactic behavior. There are certain syntactic characteristics that are typical of the modals. They can take an infinitive complement, modals can invert with the subject, and they can take a negative contraction in some cases. But in addition, speakers know that auxiliaries tend to co-occur with some elements more frequently than with others. Some of the associative links here are stronger than others, and this I think is part and parcel of linguistic knowledge.

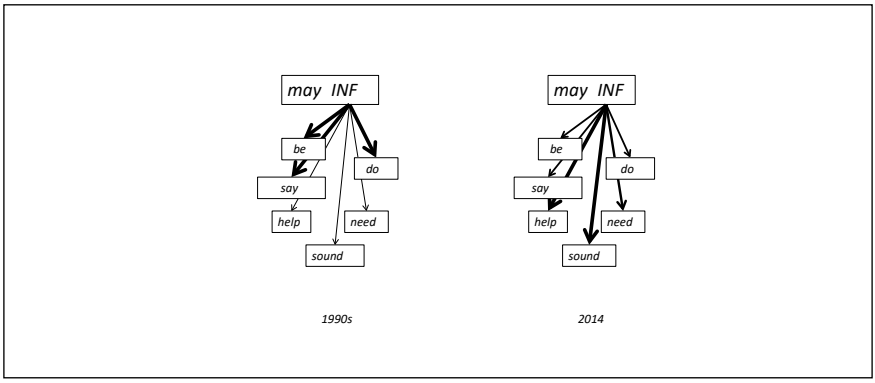
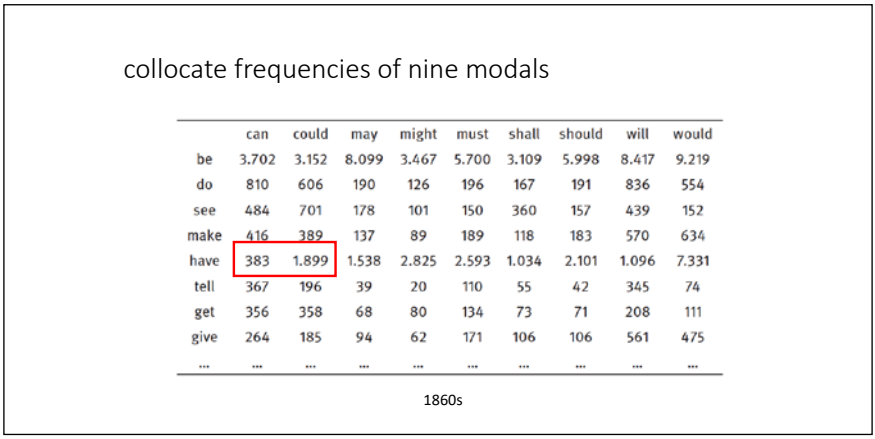


FIGURE 13

Over time, these patterns of association can change, so that a modal auxiliary such as *may* is subject to changes, to the effect that some connections grows stronger and others weaken, and perhaps even disappear. This goes back to the matrix of changes I presented earlier. Connectivity changes may yield the result that some connections strengthen over time, while other connections weaken. This is what you see represented in these two graphics here.

Corpus data allows us to study these changing patterns, and I would like to show you some results along these lines, for which I have used the Corpus of Historical American English, which is a corpus of different written genres, spanning some two hundred years of text, which are divided into decades.

Let us look at a data example about connectivity change in the English modals. From the COHA, I extracted data for nine different English auxiliaries, each of them followed by a verb in the infinitive. The data that I extracted from the corpus are the frequencies of those infinitive verbs.



	can	could	may	might	must	shall	should	will	would
be	3.702	3.152	8.099	3.467	5.700	3.109	5.998	8.417	9.219
do	810	606	190	126	196	167	191	836	554
see	484	701	178	101	150	360	157	439	152
make	416	389	137	89	189	118	183	570	634
have	383	1.899	1.538	2.825	2.593	1.034	2.101	1.096	7.331
tell	367	196	39	20	110	55	42	345	74
get	356	358	68	80	134	73	71	208	111
give	264	185	94	62	171	106	106	561	475
...	...	...	...	...	...	...	...	...	...

FIGURE 14

The data that I retrieved was organized in tables that look like this. Here we have data for the 1860s. In the columns we have the nine modals, *can*, *could*, *may*, *might*, and so on and so forth. In the rows, there are the frequencies of the lexical verbs that collocate with these modals. You see that highly frequent verbs such as *be*, *do* and *have* are frequent with all of the nine modals, but some interesting asymmetries are already apparent in the raw frequencies. For instance with *can* and *could*, *could have* is a very frequent, idiomatic collocation of the auxiliary and a lexical verb. *Can* has just about the same overall frequency as *could*, but *can have* is a lot less frequent. Asymmetries of this kind can be perceived just by looking at the raw co-occurrent frequencies that ultimately inform our understanding of how these modals differ in terms of their collocational profiles.

This kind of data allows us to make systematic comparisons between each pair of modals. The reasoning is that if two modals occur with similar sets of collocates at similar frequencies, they stand in a close semantic relation. I have talked about paradigms and paradigmatization yesterday. A perfect paradigm would be one in which each member of the paradigm has just about the same co-occurrence frequencies. We can use this data for the purpose of contrasting the nine modals in terms of their collocates and see which ones pattern in similar ways. If, for instance, we look at *could* and *may*, and we compare the frequency differences between all the different collocates and add them up, we arrive at a measure of how similar or how different they are. If we compare that measure against, for example, the differences between *should* and *must*, we can actually see which of the two pairs is semantically more strongly related. All of this relates to the idea that you can use collocates as a similarity measure. John Rupert Firth (1957) has coined the slogan, "*You shall know a*

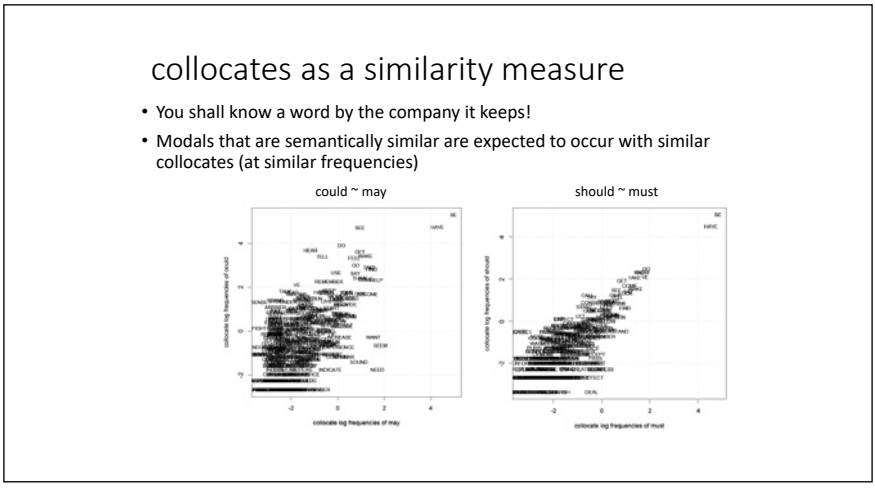


FIGURE 15

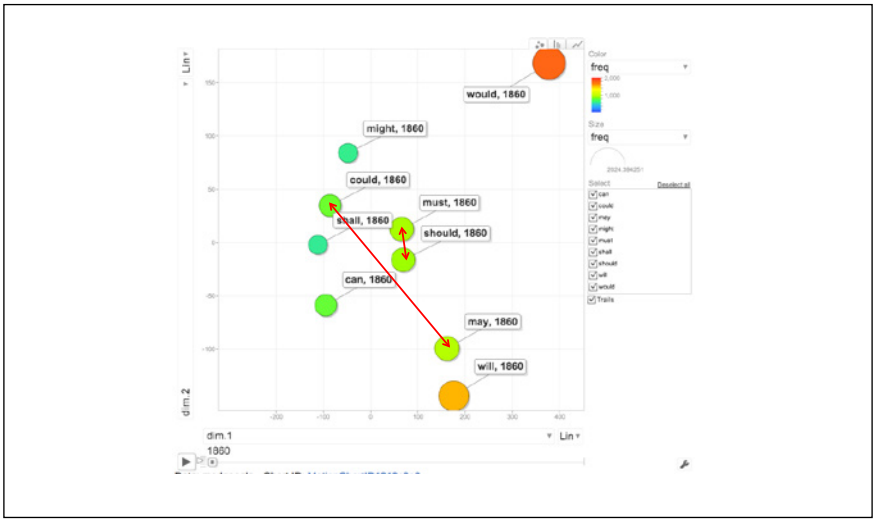


FIGURE 16

word by the company it keeps", and this is an instantiation of this idea. Modals that are semantically similar are expected to occur with similar collocates at similar frequencies.

What you can see when you contrast the modals in this way is that some pairs are quite different. For instance, here is a graphic that shows you the collocate frequencies of *could* and *may* with regard to the infinitive verbs. You see that the frequencies do not form a straight diagonal. In other pairs that are semantically more closely related, we see stronger correlations. *Should* and *must* both encode obligations, and here the diagonal is much cleaner.

If we conduct pairwise comparisons for all nine modals, we can quantitatively assess the mutual similarities and use analysis techniques such as multidimensional scaling to represent those differences in a graph. Here you see the nine modals, *would*, *might*, *could*, *must*, *shall*, *should*, *can*, *may*, and *will*, arranged in a two-dimensional map that reflects their collocational behavior. Each modal construction is a bubble, and bubbles that are close to one another occur with similar sets of lexical verbs at similar frequencies. For example, *must* and *should* are very close together, in a similar position in the graph. That means they occur with similar verbs at similar frequencies. They have a collocational profile that is largely identical. By contrast, *could* and *may* are there very far apart, and that would mean that they encode different meanings, different ideas. It also means that they occur with very different sets of verbs. You could say they are not in as tight a paradigmatic relationship as *must* and *should*. Distance represents differences in terms of collocational behavior.

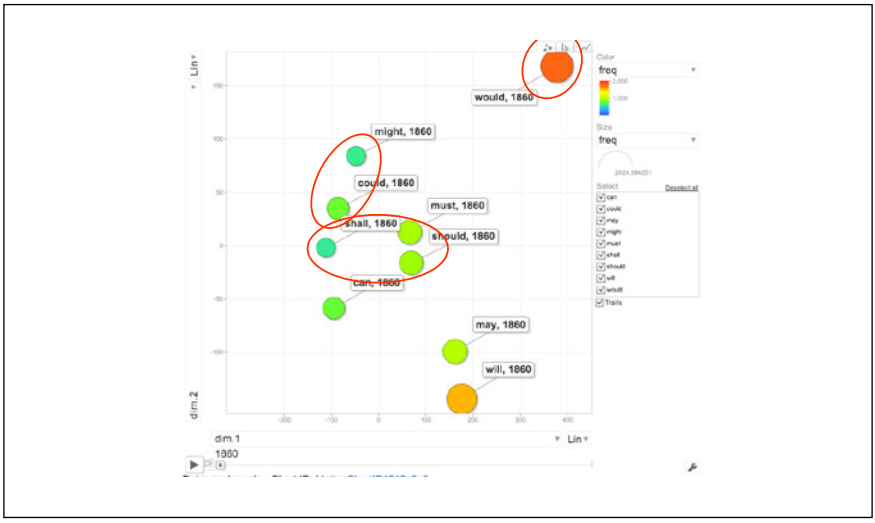


FIGURE 17

There is another meaningful element of this graph, namely, the size of the bubbles. Size represents normalized corpus frequency. If a modal auxiliary is represented by a large bubble, that means that it occurs very frequently in the corpus. If it is represented by a smaller bubble, that means that it is not as frequent in the corpus as the others. For example, the modal auxiliary *shall* is relatively infrequent.

There are a couple of things that we can see in this graph that I would like to point out. We can see, for instance, that *must*, *should* and *shall* pattern relatively closely together. These are modals that encode obligations. We see that *might* and *could* also pattern together; *might* and *could* encode logical possibilities. We further see that *would* has a profile that is very different from all the rest of the modals. What I would now like to show you is how these modals developed over the past one hundred fifty years.

I would like us to focus on the modal *may* and how it develops over time. In the overall development that takes place, it is clear that the modals are on the move. The English modals have been developing over the past a hundred and fifty years in terms of their semantics and in terms of their collocational behavior. This is corroborated in this analysis here. In the beginning, 1860s, *may* is positioned towards the bottom of the graph. As time goes on, its overall trajectory is upwards and towards a cluster of modal auxiliaries that comprises *should*, *must* and *might*. In a way, *may* integrates itself and forms a tighter paradigm with these elements. You may wonder, what has happened to *may*? Why has it made this journey? What has happened to *may* has elsewhere in the

literature been described as a development towards more and more epistemic meaning (Millar 2009).

I have talked about different meanings of modal auxiliaries yesterday. *May* can express permission, as in *You may kiss the bride now*, or it can express logical possibilities, for example when I say *That may be a good idea*. To find out whether *may* increasingly adopts epistemic uses, I decided to study this a little more closely.

I conducted what is called a distinctive collexeme analysis (Gries and Stefanowitsch 2004). The logic of comparing frequencies in one construction to another construction allows us to figure out which lexical elements are particularly typical for one construction, as opposed to the other. You see a toy

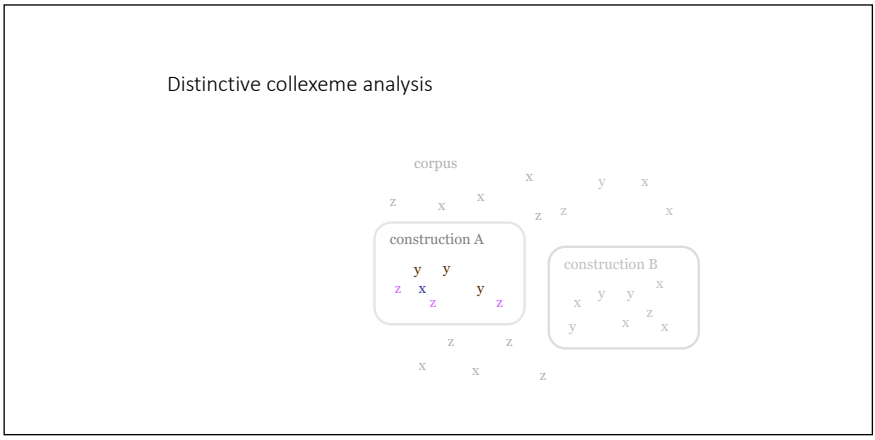


FIGURE 18

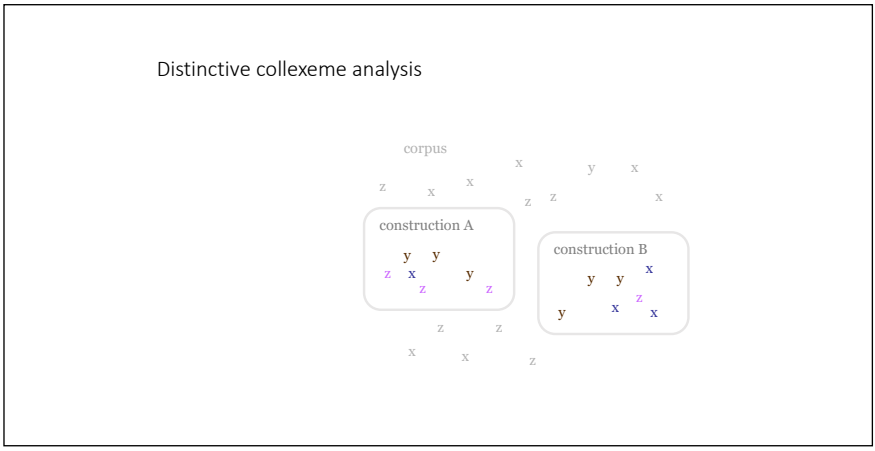


FIGURE 19

example of that on this slide here. We have a corpus with lots of words in it. The words are represented by *x*'s and *y*'s and *z*'s. Construction A has a particular collocational profile. It occurs three times with *y*, three times with *z*, but only once with *x*. If you were look at that by itself, you would conclude that for this construction, *y* and *z* are equally important because we have three of each of them.

However, when we draw a comparison between construction A and construction B, construction B occurs three times with *y*, three times with *x*, and once with *z*. In this comparison, it turns out that *y* is not really typical for just one of them, but rather it is common to both. What distinguishes A and B is that construction A occurs a lot more with *z*, and construction B occurs a lot more with *x*. That is the general logic of distinctive collexeme analysis. You highlight the differences in collocational profiles and single out those elements have the most asymmetrical distribution across the two.

How did I apply this method in this case? I was actually not investigating two different constructions, but I was comparing the same construction across earlier corpus data and later corpus data, trying to find out how *may* would have changed diachronically with regard to its collocates.

Let's see how many collocates of *y may* has in period A and how many collocates of *y* it has in period B. It turns out that they are both the same. That collocate is not an element that distinguishes between the two periods. That collocate is something that stays constant. By contrast, the collocate *x* shows a frequency increase over time. The goal of a diachronic distinctive collexeme analysis is to find out which collocates are maximally uneven in their distribution across different historical periods.

Let me show you some of the results of my analysis. I contrasted the verbs that occur with *may* in the 1860s against the verbs that occur with it in the 2000s, and the verbs that you see in this table are the verbs whose distribution is maximally asymmetric across the two decades. For instance, *may say* occurs very frequently in the early periods, about 300 times in the 1860s data. It is vastly overrepresented. The expected frequency is much lower than what we actually observed. Conversely, if we look at the data from the 2000s, the observed frequencies of *may help* are more than double what the expected frequencies are. *May help* is very typical for the late data. *May say* is very typical for the early data. This allows us to say something meaningful about how *may* developed as an auxiliary. Let's look at some actual examples with these verbs.

If we compare examples with distinctive verbs from the 1860s to examples with distinctive verbs in the 2000s, it turns out that examples with *say*, *do*, *add* and *judge*, which are typical for the 1860s, tend to express permissive meaning, as in "*If I may say so*", if I have the permission to say so, or "*You may do that if*

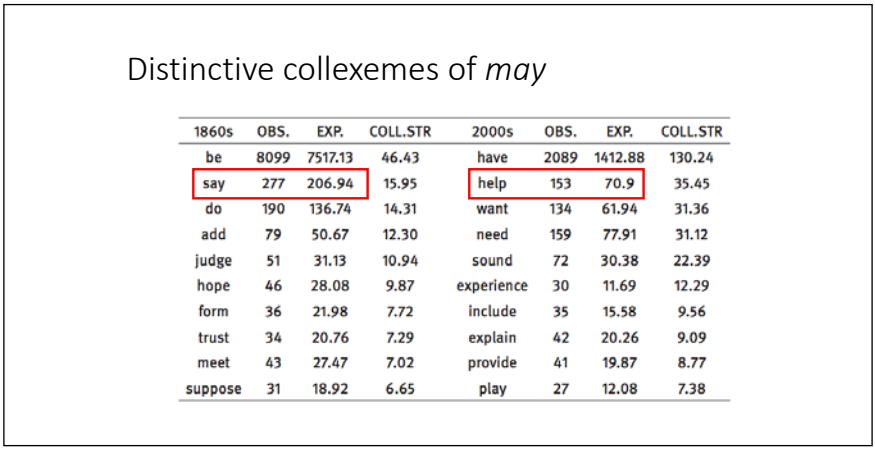


FIGURE 20

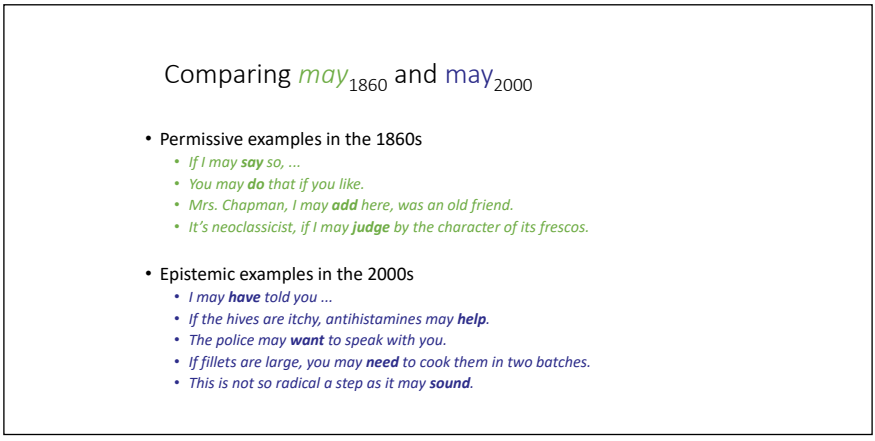


FIGURE 21

you like", you have the permission to do it. If we contrast that with the preferred or distinctive verbs for the 2000s, i.e. *have*, *help*, *want*, *need* and *sound*, these examples encode possibilities rather than permission. The utterance "*I may have told you*", is not about me being allowed to do something, it encodes that I probably told you. "*If the hives are itchy, antihistamines may help*", encodes the meaning that there is a drug that is a possible cure for your allergies. "*The police may want to speak with you*" encodes that it is possible that they would actually like to talk to you. In the light of this additional evidence, I feel comfortable interpreting this movement of *may* as a move into epistemic territory.

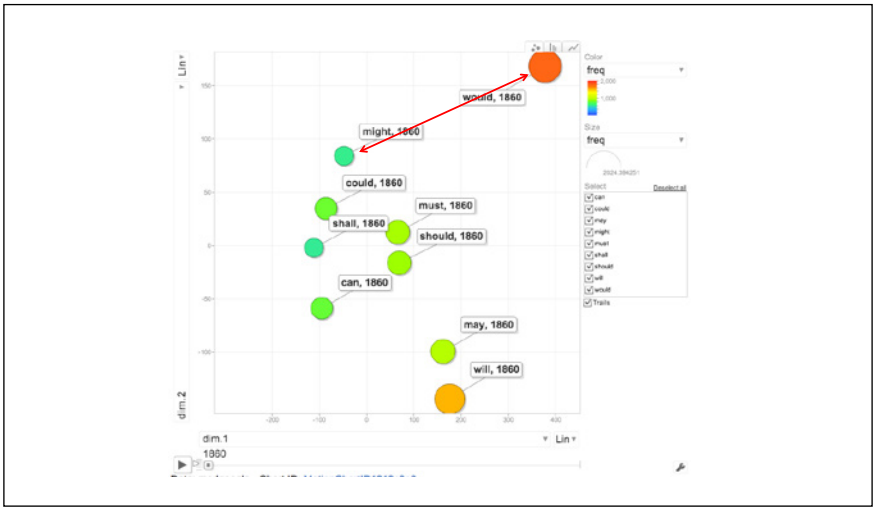


FIGURE 22

I would like to come back to this chart to discuss another contrast between two modal auxiliaries. I mentioned before that *would* patterns quite unlike the rest of the modals, and a good element to compare it to would be the modal *might*, which is situated more towards the left side of the graph, relatively far apart in terms of where it is on the x axis.

The two are sometimes used synonymously. You can say "*That might be a good idea*", and "*That would be a good idea*", and these two sentences convey roughly comparable ideas. That made me want to compare those two. I wanted to find out what the differences between *might* and *would* are and what this axis reflects.

Distinctive collexemes of *might* and *would* (1860s)

MIGHT	OBS.	EXP.	COLL.STR	WOULD	OBS.	EXP.	COLL.STR
judge	13	3.36	7.65	like	457	352.34	37.89
happen	52	26.34	7.32	make	634	536.3	18.84
have	2825	2622.63	7.10	give	475	398.33	15.87
befall	12	3.1	7.06	seem	437	382.01	8.54
expect	27	10.85	6.75	permit	76	57.86	7.53
see	101	65.33	6.21	allow	77	59.34	6.80
get	80	49.32	6.02	do	554	504.4	5.43
be	3467	3275.96	5.73	pay	74	58.6	5.11
live	39	20.14	5.39	leave	114	94.2	5.00
prove	66	40.03	5.38	care	34	25.22	4.41

FIGURE 23

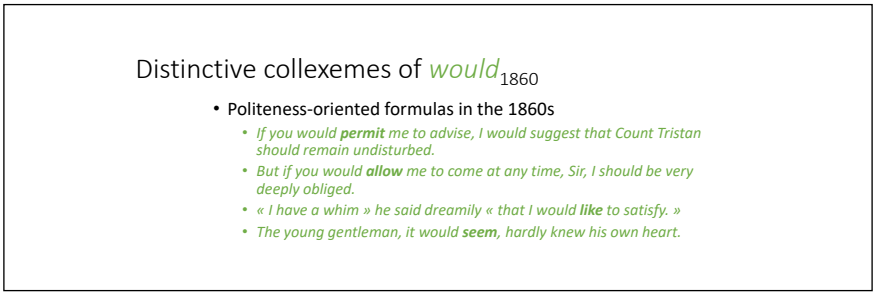


FIGURE 24

I ran another distinctive collexeme analysis, this time of *might* and *would* in the 1860s. If you take a look at the verbs that characterize *would*, like *make*, *give*, *seem*, *permit* and so on and so forth, you can identify a good number of verbs that form part of politeness-oriented phrases like *I would like*, or *it would seem*, or *if you would permit* and *if you would allow me*.

Actual examples with *would* in the 1860s confirm this idea. Here are some politeness-oriented formulas from the 1860s, "*If you would permit me to advise*", or "*If you would allow me to come at any time*", or "*I have a whim that I would like to satisfy*". These kinds of phrases are what make *would* stand out from the rest of the modal auxiliaries.

Now, in the rest of the time that I have for this lecture, I would like to present another study of shifting collocational preferences, this time focusing only on a single modal. How has the associative network of *may* shifted over the past two hundred years? The collocational analysis that I have shown you up to this point reveals only the peak of the proverbial iceberg. If we are only looking at the verbs that are maximally uneven in their distribution, we only see those elements that change in the strongest way, but there are other verbs that change as well. We might be interested in seeing the broader picture, since there is much more going on than what we see in the top ten verbs of a distinctive collexeme analysis.

For this analysis, I used a different method. I constructed what is called a semantic vector space of the 250 most frequent verbal collocates of *may*. In principle, this analysis is not that different from what I presented for the modal auxiliaries. In the study I presented earlier, the items for comparison were the nine different modal auxiliaries, which were compared on the basis of the co-occurrence frequencies of all the verbs that are found with those modal auxiliaries in a corpus. This analysis draws on the same kind of data, but the items that we want to compare are 250 different lexical verbs that occur with *may*. The analysis is based on the collocates of those verbs, and the frequencies of

these collocates are used for the purpose of comparing these verbs in terms of their semantics.

For each verb that enters into the analysis, I collected data, so that each verb can be characterized by the words that occur around it in corpus data. I have used a context window of four words to the left and four words to the right. I deleted common words such as *the* or pronouns, so that was only left with lexical elements that are highly contentful, and I did not take raw frequencies, but rather I weighted those frequencies with a collocation measure, namely Pointwise Mutual Information. I will be happy to talk more about this method in detail, but here I just want to show you the results.

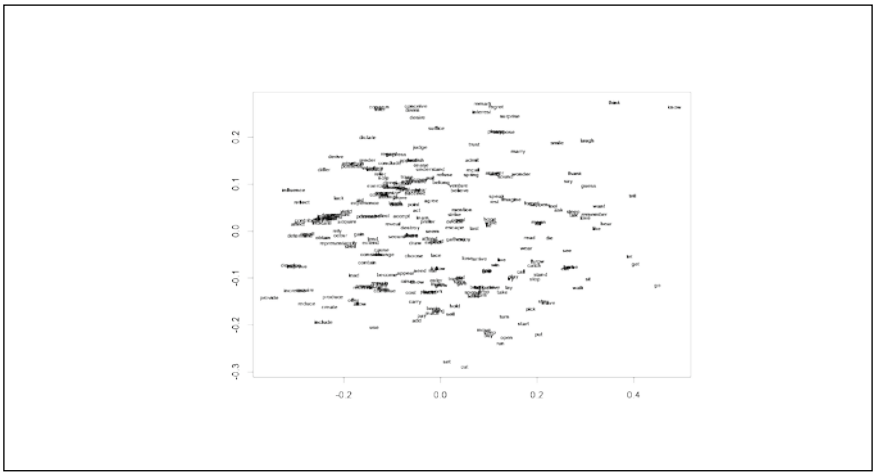


FIGURE 25

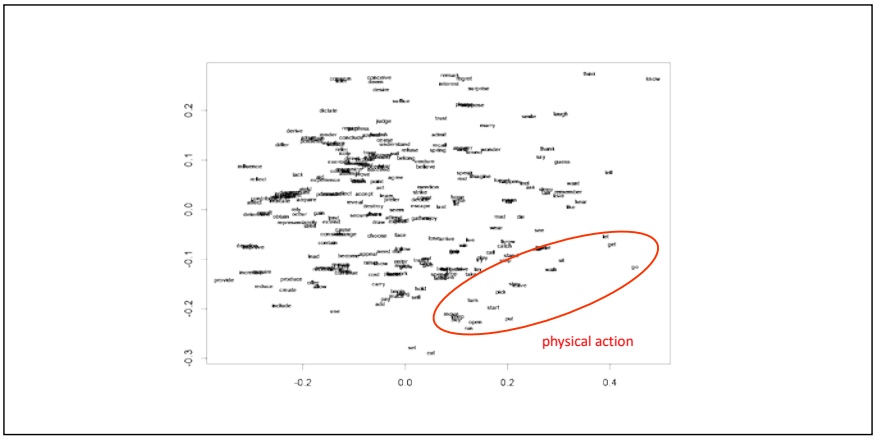


FIGURE 26

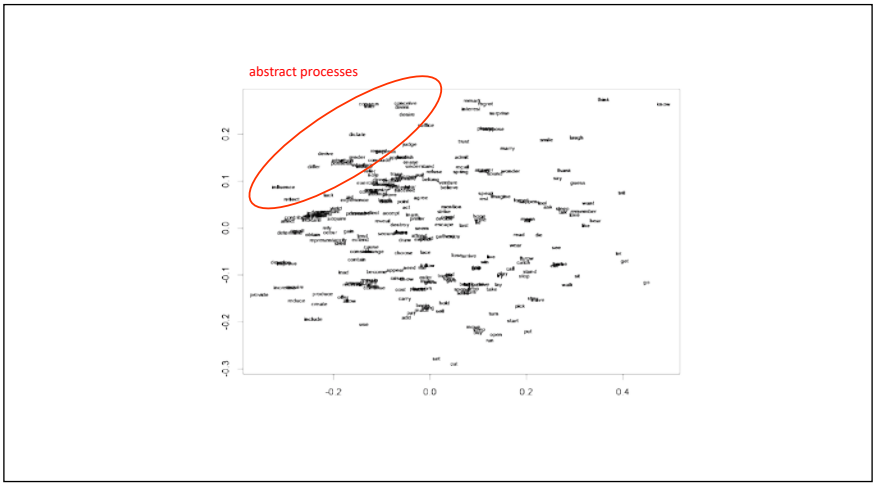


FIGURE 27

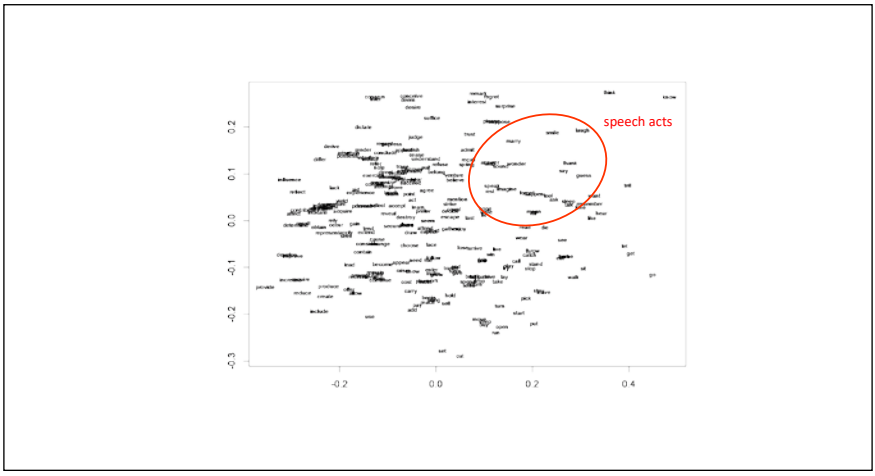


FIGURE 28

The graph you see on this slide is not unlike the graph with the different modal auxiliaries where you saw the bubbles, except that this one will not move. Here we have 250 different verbs that are arranged in terms of their semantics. It is impossible to read the labels and you do not have to. I will just explain to you what can be seen in this distribution.

In the lower right area of the graph, we have verbs such as *put*, *pick*, *walk*, *sit*, and *go* that encode physical action.

Towards the upper left we find verbs such *as differ*, *derive*, *concern*, *influence* and *reflect*. Those are verbs that encode abstract processes. If we draw

a line across the physical action verbs and the abstract processes, we have a continuum from concrete and physical actions to abstract processes. There is more that you can see.

In the upper right corner, we have speech act verbs such as *say*, *thank*, *answer*, *tell*, or *speak*. There are of course other verbs that you also find, which are not speech act verbs. The graph is not a perfect characterization of verbal semantics, but still you can see a lot of semantic structure in the distribution of verbs in this graph.

Overall, the further you go to the right in the graph, the more likely you are to find a verb that is very concrete, like *sit* or *walk*, and the further left you are on the graph, the more abstract the verbs tend to be, so *influence* and *reflect* and *provide* and *reduce*, those are relatively abstract verbs.

There is also a meaning that we can assign to the y axis of the graph. Further down on the graph, we find verbs that tend to be volitional, verbs like *put* and *run* and *open* and *cut*, which express volitional, intentional activities. Higher up in the graph, we have things like *regret* or *suffice* or *desire*, which are activities that are involuntary. I cannot decide if I desire something. It just happens to me. We have the semantic spectrum of verbs, and those are the collocates that occur with the modal *may* at a given point in time. What we can do now is trace the history of *may* and its connection to these verbs by overlaying this semantic space with the frequencies with which *may* occurs with all of these verbs. Let me show you how I did that.

The diagram on this slide shows where in the semantic space *may* select verbs most often. You see certain peaks here. In the middle of the graph, we have a peak around the verb *see*, indicating that *may see* is a very frequent combination. Another frequent area can be seen around the verbs *say*, *thank*,

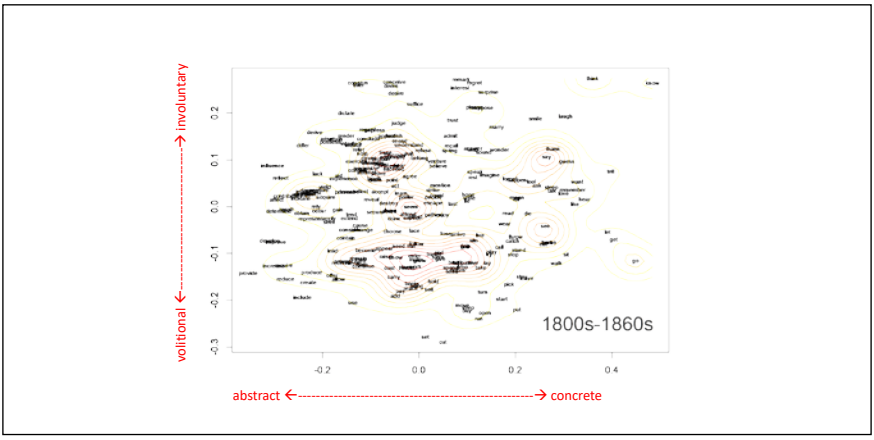


FIGURE 29

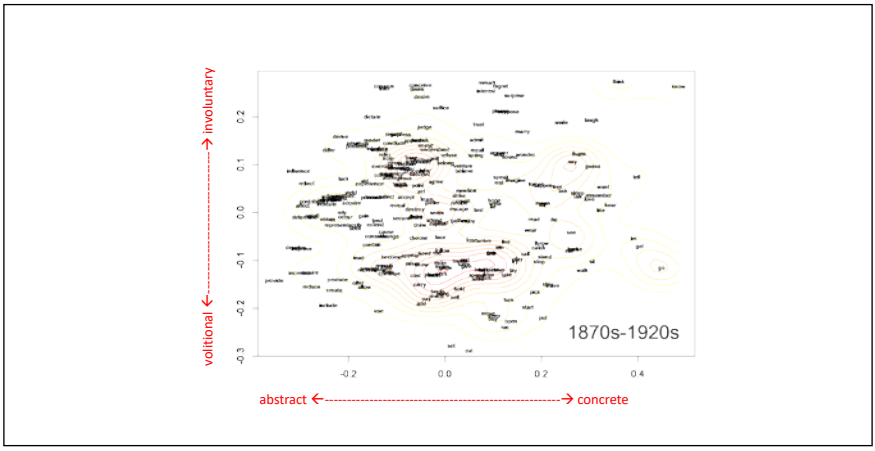


FIGURE 30

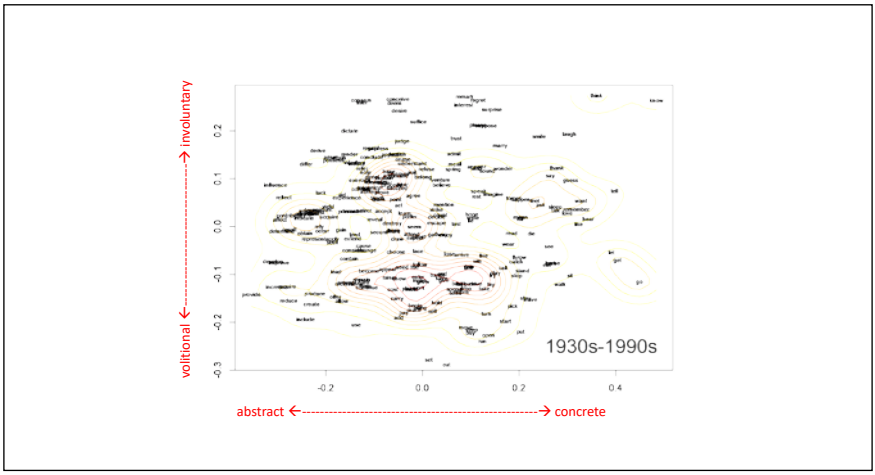


FIGURE 31

and *guess* towards the upper right. Right at the center is the verb *seem*, which is very frequent. In the outer areas, we have verbs that are less associated with *may* and that are less frequently used with it.

Overall, this kind of frequency profile, this kind of landscape represents *may* as it is being used in the 1800s to 1860s. As we move along in time, the frequency profile changes. Let's look at this in some detail. The current slide visualizes corpus data from the 1800s to the 1860s.

When we move on to the 1870s to 1920s, the collocate frequencies change, so that the peaks and valleys in the semantic space now appear a little differently.

Moving further along in time, these are the 1930s to 1990s. It is necessary to inspect these maps qualitatively for some time to actually take in all the

changes that are under development. One change that I would like to focus on is what happens to the frequency of *say*. We saw that *say* is an element that is distinctive for early periods, and this is corroborated by the present analysis. You see that in the first period, *may say* is a very frequent combination, and then as we move to the next period, this peak is already getting flatter, and in the last period it is gone. The high token frequency that was associated with *may say* has worn off over time and has become flat now. This would be an instance of the semantic landscape of *may* changing like a mountain range. It takes evolutionary time to really change a mountain range and what it looks like, but this is something that we can observe on the basis of historical corpus data with constructions like modal auxiliaries and their associations with lexical patterns.

In conclusion, the three questions that I had for this morning were the following: First, what's the object of studying Diachronic Construction Grammar? Second, when is a new construction a new construction? And third, what knowledge is represented by the nodes and by the connections between them? One important issue in this context is the issue of connectivity changes. How do new links emerge in the constructional network through meaning extension and other processes? One general conclusion that I would like to advance here is that research in Diachronic Construction Grammar and Construction Grammar more generally has a lot to gain by focusing more on connections. For the project of Diachronic Construction Grammar, this means engaging head on with the phenomenon of connectivity change. In principle, changes in connection strength, the empirical results that I have shown you today, chiefly relate to changes in connection strength, how associative links between two constructions can become stronger or weaker, and how construction-internal links between form and meaning can become stronger or weaker.

I do want to stress the fact that there is exciting ongoing work that has very similar goals. Yesterday in Lecture 1, I mentioned constructional contamination (Pijpops and Van de Velde 2016). There is also interesting work by Tiago Torrent (2015) who has developed two hypotheses, "the constructional convergence hypothesis" and "the construction network reconfiguration hypothesis". This kind of work I think goes precisely into the direction that Diachronic Construction Grammar should take, namely, towards the formulation of testable hypotheses about changes in the constructional network. In the next lecture, I will address in more detail how shifting patterns of associations between constructions and lexical items can be analyzed, and what theoretical conclusions we can draw from such analyses. I hope that the taste that I have given you in this lecture has already prompted you to think about how these ideas could be explored further. With that I would like to thank you for your attention, and I am looking forward to the next lecture this afternoon. Thank you.

LECTURE 4

Shifts in Collocational Preferences

Welcome back to Ten Lectures on Diachronic Construction Grammar. In the lecture this morning, one of the issues that I was concerned with was the open question in Diachronic Construction Grammar of how we can make our analysis more focused on the connections between constructions, rather than maintaining a focus on the internal structure of the nodes in that network. I called that the "fat node problem". To address that problem, I have discussed the example of the English auxiliary *may*, which has over time come to be used with a very different set of lexical collocates. You remember the journey of *may* in the system of modals that reflects its changing connections with the lexical constructions that it occurs with. In this lecture, I want to expand on the issue of shifts in collocational preferences. I want to discuss a number of case studies that allow us to assess the theoretical conclusions that we can draw from observing shifts of this kind.

One important point is that in the development of grammatical constructions, these collocational shifts are not just random events. In the lexical domain, by contrast, anything can happen. You remember from yesterday the collocate display of *gay*, and its early collocates like *gay colours* and its later collocates like *gay community*. The collocational shifts show us that some meaning change has taken place, but there aren't any broader implications of this beyond that. This is what happened to the adjective *gay*, and other lexical elements may change in completely different ways.

With grammatical constructions, however, shifts and collocates tend to reflect more systematic patterns of change, such as the ones proposed by grammaticalization theory.

Remember also the proposal by Traugott and Trousdale (2013) that grammatical constructionalization involves an increase in schematicity. We expect grammatical constructions to broaden in their collocational behavior over time. We expect them to occur with more types of lexical elements. We expect

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691059.

- Israel, Michael. 1996. The *way* constructions grow. In Adele E. Goldberg (ed.), *Conceptual structure, discourse and language*, 217-30. Stanford: CSLI Publications.

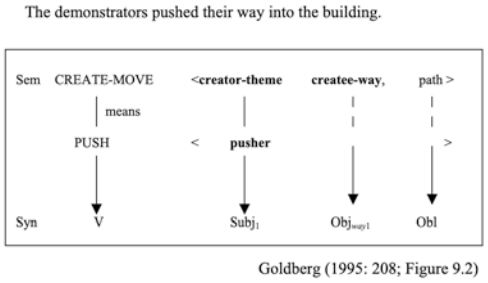


FIGURE 1

these types to have different meanings, or an increasing range of lexical meaning. That is the kind of phenomenon that I want to look at in this lecture. With all of that in mind, let's get on the way. Coincidentally, what got me on my way with regard to Diachronic Construction Grammar was a paper by Michael Israel (1996) with the curious title "The *Way* Constructions Grow". This title very cleverly blends two ideas. On the one hand, the paper addresses by now well-known pattern that is called the *way*-construction. It presents a study of *way*-constructions and how these constructions grow. At the same time, the paper explains how this process unfolded, as it discusses the way in which constructions grow. It is terrible of me to explain the joke, you simply have to forgive me.

What you see on the slide is an example and a representation of the *way*-construction, taken from the work of Adele Goldberg (1995). The example is "*The demonstrators pushed their way into the building*". The diagram, the box with the semantics and syntax and the arrows and the arguments, that is meant to capture that the construction can take a verb such as *push*, that has the meaning of creating something and moving something at the same time. The construction adds two arguments that are not normally called for by the verb. That is what the dotted lines mean. *Push* requires a pusher, i.e. the subject. But the two other constituents are actually added by the construction. They are not part of the ordinary argument structure of the verb. The first of the two arguments that are added by the construction is an object that encodes a created way. Goldberg calls this the createe-way argument. Then there is a

prepositional phrase, Goldberg calls it the path here, that expresses either a path or a location. In this case, it is a path: "into the building".

How did Michael Israel analyze this construction? His starting point is an observation that Goldberg made about the construction. The construction is polysemous. It can express two basic ideas. It can express manner of motion as well as a means of motion. Manner of motion is expressed in examples such as "*The wounded man limped his way across the field*". The movement is carried out in a limping, effortful manner. By contrast, in an example such as "*Joe cheated his way into law school*", cheating is a means to achieve a result. There is a metaphorical movement. Joe is metaphorically moving into law school and cheating allows the subject to carry out that movement. The *way*-construction is what Goldberg calls an argument structure construction, which is to say that in synchronic usage, the construction can modify the argument structure of the verb that it takes. It imposes an unusual argument structure on the verb, in these examples the verbs *limp* and *cheat*. You notice that I can use the verbs *limp* and *cheat* intransitively. I can say "*The man was limping*" or "*John, he is always cheating*". There are other argument structures that go with these verbs, but I can use them intransitively, and the *way*-construction allows me to use new and additional arguments with these verbs, namely a created way argument and a path or a goal argument.

Michael Israel looked at this construction by tracing its usage history in historical texts. He used a particular data resource for this, the Oxford English Dictionary, which is a dictionary that not only lists words and definitions. It also provides authentic bits of texts that illustrate how these words are used. What he found was that all the examples from the Oxford English Dictionary show that the construction established itself first with verbs that encode paths. What he found was that the early manner interpretation corresponds to examples with general motion verbs. Here is an example of this, "*The kyng took a laghtre and wente his way*". *Went* is just the past tense of *go*, a movement verb, the king laughed and went his way. The construction thus emerges in a very specific lexical context with verbs that harmonize with the overall meaning of the construction. That is not only true of the manner interpretation. It is also true of the means interpretation. Here the construction is used early on with verbs that explicitly encode the creation of path. Here we have an example, "*Arminius paved his way*". You pave the surface and after that you have a way to move along on.

Now, looking at these examples at this stage, you could argue that the *way*-construction has not fully constructionalized in the sense of Traugott and Trousdale, but rather what we have here is a more or less transparent

compositional use of a particular argument structure and verbs that harmonize with that argument structure. This, if you like, is the *way*-construction coming into being, constructionalizing in the sense of Traugott and Trousdale.

As time goes on, the construction goes through a type frequency increase. More and more different verb types occur in that particular argument structure pattern with the manner interpretation. Michael Israel counted all the verbs. Up to 1700, the manner interpretation is relatively sparse in the data. There are only 16 different verb types, including *go*, *pass*, *run* and a few others. Then, from the 19th century onward, there are 38 more that he finds. *Worm* here means to move like a worm. *Fumble* is a slow and clumsy movement. With regard to means, this starts earlier, so Israel finds examples as early as 1650. There we have the path-creation verbs such as *pave* or *smooth* or *cut*. You cut your way through a jungle or through some difficult terrain. Later on in the 18th century, he finds verbs that metaphorically extend this kind of path-creating idea. There is *battle*, there is *fight*. Then from the late 19th century onward, we find verbs that we find also nowadays in the *way*-construction, verbs like *elbow*, *shoot*, *spell*, and even things like *write*. You can "write your way out of a difficult situation". You've offended someone and you write them an apology. You can write your way out of a hot mess that you've gotten yourself into. The increasing type frequency maps onto increasing degrees of schematicity and abstraction. This instantiates one of the pathways that Traugott and Trousdale proposed for grammatical constructionalization, i.e. increases in productivity, increases in schematicity, and decreases in compositionality. When we have something like "*He elbowed his way out of the subway*", the meaning component of difficult laborious motion is carried holistically by the entire construction, rather than by any individual element on its own. When I read Michael Israel's paper, I took away three lessons. Those lessons for me were the following.

First of all, I understood that constructionalization happens in the context of particular collocations. There are certain words and constructions that are used together. This starts the process. Second, shifts in collocational preferences reflect changes in constructional meaning. As we use a construction with more and more different verbs, our idea of what this construction can do changes along with it. Constructional change, and that is the third conclusion, shows itself in changing relative frequencies and type frequencies of lexical collocates, not necessarily in their absolute frequencies. One curious thing about the *way*-construction is that for the longest time, *make* actually has been the most frequent verb in this construction. That has been relatively constant. But there have been a lot of developments going on under the surface of the most frequent elements.

There is one further conclusion from the paper that I would like to read to you in the words of Michael Israel himself. This captures very much the way I have come to think about constructional change. Here it is.

> The way-construction emerged gradually over the course of several centuries. There is no single moment we can point to and say, 'This is where the construction entered the grammar.' Rather, a long process of local analogical extensions led to a variety of idiomatic usages to gradually gain in productive strength even as they settled into a rigid syntax. As the range of predicates spread, increasingly abstract schemas could be extracted from them and this in turn drove the process of increasing productivity.

Here Michael Israel actually prefigures the main aspects of Traugott and Trousdale's concept of constructionalization, with its aspects of compositionality, schematicity and productivity. More specifically, the development of the *way*-construction clearly illustrates the phenomenon of what I talked about as host-class expansion yesterday and earlier today. I would say that Michael Israel's paper on the *way*-construction was one major inspiration for the research that I then did in my doctoral dissertation.

In my dissertation, I investigated future tense constructions across a range of languages from the Germanic family. The main overarching aim of that book was to see if we can use shifting collocational preferences to study on-going semantic change. Here is where the theoretical parts of what I have been talking about so far meet the methodological parts. I chose future constructions for my study because a lot is known about the grammaticalization of future tense markers. They tend to come from a handful of lexical sources such as movement verbs (English *be going to*), verbs of desire (English *will*, which derives from a verb meaning 'want' and 'desire'), or verbs of obligation (English *shall* falls into that category). There are verbs that mean 'turn' or 'change'. There is no English construction for that, but German has one, namely *werden*, which meant 'turn' originally. There are verbs that express intentions or at least relate to intentions. Swedish, for example, has a future construction with a verb that means 'think'. When I say *I think going to the movies*, it over time turns into an expression that tells my interlocutor I will be going to the movies later today.

Grammaticalization scholars have proposed very specific developmental pathways for these constructions. Movement verbs are known to turn into markers of intention, and after that into markers of future time, and after that into yet other meanings. I was curious to see whether these proposed grammaticalization pathways could be shown to be reflected in historically shifting

patterns of collocations. This turned out to be true. You can actually see these trajectories.

My study compares future constructions from five different Germanic languages. It is based on historical corpus data exploring both the synchronic meaning and the historical development of these future constructions. It applies a method that I briefly talked about this morning, collostructional analysis. I will say more about this as we go along here. The method allows us to measure how a particular future marker is connected to lexical verbs that occur with it. It allows us to determine what associations exist and how strong they are.

More importantly, we can also use it to investigate how these patterns of association change over time. Michael Israel looked at this from a qualitative point of view. He took examples from different historical periods. He noted what kind of verbs came first, what kind of verbs were added later, and how the changes reflected increasing degrees of productivity. When I say from a qualitative point of view, that is not entirely true. He did count the types, but that very much stays at the level of descriptive statistics, not inferential techniques.

In my studies, I have been fortunate to have two great teachers and mentors, Anatol Stefanowitsch and Stefan Gries, who developed a technique for the analysis of collocational relations between constructions and lexical items. This method became available just when I needed it. It fell into my lap at the exactly right time. I am of course talking about collostructional analysis. This is a method that allows you to quantify how strongly a set of lexical items is associated with a grammatical construction that has an open slot for these lexical items. Stefanowitsch and Gries developed collostructional analysis for the study of such associations in synchronic present-day usage. But one day, when I was riding my bicycle home from university, I wondered whether it would be possible to tweak the method just a little bit, so that I could use it to study change over time.

I wrote an email to Anatol later that day and asked him whether he thought it could be done. He emailed me back and told me to try it. As a dutiful student, I followed his advice. I developed the idea and publish a short paper with the title "Distinctive Collexeme Analysis and Diachrony" (Hilpert 2006), and in the same issue of the journal then Anatol published a short reply (Stefanowitsch 2006), summarizing all the points that he found problematic about it. I am getting ahead of myself here. Let me explain what collostructional analysis is all about.

The basic idea is one that I mentioned in lecture one, namely that constructional meaning is reflected in associations between syntactic patterns and lexical elements. The fact that *give* is the most frequent verb in the ditransitive

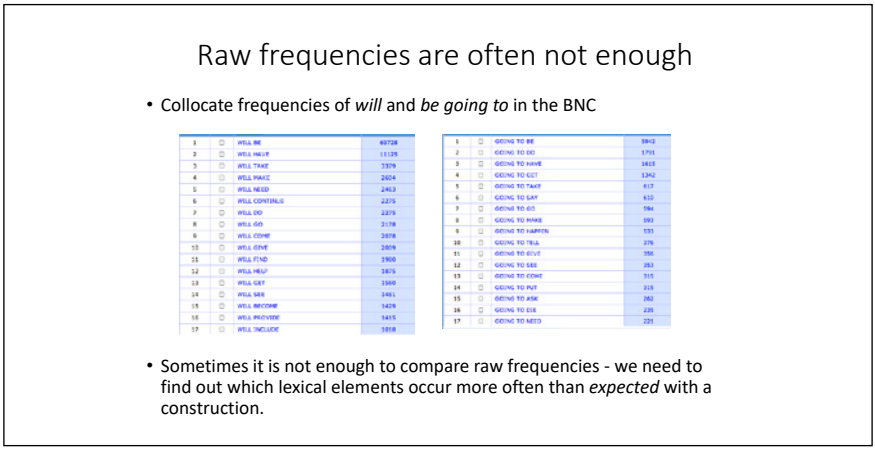


FIGURE 2

construction, and the fact that we find the verb *elbow* in the *way*-construction, that illustrates this harmony in meaning between grammatical constructions and the lexical items that occur within them. Let me just read this to you again. Stefanowitsch and Gries (2003: 236) motivate this in the following way: "*If syntactic structures served as meaningless templates waiting for the insertion of lexical material, no significant associations between these templates and specific verbs would be expected*". But no matter where we look, pretty much any syntactic pattern that has been investigated in this way shows asymmetries that demonstrate that these distributions differ from chance in many ways. There is no random distribution of lexical elements across syntactic constructions.

This also relates to the controversy of collostructional analysis versus raw frequencies. This is actually a good moment to explain what is really at the heart of this controversy. Sometimes it will turn out that when you look at the lexical elements in a grammatical construction, the raw frequencies won't tell you a great deal. Sometimes looking at raw frequencies is just not enough. As an example for that, this slide presents two lists of collocate frequencies from two near-synonymous grammatical constructions of English, namely the English *will* future and the English *be going to* future, and the most frequent lexical verbs that occur within them. If you go through these two lists, you notice that they are almost identical. We start with very frequent verbs like *will be*, *will have*, *will take*, *will make* and so on and so forth. With *going to*, we have *going to be*, *going to do*, *going to have*, *going to get.* Basically, you're just getting these long lists of semantically light, very general, very frequent verbs. You might look at these two lists and conclude that they are almost indistinguishable. These two constructions have similar functions, so they occur with similar sets

of verbs. However, if we find out which elements are not only frequent, but actually more or less frequent than expected, then we can say something more about these constructions. I mentioned collostructional analysis briefly earlier this morning. Let me go back to this and explain the general logic.

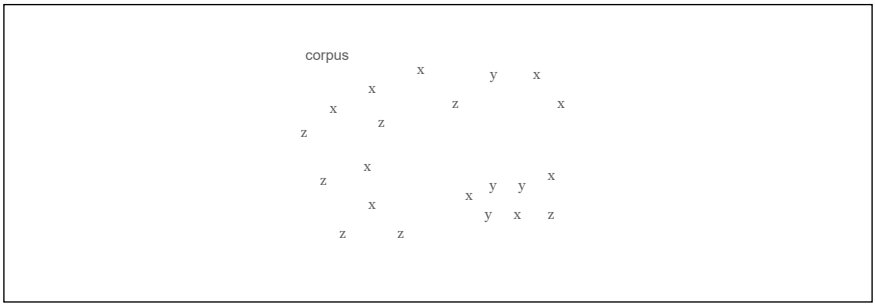


FIGURE 3

Let's say that you have a corpus that has lots and lots of different word types in it. Here I have symbolized them with letters *x*, *y* and *z*. Let's say that you extract from that corpus all the instances of a construction and count the lexical elements that you find in that construction. In this case, we have a very small sample.

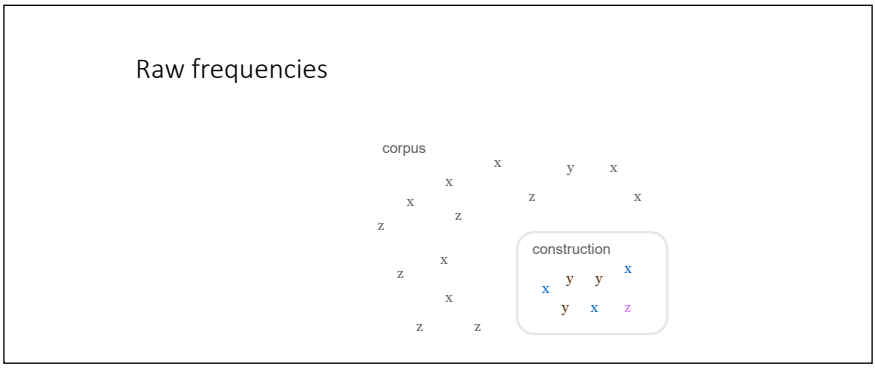


FIGURE 4

The construction occurs with seven lexical items. These seven lexical items fall into three types, *x*, *y* and *z*. Three times *x*, three times *y*, and one *z*. If we were to stay with these raw frequencies, we would say that, for this construction, *x* and *y* are of approximately equal importance.

However, as soon as we compare the frequencies of these items within the construction to the frequencies of these items outside of the construction in the corpus, it becomes apparent that *x* is an element that we find throughout

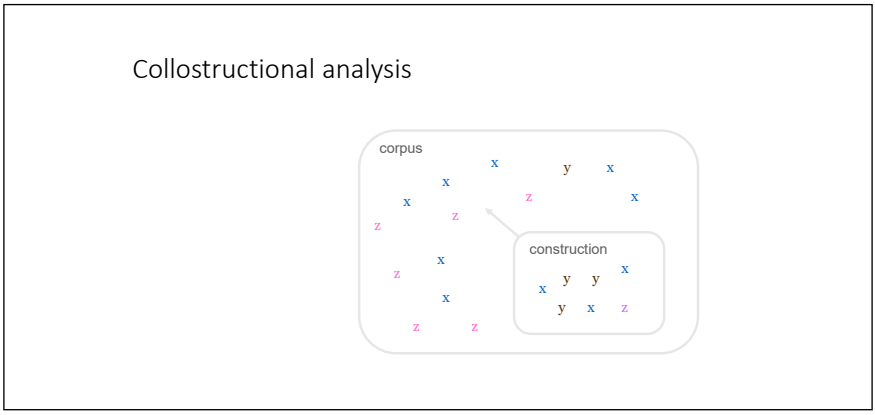


FIGURE 5

the corpus with high frequency. This is in contrast to *y*, which we find three times in the construction, but only once in other contexts. This would be similar to the case of *elbow* in the *way*-construction. *Elbow* is a very infrequent verb. To the extent that we find it in a corpus, it will tend to appear in the *way*-construction, but not anywhere else, at least not very frequently.

This, in a nutshell, brings me to the controversy between collostructional analysis and raw frequencies. The view of focusing on raw frequencies only chooses to ignore how often a lexical element appears not only in the construction, but also in the corpus as a whole.

I have a quote from Gries *et al.* (2005: 665) here, where they say, "*Arguing and theorizing on the basis of mere frequency data alone runs a considerable risk of producing results which might not only be completely due to the random distribution of words* [*in a corpus*]*, but which may also be much less usage-based than the analysis purports to be*." It may be questionable to say that words in a corpus are randomly distributed. Language does not work that way, but the point stands that we want to account for these occurrences that occur outside of the construction. Gries et al. (2005) have done empirical work that supports the better adequacy of collostructional results over raw frequencies. Let me review that for a minute.

When we have a verb as a cue for a construction, what is it that determines the validity or the usefulness of that cue? According to Bybee's view, the most frequent verb that appears in a construction should be the best cue for that construction. According to Stefanowitsch and Gries (2003), it might not be the most frequent verb, it might be the verb that is most strongly attracted to the construction that provides the best cue, even if that verb is not very frequent. If all of its instances are found with the construction, as in the case of *elbow* and the *way*-construction, that would make for a very reliable, very useful cue.

Let me give you an example. The most frequent verb in the *way*-construction is the verb *make*. Now, when I say *I made*, does that make you think of the *way*-construction? Does it make you want to continue that sentence with *my way through the city*? *I made* can be continued in lots of other ways. *I made a mistake*, *I made her a sandwich*, and lots of other things. It is not the best cue for the *way*-construction.

By contrast, the verb *elbow*, if I start an utterance with *I elbowed*, there are not many ways to continue that sentence. *I elbowed my way out of the subway*, that would be a very natural continuation of that sentence fragment. This is the methodology that Gries and colleagues applied to investigate which verbs in a carrier phrase would prompt speakers to continue with a specific construction. The construction that they used as an example is a construction that they called the English *as*-predicative construction. It is instantiated by sentences such as "*The idea was perceived as too radical*". We have a verb like *perceive*, and then a prepositional phrase with *as*, and a certain predicate, so *radical* is predicated over the idea.

Here are three examples of the *as*-predicative construction: *The proposal was considered as rather provocative*; *I had never seen myself as being too thin*; *California is perceived as a place where everything is possible*. There are different verbs that appear in this construction. One of them is the verb *see*, as in sentence fragments such as *I have never seen* or *I had never seen.*, which you could continue that with the *as*-predicative. Some verbs give you a very strong cue, like the verb *hail*. When I start a sentence fragment with *The idea was hailed*, the *as*-predicative almost forces itself upon my mind. So, some verbs make you think about the *as*-predicative in very direct ways, and other verbs do not. The question now is which verbs are which. What determines cue validity? Is it frequency in the construction, or is it attraction to the construction that we measure via collostructional analysis?

four types of verb

	Relative Frequency	
	High	Low
Surprisingly high frequency in the As-Predicative	define, describe, know, recognize, regard, see, use, view	acknowledge, class, conceive, denounce, depict, diagnose, hail, rate
Surprisingly low frequency in the As-Predicative	keep, leave, refer to, show	build, choose, claim, intend, offer, present, represent, suggest

FIGURE 6

Gries and colleagues compared four different types of verb for their design. First, they took sets of verbs that are generally frequent in English. These are verbs such as *define*, *describe*, *know*, *recognize* and verbs like *keep*, *leave*, *refer to* and *show*. Some of them are surprisingly frequent in the *as*-predicative, for example *define* and *describe*. Some of them are very frequent in general but surprisingly infrequent in the *as*-predicative. This is true for *keep*, as in *He was kept as a slave*. That does not appear very often. In the right column of this table, you see verbs that are low in frequency. Some of them are surprisingly frequent in the *as*-predicative, for instance *conceive*. It is not a very frequent verb, but I can say "*This was conceived as the solution for that problem*". *Depict* is another surprisingly frequent collocate, and *hail* is yet another one. Finally, some verbs are infrequent in general, and also surprisingly infrequent in the *as*-predicative construction. This includes *suggest*, "*This was suggested as a possible solution*".

On the raw frequency view, high frequency verbs should pattern alike, no matter whether they are surprisingly frequent or surprisingly infrequent in the *as*-predicative When speakers see a sentence fragment with a high frequency verb, there should be a high probability that they continue the fragment with the *as*-predicative.

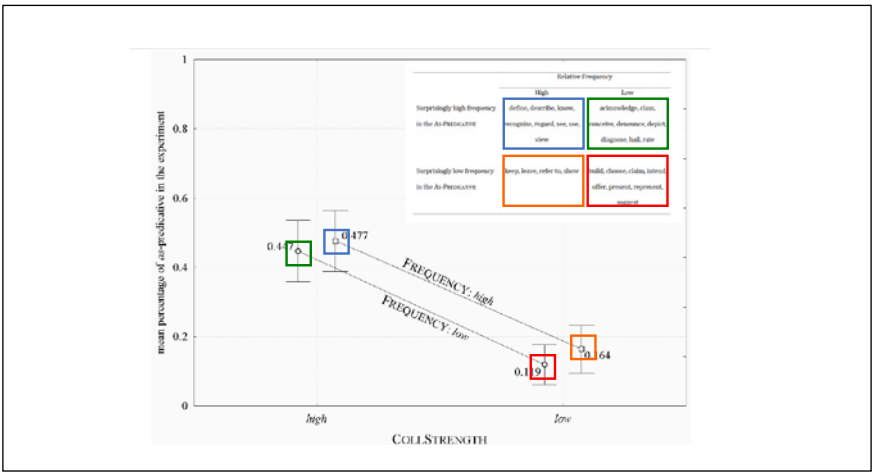


FIGURE 7

On the collostructional hypothesis, however, attraction to the *as*-predicative should matter more than just raw frequency. The verbs that are surprisingly frequent in the *as*-predicative should pattern together with low frequency verbs such as *hail* and *depict*. That means that when speakers see a sentence fragment with these verbs, they should be more likely to continue the fragment with an *as*-predicative construction.

Here's what came out of the experiment. Here you see a chart (Gries et al. 2005: 659) that shows the rate of completion with an *as*-predicative construction that was obtained in the experiment. The *y*-axis shows you how many of the participants chose to continue a given sentence fragment with the *as*-predicative. You see that there are two types of verb that are higher up and two types of verb further down.

Higher up you see the verbs that are strongly associated with the construction. There are the frequent verbs and the infrequent verbs, but both of them are surprisingly over-represented in the construction. Down below you see the verbs that are not strongly associated with the construction. If we look at the orange box here, these verbs are frequent in the *as*-predicative. Bybee would predict that those should trigger a lot of completions with the construction, but it does not. There are not many completions with the *as*-predicative. Up here is the green box. These verbs are infrequent in the *as*-predicative, but they are strongly associated with them. This would be similar to the case of *elbow* in the *way*-construction. For the *as*-predicative, it is *hail*. It is not very frequent, neither inside nor outside the construction, but when you see *hail*, then it is a very reliable cue for the *as*-predicative.

In summary, collostructional strength, the strength of association between a construction and a lexical item, matters. Raw frequencies matter too. You see that the frequent verbs are slightly above the infrequent verbs, but that is a secondary factor.

Let me come back to how collostructional analysis works. What I have shown you so far is what is known as collexeme analysis, the basic type of collostructional analysis that compares frequency within a construction against lexical item frequency in the corpus as a whole.

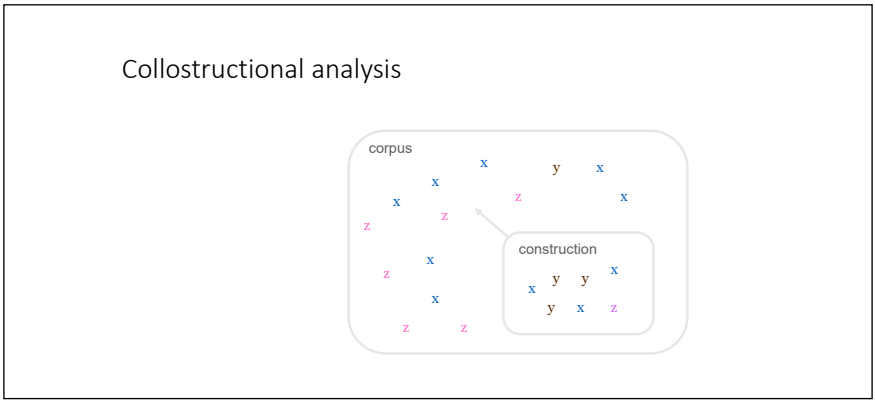


FIGURE 8

The analysis type that I briefly discussed this morning is distinctive collexeme analysis, where we compare collocate frequencies across two different constructions. We have construction A, which has a number of collocates with different frequencies. We have construction B with the same types at different frequencies. We can figure out which elements are maximally uneven in their distributions. Which elements have the greatest frequency asymmetry between construction A and construction B? That can give us a cue as to what makes these two constructions different. That kind of contrast, that kind of analysis makes sense when we are dealing with constructions that are related in some way, that have similar functions like *will* and *be going to*. Or think of different complementation patterns, *that*-clauses and *ing*-clauses, or the *get*-passive and the passive with *be*. You could also contrast broader tense patterns like the simple present and the present progressive. That would show you the different collocational preferences that these construction types have. Depending on the level of abstraction of these constructions, there are different observations you could make. So far, I have presented the synchronic way of conducting collostructional analysis. Let me get to its diachronic application.

The idea that I had on my bicycle, going home from university, was that we have a diachronic corpus with data from different historical periods, for example starting in the 1600s, then the 1700s and then 1800s and so on and so forth. We look for the same construction across different time slices of the same corpus. That enables us to identify asymmetries across these time slices.

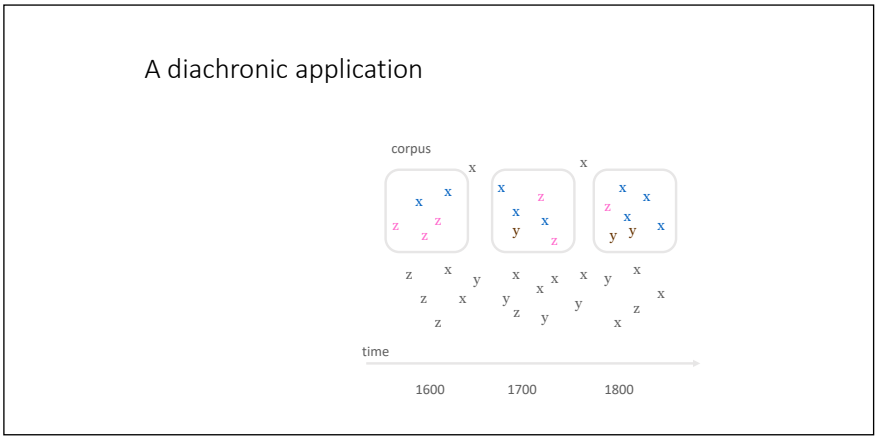


FIGURE 9

We can examine the collocates that a construction has in the 1600s, and we can ask whether they are the same as the ones that we find in the 1700s and in the 1800s. We can identify elements that have become more or less frequent

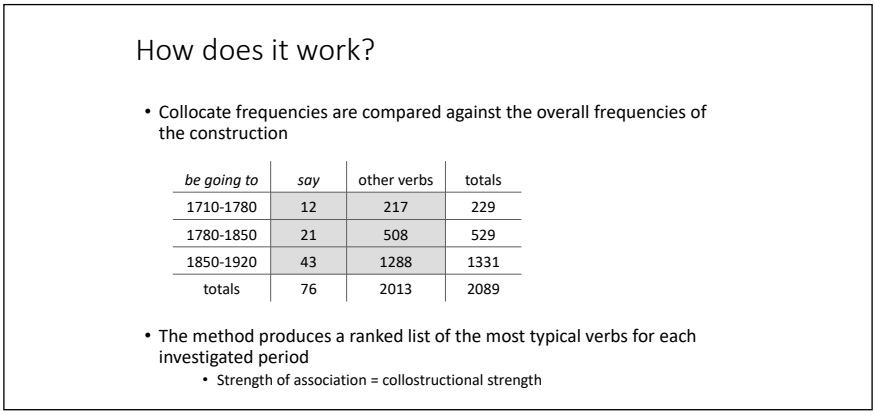


FIGURE 10

over time. We can identify elements that are typical for a particular section of the corpus, and we can draw conclusions from that.

The way it works in practice is that the collocate frequencies of a construction are compared against the overall frequencies of that same construction for each collocate that you observe. You see an example here of the *be going to* construction that is used with the verb *say* across three periods of time.

We observe the raw frequencies of *say*. They happen to increase. From 12 to 21 to 43, but then you see that also the construction as a whole increases in frequency as well. It goes from about 230 examples to 530 examples to more than 1300 examples. So *say* goes up in frequency, but also the construction itself goes up in frequency. Does that mean that *say* becomes more or less attracted to the construction? Does that mean that the attraction stays at the same level? This is something that we can figure out on the basis of a statistical analysis that takes all the verbs and their frequencies into account. The method produces a ranked list of the most typical verbs for each investigated period. I have used the term collostructional strength without properly defining it. Collostructional strength would be the strength of association between a construction and a lexical item that occurs within it, as measured by a collocation statistic.

What does this type of analysis show? It determines the elements that are most typical for a construction in modern usage, when you analyze a synchronic corpus. Applied diachronically, we can also determine the most typical elements for a construction in any given historical period. The collocational preferences can be used to describe the modern semantics and how it came to be that way. That was important for me. One other thing that attracted me to the method was that if we find that there are systematic changes, that there are attracted sets of semantically related collocates, not just individual items,

then that would suggest that there is a broader trajectory of semantic change going on that reaches beyond just individual words, individual histories of lexical items.

Why would this be useful? The method allows explorative analyses of constructions and their diachronic variation, how they have developed. More usefully perhaps, it allows us to test proposed semantic developments. I mentioned grammaticalization paths that thave been proposed for the development of future constructions. My idea was to go into the data and see if these proposals that have been made could be falsified or substantiated. This is especially useful if we want to distinguish between competing hypotheses, which you can often find in the grammaticalization literature. One account might claim that this future construction developed in this way, and another account might propose a very different semantic pathway. How do we decide between the two? This method actually can get to the bottom of what kind of meaning came first, what kind of meaning it developed later, and how it all ended up.

Semantic classes of distinctive verbs would indirectly reflect stages of grammaticalization paths that have been proposed for the development of future markers, for example, in the work by Joan Bybee and colleagues (1991, 1994), who have argued for a path from the meaning of intention to the meaning of future and from thereon to epistemic or speaker-oriented modality.

To give you a taste of how all of this can be applied to concrete case studies, let's look at an example. So far, I have exclusively shown you data from English, so I think it is high time to broaden the outlook a bit. You know that English has a future with the modal auxiliary *will*. What you perhaps do not know is that a small language close by, Danish, has a *vil* future as well. The word looks almost exactly the same, etymologically it is the same.

There are further parallels in that the construction in its synchronic usage shows a preference for certain types of verbs. It is highly productive. It is highly general, but there are preferences, there are asymmetries. The construction prefers abstract atelic verbal complements such as verbs like *require*. There is nothing dynamic or agentive about *require*. It is a state. A Danish verb that means 'to be' is also among the most attracted items.

My questions for this particular case study were the following: Can the semantic developments be described in terms of shifting collocational preferences? Were there certain semantic verb classes that were central to the development? Can the collocates address the hypothesis that future constructions of this kind develop out of markers of intention? That would be the default hypothesis. If we have a verb that means "want", then it is a marker of intention that may eventually shade off into a use that encodes just future time reference, so that also inanimate subjects can be. used with it. I can say things like, *It will rain*, which does not mean that *It wants to rain*, but rather that it will happen in the future.

Data

- raw-text historical corpora

Corpora	Century	Size	Search	Ex.
ACOD	12-14	36 k	*vil*, *vill*,	99
DSST 1	15-16	400 k	*wil*, *will*	642
DSST 2	17-18	240 k		408
Gutenberg	19-20	700 k		999
Total		1.4 M		2,148

FIGURE 11

Here's an overview of the kind of data I used. A collostructional analysis does not require you to have millions and millions of words. In this case, I had four different historical periods ranging from the 12th to the 20th century. It is a long and thin corpus really, with all in all about 1.4 million words. That, by today's standards, is considered very small. I looked for different orthographic variants of the auxiliary. I had a total of some 2000 examples to work with.

Absolute frequencies

ACOD (12-14)			DSST (15-16)			DSST (17-18)			Gutenberg (19-20)		
Verb	**Gloss**	**N**	**Verb**	**Gloss**	**N**	**Verb**	**Gloss**	**N**	**Verb**	**Gloss**	**N**
give	*give*	9	være	*be*	44	have	*have*	27	være	*be*	126
have	*have*	9	gøre	*do*	42	være	*be*	27	sige	*say*	84
tage	*take*	8	give	*give*	38	gøre	*do*	20	have	*have*	82
vide	*know*	5	have	*have*	31	sige	*say*	20	blive	*become*	41
lade	*let*	4	sige	*say*	26	give	*give*	19	se	*see*	29
søge	*seek*	4	lade	*let*	19	gå	*go*	11	gå	*go*	27
fare	*travel*	4	blive	*become*	16	bevise	*prove*	10	gøre	*do*	21
gøre	*do*	3	tale	*talk*	15	lade	*let*	10	få	*get*	19
svare	*answer*	3	bevise	*prove*	10	tro	*believe*	10	tage	*take*	19
mæle	*speak*	2	ride	*ride*	10	forlade	*forgive*	9	komme	*come*	14
bytte	*swop*	2	skrive	*write*	10	komme	*come*	9	finde	*find*	13

FIGURE 12

When we examine the absolute frequencies of the most frequent collocates across the four periods, there are already a few observations that we can make.

For example, there is a verb meaning "give", which is the most frequent one in the first period, after which it decreases. Eventually it disappears from the list of the most frequent elements.

In the opposite direction, there is a verb meaning "say" that does the exact opposite. First, it is not among the most frequent verbs, but then it gradually works its way up the list. In the last period it is on position two. But overall,

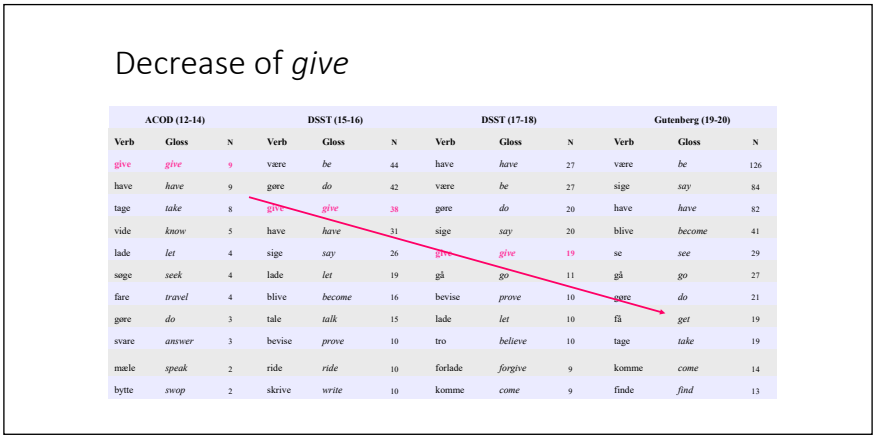


FIGURE 13

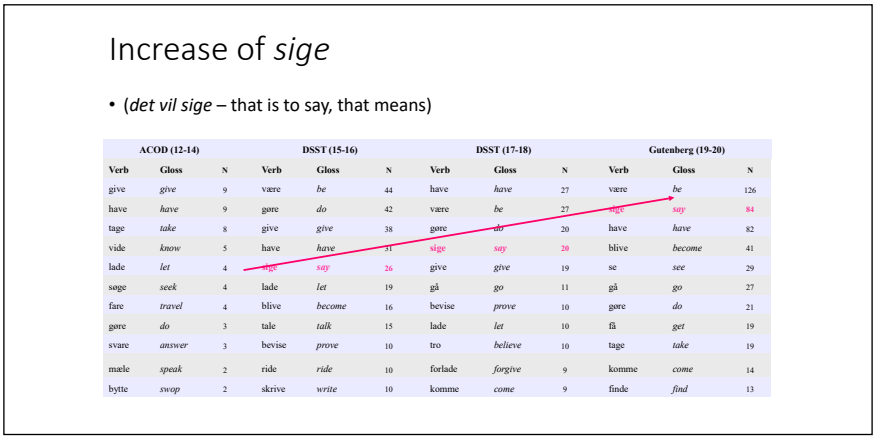


FIGURE 14

when we examine the raw frequencies, most of what we see are highly frequent verbs, like *have*, like *be*, *do*, *give*, *take*, and *go*. These verbs, for the most part, do not have a whole lot of semantic substance, and thus they do not allow us to say much about how the construction develops semantically.

The absolute frequencies revealed some tendencies, but not tangible developments. There is a constant of light verbs with high absolute frequencies, which could be taken to suggest that the changes that happened were either non-substantial or unsystematic. These could be seen as chance fluctuations in the lexical domain. If I had been working with a raw-frequency approach, I would have given up at this point with the conclusion that nothing has happened. However, I was curious to see if the collostructional approach would yield a different outcome, which it actually does.

Distinctive collexemes

ACOD			DSST 15-16			DSST 17-18			Gutenberg 19-20		
Verb	**Gloss**	**CS**	**Verb**	**Gloss**	**CS**	**Verb**	**Gloss**	**CS**	**Verb**	**Gloss**	**CS**
tage	*take*	3,5	ride	*ride*	5,2	bekende	*confess*	2,9	se	*see*	6,7
fare	*travel*	3,4	skrive	*write*	4,3	bevise	*prove*	2,8	være	*be*	6,2
søge	*seek*	2,8	gøre	*do*	4,3	tro	*believe*	2,8	sige	*say*	4,3
bytte	*swop*	2,7	give	*give*	3,8	forlade	*forgive*	2,5	spille	*play*	4,0
bøte	*pay*	2,7	antegne	*note*	3,7	føre	*lead*	2,3	blive	*become*	3,5
fange	*catch*	2,7	forklare	*explain*	3,7	æde	*eat*	2,3	få	*get*	3,4
løse	*solve*	2,7	råde	*advise*	3,7	nægte	*deny*	2,3	rejse	*travel*	3,0
mæle	*talk*	2,7	slås	*fight*	3,7	skikke	*send*	2,2	finde	*find*	2,8
rebe	*tie*	2,7	love	*promise*	2,1	vise	*show*	2,2	høre	*hear*	2,8
vide	*know*	2,6	tale	*talk*	1,9	bo	*live*	2,1	bringe	*bring*	2,4

FIGURE 15

Here's a table with the distinctive collexemes for each of the four periods. For each period you see the elements that are maximally over-represented in that particular corpus period. Let me briefly go through the periods individually and point out some developments along the way.

Period 1: ACOD

Verb	Gloss	CS
tage	*take*	3,5
fare	*travel*	3,4
søge	*seek*	2,8
bytte	*swop*	2,7
bøte	*pay*	2,7
fange	*catch*	2,7
løse	*solve*	2,7
mæle	*talk*	2,7
rebe	*tie*	2,7
vide	*know*	2,6

- Distinctive verbs require animate, intentional subjects.
- Examples with distinctive verbs express intention, not future.

tha a bonden siæluæ doom um at standæ hwilkæ bøtær han wil takæ

'Then the farmer can decide on his own what fees he wants to take.'

Wil man witæ of manz konæ gør hoor læggæ thænnæ steen undær hænnæ houæth.

'If you want to know whether your wife is cheating, put this stone under her pillow.'

FIGURE 16

In the first period, this is Old Danish, as it was written between the 12th and 14th century. We find exclusively verbs that require animate intentional subjects, things like *take*, *travel*, *seek*, *pay* and *catch*. Human beings can do that, inanimate beings cannot. If we are examining the examples with distinctive verbs, we find that these examples express intention, not future. We have things like *The farmer can decide what fees he will take*, that is, what fees he wants to take.

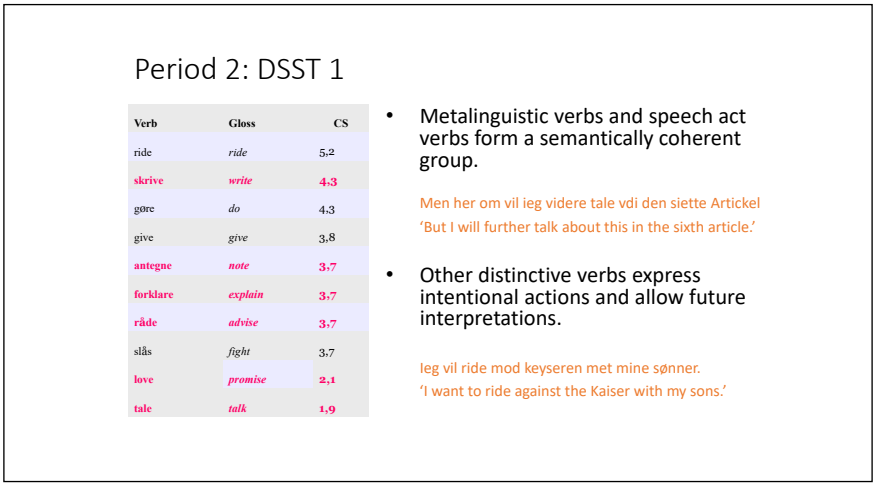


FIGURE 17

That will happen in the future, but the main idea is intention or volition. I love this example here: *If you want to know whether your wife is cheating, put this stone under her pillow*. *If you want to know*, you will know eventually. I am not exactly sure what's supposed to happen with that stone, though.

Moving on to Period 2, we find that the profile of verbs changes in that we see an over-representation of verbs that are metalinguistic, that encode speech acts. There are quite a few of them here. We have *write*, *note*, *explain*, *advise*, *promise* and *talk*, among the most attracted ones. There are a few other distinctive verbs that express intentional actions and that allow a future interpretation. The top collocate here is the word that means "ride", as in *ride a horse*. The relevant example that you have at the bottom of the slide here translates as "*I want to ride against the Kaiser with my sons*". Someone wants to do something, but it is also clear that this will happen in the future. This you can see as the future interpretation making inroads and establishing itself more strongly across the semantic spectrum of the construction.

In the last period, this trend of speech act verbs being strongly represented continues. Speech act verbs are still the largest coherent group of distinctive collexemes. The verbs *confess*, *deny*, or *forgive* express actions that you do linguistically, but the interesting thing that we see in Period 3 is the first occurrence of inanimate subject referents: *If your sins are about to lead you into desperation*. Your sins do not want anything. Your sins are actions, not intentional beings.

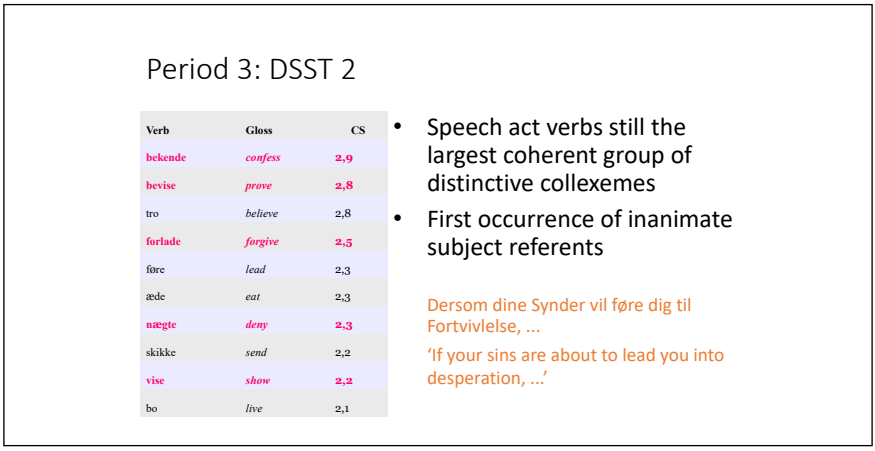


FIGURE 18

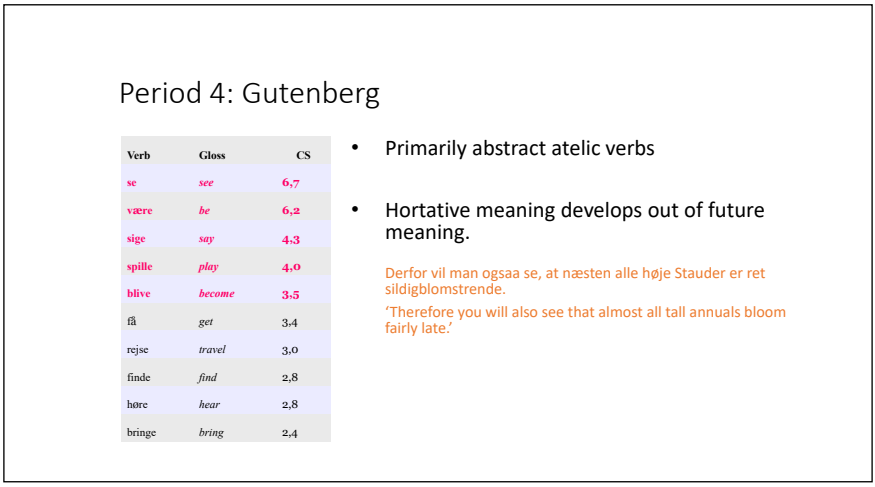


FIGURE 19

Finally, in the fourth period, which represents Present-day Danish, we see the profile that we are used to seeing in modern usage of the construction. The top collexemes are abstract atelic verbs, like *see*, *be* and *say*. There are a couple of others, and there are even extensions out of future meaning. There is a certain type of meaning that we can call hortative, which encourages someone else to do something. This occurs primarily in the context of *see*, *if you will see*. *You will see that almost all tall annuals bloom fairly late*. This is someone who is giving a piece of advice and who's inviting someone else to take this particular point

of view. That is something that does not necessarily encode future time, but it is an extension out of that temporal meaning.

What can we conclude from this? These data are corroborating the idea that future meaning with this construction developed out of intentional meaning. We see that very firmly ingrained in the construction's profile early on. You could say this is something that has been predicted all along, and we have lots of typological evidence for this particular pathway. What is the big deal? I would say the big deal is that the corpus-based material gives us a lot more detail than secondary resources that we can glean from descriptive grammars or even from native speaker's intuitions. One thing that definitely goes beyond the generic story of future tense development that has been proposed in grammaticalization studies is that we have this group of speech act verbs, which seem to be central to the development of this construction here.

There is a second example that I want to mention. For that, we will go from *will* futures to *be going to* futures. There is the English *be going to* future of course, but there is another small language close by, Dutch, which has its own *be going to* future, even though that form is a little different. It is not a progressive *going* type form, but it is a basic verbal form of a verb (*gaan*) that means 'go'.

These two constructions, Dutch *gaan* and English *be going to*, are often seen as more or less equivalent. They translate into each other. I can say "*It is going to rain tomorrow*" and there is an equivalent sentence in Dutch. I can say "*This is what we are going to do*". Again, there is a way to say that in Dutch. If I say "*I think that is going to happen*" with an inanimate subject, also this I can translate into Dutch almost word for word. The constructions share characteristics such as an orientation to the present, a preference for intentional or premediated actions, and there is a preference for dynamic events as opposed to the stative and atelic predicates that *will* in English prefers.

That was my starting point. I thought this would be a good opportunity to develop a parallel case study, but then I looked at corpus data, and mutual translations of these two constructions. I examined corpus data from European Parliament Proceedings, which features speeches that are held in different languages and which are translated into several other European languages. The same text is thus available in different languages.

I looked at a data set of more than 7500 examples of *be going to*. Out of those 7500 there are only about 1000 that are translated into Dutch with *gaan*. That is 15%, which is not much. Had it been half, I would have suspected a stylistic reason, so that translators want to avoid colloquialisms and therefore choose the more conservative construction with *will*. A rate of 15% however suggests

to me that the two constructions are functionally different. It also made me wonder how much overlap I should expect in the first place.

I looked at two other future constructions that are not etymologically related and that would not be considered as translational equivalents. I looked at the *will* future and the Dutch *zullen* future, which is etymologically related to English *shall*. I took a smaller sample of 500 examples, and I found 215 translations with *zullen*, so 43%. The overlap between these unrelated construction is significantly larger than the overlap that we find between *be going to* and *gaan*.

That sparked my interest. What semantic characteristics of *be going to* and *gaan* work against a mutual translation in present-day usage? Why are they not translated into each other more than they are? Second, did the two constructions drift apart only recently, or did they grammaticalize in different ways? That would be something of a surprise, because movement-based futures are thought to develop along similar lines.

My analysis was based on synchronic and diachronic corpus resources from English and Dutch. For modern English, I used the British National Corpus (BNC), a 100 million words corpus. For the historical part, I use the CLMET, which is a 9.5 million words corpus. I also used the Oxford English Dictionary, the source that Michael Israel used for the *way*-construction. The corpora for Dutch are smaller. The modern data comprise about 8 million words and the historical data from the Project Gutenberg hold about 4 million words. I exhaustively retrieved all examples with *going to* and *gonna* plus infinitive, and then the Dutch forms of the verb *gaan*, for which there are several morphologically inflected variants. I analyzed the data both in a synchronic way, with a collexeme analysis, and then diachronically with a diachronic distinctive collexeme analysis.

Comparing the collexemes

be going to		***gaan***		
Verb	**CollStr**	**Verb**	**Gloss**	**CollStr**
do	Inf	regenen	*rain*	73.56
get	Inf	praten	*talk*	32.80
happen	Inf	gebeuren	*happen*	32.02
say	168.39	kosten	*cost*	29.42
die	125.70	waaien	*storm*	26.26
cost	93.45	werken	*work*	25.56
put	91.57	samenwerken	*collaborate*	21.81
ask	59.91	zitten	*sit*	18.83
go	58.13	onderzoeken	*analyze*	18.52
marry	52.95	schijnen	*shine*	17.99

FIGURE 20

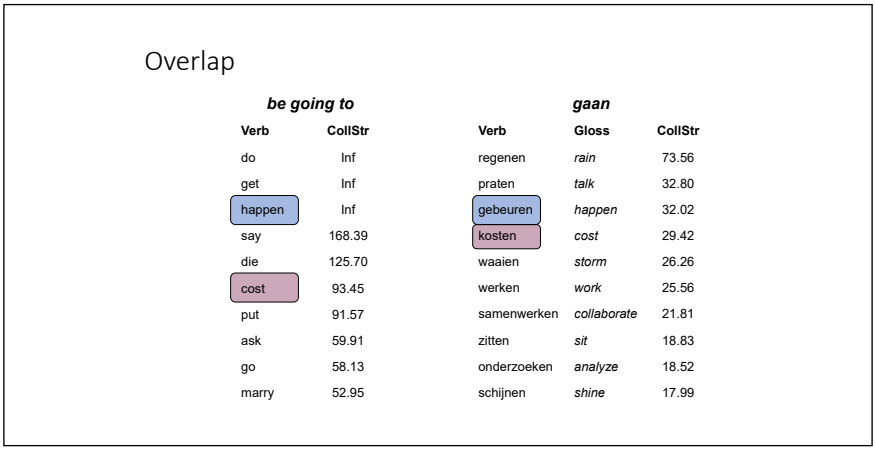


FIGURE 21

Let me show you the synchronic results first. On this slide, you see two lists with the synchronic collexemes of *be going to* and *gaan*. The lists show the most attracted lexical verbs for both *be going to* and *gaan*.

Among the top collexemes, there are exactly two verbs that match. With *be going to*, we have *happen* and *cost*. With *gaan*, we have two Dutch verbs that also mean "happen" and "cost", but that is where the similarities end.

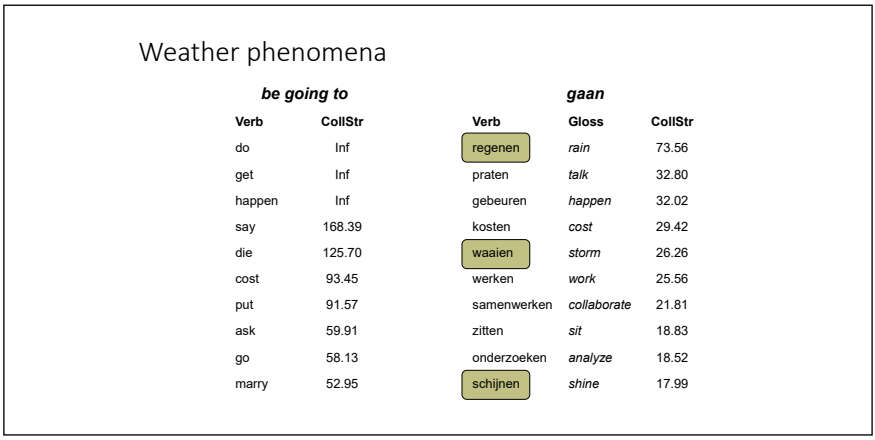


FIGURE 22

It is instructive to look at the different verb types that we find across the two constructions. In Dutch, there are quite a few weather phenomena, such as *rain*, *storm*, and *shine*. That is not the typical kind of verb that you expect to see with a movement-based future. Rain is not an intentional agent, it is a natural phenomenon.

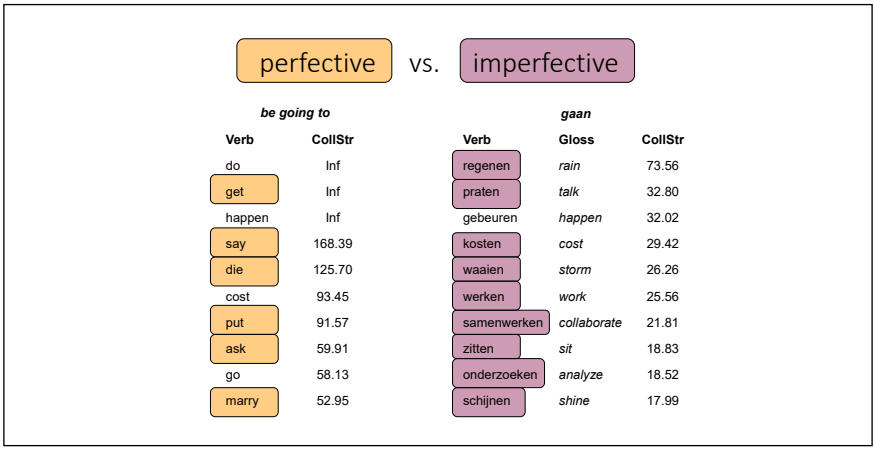


FIGURE 23

What struck me more though is that there is a great asymmetry with regard to the lexical aspect of verbs that we find in either construction. *Be going to* has a strong preference for verbs that I describe here as perfective. They encode a particular start point or end point, or they are even just happening at one singular point in time. Consider a verb such as *get*, which describes a punctual event. One moment you do not have it, then you get it, and then you have it. The same is true for the verb *say*. The verb *die* is perhaps the most drastic of them all. *The verb marry* also encodes a clear, punctual division of before and after.

With *gaan*, the most strongly attracted elements are imperfective. *Rain* can go on for a long time. Verbs such as *talk*, *storm*, *work*, *collaborate*, *sit*, or *analyze* encode activities that I can carry on and continue for an unlimited amount of time, or for as long as I choose to do so. That is one considerable difference.

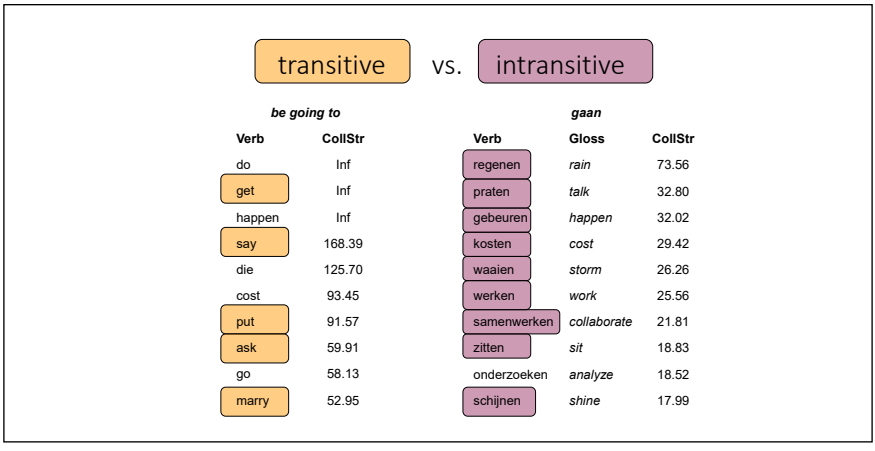


FIGURE 24

There is another difference in terms of transitivity. *Be going to* has a strong preference for transitive predicates. *I get something*, *I say something*, *I put something somewhere*, *I ask someone*, *I marry someone*. These are transitive verbs. The most attractive elements in the Dutch construction are intransitive, i.e. *rain*, *talk*, *happen*, *cost*, *storm*, *work* and so on and so forth.

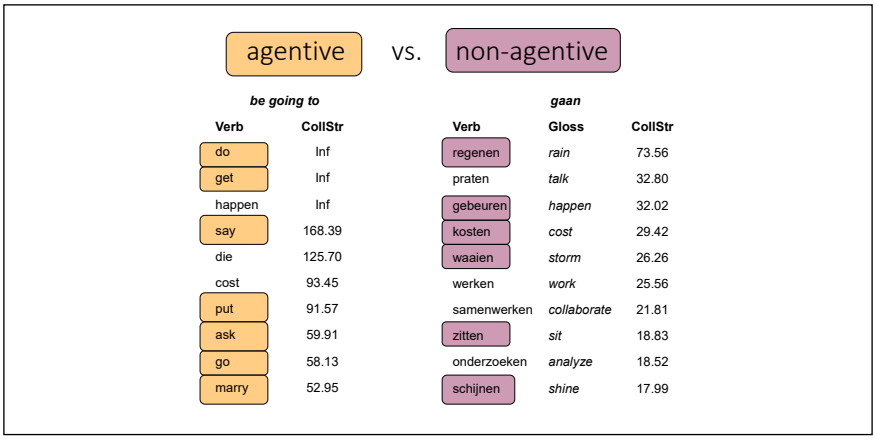


FIGURE 25

Another parameter that is strikingly asymmetric is agentivity: *I do something actively*, *I get something*, *I say something*, *I put something somewhere*. Things like *rain* or *happen* or *sit* are not agentive in the same way. There is no patient argument that would be affected by these activities.

To summarize, when we compare the collexemes, the main differences in synchronic usage are concerned with lexical aspect, transitivity and agentivity. These are aspects of transitivity that have been described famously in a paper by Hopper and Thompson (1980). I concluded that these grammatical differences really motivate the low rate of mutual translations. Grammatically, the two constructions both refer to future time events, but that is where the similarity ends. With regard to what kinds of events are encoded, the two constructions are almost diametrically opposed to each other, which brings me to the historical development. How did all this come to be? Let's look at the distinctive collexemes historically of English *be going to*.

This table shows you the overall development across the three periods with lists of the most attracted verbs for each historical time slice. Let me go through each period on its own.

We start with data in the 1700s, so 1710 to 1780. Here the most attracted collexemes encode intentional activities, and movement is often still a possible interpretation. When I am saying something like "*You're going to fight for your country*", you're not going to do it in your living room. You're going somewhere

English *be going to*

1710-1780		1780-1850		1850-1920	
Verb	**CollStr**	**Verb**	**CollStr**	**Verb**	**CollStr**
fight	3.48	hunt	2.31	be	10.66
publish	2.32	speak	2.21	do	4.81
answer	1.94	commence	1.79	get	4.18
observe	1.94	expose	1.79	have	3.19
embrace	1.92	part	1.79	try	2.94
ravish	1.92	strike	1.79	die	2.30
relate	1.69			stay	1.76
begin	1.56			happen	1.74
visit	1.48			run	1.57
				talk	1.37

FIGURE 26

English *be going to*

1710-1780	
Verb	**CollStr**
fight	3.48
publish	2.32
answer	1.94
observe	1.94
embrace	1.92
ravish	1.92
relate	1.69
begin	1.56
visit	1.48

- Intentional activities, movement still a possible interpretation

  And now my boy, I cried, you are going to fight for your country.
  I am going to visit the Marquis, and will talk further with thee at my return.

- Metalinguistic verbs

  By the circumstances of the story which I am going to relate, you will be convinced of my candour.

  As he was going to begin his narrative, Rasselas was called to a concert.

FIGURE 27

else to do the fighting. In "*You're going to visit someone*", it is implied that this visiting will take place somewhere else. Again, there are verbs that encode speech acts. There are a number of metalinguistic verbs that figure in this early period. Metalinguistic verbs feature in examples such as *the story which I am going to relate* or *as he was going to begin his narrative*. These verbs are attracted to *be going to* during this early period.

In the second period, all distinctive collexemes are compatible with intentional actions. Speech act verbs still continue to be represented. But we find the first events that are independent of human actions, as for instance the distinctive collexeme *strike*. While *strike* could be viewed as an agentive, intentional verbs that requires a human agent who is carrying out an action, the example on this slide shows a different meaning. In the example "*When ten*

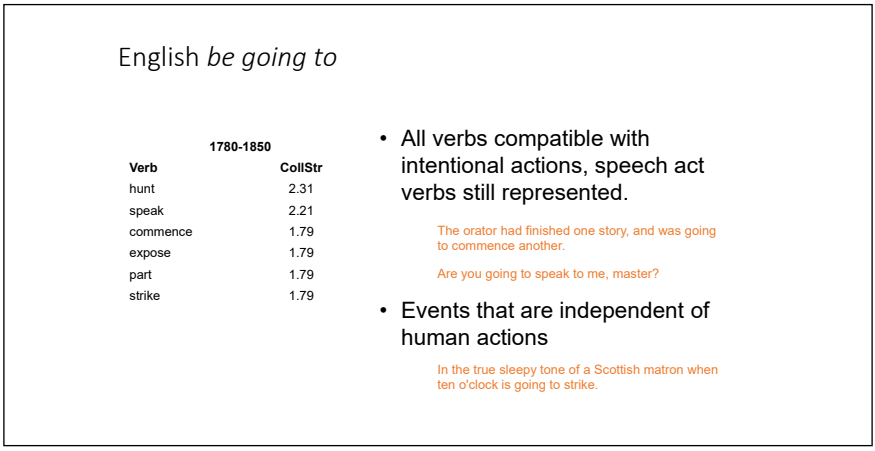


FIGURE 28

o'clock is going to strike", it is a clock doing the striking, not a human being. This opens up a pathway to other inanimate entities accomplishing actions, and eventually the construction broadens semantically.

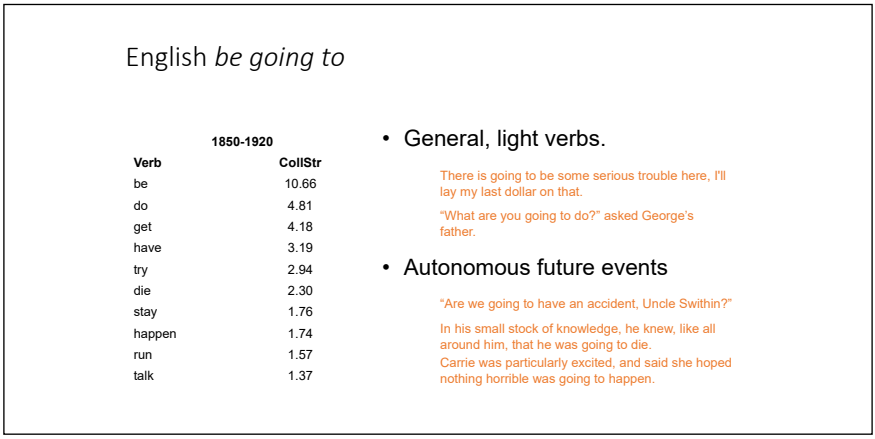


FIGURE 29

Moving on to the third period, here we find again the present-day profile of what *be going to* is like. There are general light verbs such as *be*, *do*, *get* and *have*, and these encode at least to some extent autonomous future events that haven't been planned and that haven't been executed by evolutional agents. Examples like "*Are we going to have an accident*" appear in the form of a question, "*He knew that he was going to die*" or "*She hoped nothing horrible was going to happen*" encode spontaneous or hypothetical events. Examples of this kind are over-represented in this last period. In the shifting patterns of collocations, we can thus see a development towards more abstract meanings.

Also with *be going to*, the intentional source of future meaning seems very solid as a hypothesis. Speech act verbs as prototypical intentional verbs are central. The development towards future meaning is accompanied by a growing preference for general light verbs. All in all, this corroborates existing accounts of the grammaticalization of *be going to*.

Let's look at the diachrony of Dutch *gaan*. This slide shows an overview, but we are going to look at each period in turn.

Dutch *gaan*

16th – 17th century			18th – 19th century			20th century		
Verb	**Gloss**	**CS**	**Verb**	**Gloss**	**CS**	**Verb**	**Gloss**	**CS**
lopen	*walk*	3.52	opzoeken	*find*	2.67	beminnen	*love*	3.44
strijken	*run off*	3.30	onderzoeken	*analyze*	2.35	denken	*think*	3.44
stellen	*put*	3.25	vertellen	*tell*	2.25	gebeuren	*happen*	2.45
reizen	*travel*	2.86	doorbrengen	*spend*	2.01	bevrijden	*liberate*	1.96
liggen	*lie*	2.46	geven	*give*	1.68	studeren	*study*	1.96
preken	*preach*	2.00	varen	*travel*	1.68	voelen	*feel*	1.96
treden	*step*	2.00	verkoopen	*sell*	1.68	eten	*eat*	1.66
verhuizen	*move*	2.00	leggen	*put*	1.50	werken	*work*	1.49
leiden	*lead*	1.48	sterven	*die*	1.49	krijgen	*get*	1.47
rechten	*straighten*	1.48	brengen	*bring*	1.49	twijfelen	*doubt*	1.47
spreken	*speak*	1.41	halen	*get*	1.41	kijken	*look*	1.32
drinken	*drink*	1.36	roepen	*call*	1.39			
bezigen	*use*	1.33	nemen	*take*	1.37			

FIGURE 30

Gaan unsurprisingly starts with predicates that are movement verbs such as *walk*, *run off*, *travel*, *step* and *move*, which are among the most strongly attracted elements in the first period. There are caused posture verbs like *put* or *straighten*. Most verbs have a imperfective aspectual profile, as for example *walk* or *travel*, which encode actions that you can continue for a long time. These verbs further encode intentional human actions.

Dutch *gaan*

16th – 17th century		
Verb	**Gloss**	**CS**
lopen	*walk*	3.52
strijken	*run off*	3.30
stellen	*put*	3.25
reizen	*travel*	2.86
liggen	*lie*	2.46
preken	*preach*	2.00
treden	*step*	2.00
verhuizen	*move*	2.00
leiden	*lead*	1.48
rechten	*straighten*	1.48
spreken	*speak*	1.41
drinken	*drink*	1.36
bezigen	*use*	1.33

- Movement verbs

  Nu willic gaan loopen al in mijn huus.
  now want.I go walk all in my house
  'Now I want to go home.'

  Daer gaet hij strijcken!
  there goes he run.off
  'There he's running off!'

- Caused posture verbs
- Most verbs are imperfective
- Intentional human actions

FIGURE 31

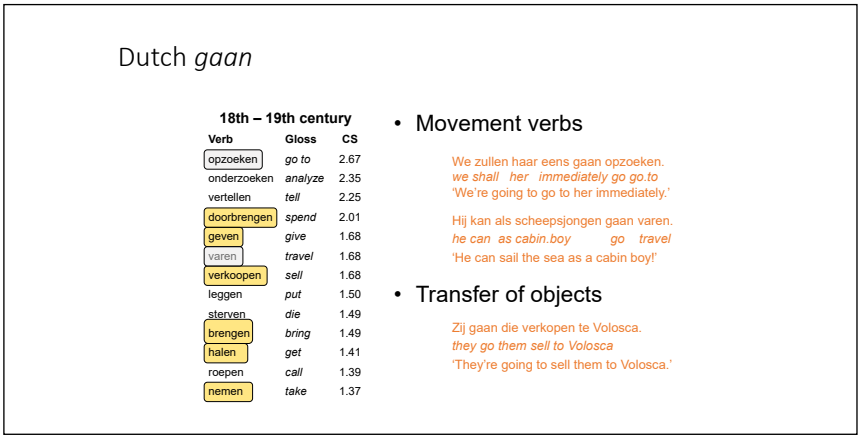


FIGURE 32

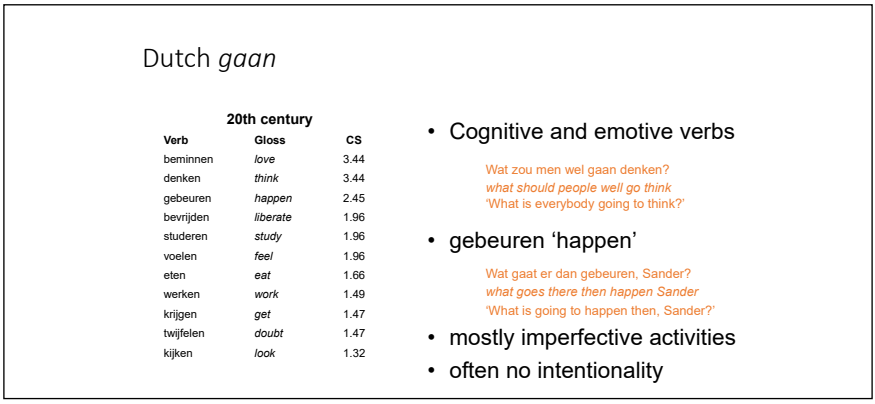


FIGURE 33

In the second period, movement verbs are still strongly represented. There is a verb that means "go to" and there is a verb that means "travel". Other distinctive collexemes encode transfers, as for example *spend*, or *give*, or *sell*, or *bring*. Suddenly these show up as a relatively homogeneous class. They are different in terms of their aspectual profile, as they have an endpoint. If you spend your money, then it is gone.

In the third period, the distinctive collexemes include cognitive and emotive verbs. A verb that means *happen* is another strongly attracted element. Most of the distinctive elements for this period are imperfective activities, typically with no intentionality at all. A verb like *love* encodes a human activity, but I can't intentionally decide to love someone, and when I love someone that is a state that has a temporal extension. That is very different from verbs such as

break or *cut*. The list of verbs on this slide includes *love*, *think*, *feel*, and *doubt*, which form the profile of this movement-based future construction.

In summary, the early usages of *gaan* commonly involve literal, intentional motion. Later, movement verbs are joined by verbs of transfer. All along the constructional meaning broadens, accommodating verbs without the meaning of intentionality and then in present-day usage, *gaan* preferentially occurs with atelic predicates, and intention is no longer a part of the constructional semantics. In synchronic usage, we find weather verbs as attracted collexemes, and in the third historical period, these are cognitive response verbs such as *doubt*.

What can we conclude from all of this? English *be going to* and Dutch *gaan* both follow the general path of movement-based future constructions, which start with the idea of motion, then merge into the idea of intention, and finally settle into future time reference. Even though they are moving along the same path that is typologically well-attested, this does not mean that they function similarly in language use. They have converse preferences for perfectivity, transitivity, and agentivity. If we are looking at the collocational patterns in their shifts, we see substantial developmental differences. *Be going to* has a preference for speech act verbs in the second period. *Gaan* starts out with movement verbs that are imperfective like *travel*.

I have talked enough about future constructions for one day. To sum up, I hope you see what I find attractive about the concept of looking at shifting patterns of associations. If you come to it for the first time you might think that this gives you a very diffuse idea of how language changes. Wouldn't it be much clearer to look at morphosyntactic changes or at first instances of a construction with a tangible meaning difference? I think that the study of collocational shifts can yield important insights.

I would submit that shifts in collocational preferences constitute one central type of change in the network of constructions. I have mentioned host-class expansion as one important concept in this context. A construction increases its number of links either to lexical collocates or to syntactic carrier phrases, constructions. I have argued that shifting collocational preferences can actually reflect systematic patterns of semantic change with a sufficient amount of accuracy, so that we can test hypotheses or test claims that have been proposed about semantic change elsewhere. Collocational preferences change as a construction develops along a grammaticalization path. Different paths then are embodied or realized by collocational changes that differ with respect to each other as well. With this, I would like to come to a close for today. Thank you for your attention.

LECTURE 5

How Constructional Networks Grow and Fade

Welcome back to this fifth lecture of Ten Lectures on Diachronic Construction Grammar. In the last lecture, I have started to discuss a number of ways in which constructionalization and constructional change can be studied empirically on the basis of diachronic corpus data. In this lecture, I will continue with that general theme. The title for this lecture is "How constructional networks grow and fade". The studies that I presented in the last lecture elaborated on the idea that constructional meaning is reflected in patterns of associations between syntactic structures and lexical elements. I explored the idea of a constructional network, in which grammatical constructions entertain many connections with lexical elements that occur with these constructions. A construction like the English auxiliary *will*, for example, has a slot for a verb in the infinitive. That slot is connected via hundreds and hundreds of associative links to different lexical verbs that can fill that slot. Importantly, these links have different strengths. Associations differ in how strong they are, also the strengths of these connections can change over time. You can examine how a given construction is connected to lexical elements in the 1800s and how that changes over the years, so that you have a very different situation in the 2000s. Collostructional analysis, the method that I have used there, allows you to investigate that.

The overall phenomenon that I have addressed is connectivity change. This is what you see displayed schematically on the slide. The examples that I have discussed were future constructions in Danish, Dutch and English that illustrate the general point. This idea of connectivity change and how you can study it was the general theme of my 2008 book. If you would like to go deeper into that issue, you can turn to the book, which presents a range of similar studies that explore that idea from different angles. In this lecture I want to put the focus on another of the ten basic ideas of Construction Grammar that I presented to you earlier. Namely, I want to focus on the idea that constructions vary in terms

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691107.

of their degrees of complexity and schematicity, which has profound implications for the way we think about constructional networks. I will talk about a perspective on changing constructional networks that I developed in my 2013 book on constructional change. Constructional change, if you remember the definition from Lecture 2, selectively seizes a conventionalized form-meaning pair. It changes a construction in terms of its form, its frequency, aspects of meaning, and its distribution in the community who uses it and any combination of these factors. The networks of changing collocational associations that I described in the last lecture mainly focused on two aspects of this definition, namely changes in frequency. We saw that different collocates become more frequent or less frequent, and this is picked up by the collostructional analysis. The second aspect is that these future constructions also change in function. As a future construction moves along its grammaticalization path, it becomes broader in its meaning potential. Its function changes. Crucially, the constructions that I described yesterday did not undergo any formal morphosyntactic changes during the historical period that I investigated. I was really exploring what Traugott and Trousdale would call constructional change, rather than constructionalization.

Today I would like to present a study of constructional change that does involve formal change alongside functional change and frequency change. I will also focus on a different grammatical domain. Yesterday, we were firmly in the domain of verbal grammar, studying the behavior of verbs and auxiliaries. Today, we turn to the nominal domain. I will be talking about English nouns, and specifically about morphology and word formation, discussing how new words enter the language.

The case study that I brought along for this morning concerns an English derivational suffix that has had an interesting life and death. The suffix by itself is *-ment*, which is a nominalizing suffix that you know from English words such as *punishment*, *treatment*, *settlement*, and many others. I became interested in that particular suffix, because I noticed something strange about it. In Present-day English, you find around 1000 words ending in the suffix *-ment*. Dictionaries list many different words of this kind. If we investigate *-ment* on the basis of corpus data, we find that many of these noun types are relatively infrequent or occur only once. A word that occurs only once in a corpus is what corpus linguists call a "hapax legomenon". Morphological patterns with lots of unique types, lots of hapax legomena, are typically very productive, because the large number of low-frequency types shows that speakers use that suffix to create new words spontaneously. That is normally the case. But with *-ment*, even though there are lots of different types and lots of infrequent types, speakers cannot produce new words on the basis of *-ment*. Let's say that I have been

emailing back and forth with a colleague this morning. I could describe this as a long *emailment*. That, however, is not a possible word of English. People would understand it if I say it. But they would also understand that I am making a language-based joke. I would be doing something that is not within the limits of what I can conventionally do with the suffix-*ment*. The same would be true of *recyclement*. I can collect my glass bottles and my old newspapers and then show them to a friend and when I say, "These here are my recyclements." My friend would understand me, but it would not make *recyclement* a usable word. I was wondering, how can I explain this discrepancy? What is going on with the suffix *-ment*?

In the time that I have this morning I want to focus on four questions. First of all, what is the *V-ment* construction? I call it the *V-ment* construction, because the stems that you find in the construction are typically verbal. A form like *punishment* has the verb *punish* and *-ment* makes a noun out of it. The word *treatment* starts with the verb *treat*, and *-ment* makes a noun out of it. How can we describe that construction?

Second, how did the construction change in productivity? It must have been productive at some point. At some point, speakers of English produced new types on the basis of that suffix, but somehow that stopped.

I also want to explore how this construction changed over time in form and function. What are the different meanings that can be described with it? What are the different forms that we find in it? For example, even though I just said that the construction normally occurs with verbs, as in *treatment* or *punishment*, there are some forms that are exceptions to that. If you think of something like *basement*, we have *base* plus *-ment*. *Base* is arguably not a verb in this context, although it exists as a verb elsewhere.

Lastly I want to bring it all back to Construction Grammar and the question of constructionalization and constructional change. How do these findings about productivity and change in form and function speak to the issue of constructional change?

Let's start with a general characterization of the construction. I have already mentioned that it is a combination of a lexical stem with a suffix that has a certain phonemic shape. The suffix is pronounced as *-ment*, and the stem strongly tends to be verbal. I have mentioned the exception of *pavement* already. There are others, for instance, *merriment*, which is an infrequent, somewhat archaic word that describes joyful activities. Meaning-wise, the construction typically conveys the meaning of an action. An *adjustment* is the action of adjusting a projector that is sitting a little bit askew. It can also be the result of an action. When I buy an *assortment* of sweets, an assortment is not the action of putting the sweets in the box, but rather it is the box that I buy and I can take

home. The construction can also express the means to accomplish an action. A *refreshment* is not the action of refreshing myself by taking a sip of water, but rather it is the water that I can drink or small snacks that I can eat.

Let us move on to what happened to *-ment* historically. This slide shows a very short summary of its history. It starts in Old English, where we can already see a few words that end in *-ment*, and those were loan words from Latin. Old English had borrowed and incorporated words from Latin that were part of ordinary language use. But those were isolated forms. The construction as a generalization came about as a contact phenomenon, due to an important event in the history of the English language, namely, the Norman Conquest of 1066. The *-ment* suffix is of Romance origin, and that is how it entered into English. The construction then became nativized. We see that in forms with *-ment* that do not have French verbal stems. Take a word with a Germanic stem like *shipment* for instance. *Ship* is a Germanic word, and we find it combined with *-ment*. We find the Romance construction with Germanic parts in it. That is how we can know that this construction was actually incorporated and nativized into English. That happened between 1250 and 1350.

Shortly after that, the construction already receded and became less productive and fewer and fewer loan words were borrowed into English. That already marks the downfall of the construction. The overall productivity receded, and what we have in Present-day English is what you could describe as a residue, leftovers of history. About thousand types remain in the language, but the construction is not productive. I have given you the examples of *emailment* and *recyclement* earlier. Today I have been going for a run that I cannot describe as a *jogment*, or if I am kissing someone repeatedly, that cannot be described as a *kissment*.

How did the construction change in productivity? How can that be analyzed? I decided to use a database for the study that I talked about yesterday in the context of Michael Israel and the *way*-construction. Michael Israel used the Oxford English Dictionary for his study, and I did the same for my investigation into *-ment*. The Oxford English Dictionary is not a diachronic corpus. At least it is a very special type of diachronic corpus. It is a dictionary that has words and definitions, but above and beyond that, it has authentic examples of texts. It has quotations. These quotations have historical dates.

Let me show you a screenshot of what the electronic version of this dictionary looks like. This slide shows the entry for the word *achievement*. You see that there is a section on the etymology of *achievement*. It comes from Anglo-Norman and Middle French. It means the action of finishing or completing something. We are given information about the general time during which that word was borrowed. Below that are authentic examples that the compilers

FIGURE 1

of the Oxford English Dictionary have collected. That allowed me to take the first date of attestation as a proxy for the time when this word would have entered the English language. Much like Michael Israel was collecting verbs with the *way construction*, I have been collecting *-ment* words from the Oxford English Dictionary, taking notes as to when these words were first attested and when they first entered the language. That allowed me to track over time how many new words with *-ment* entered the language at any given point in time.

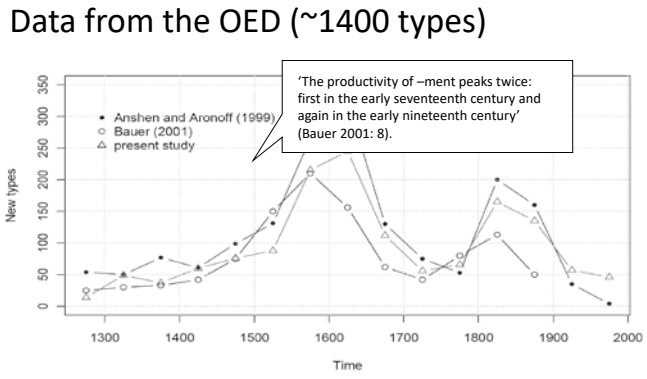


FIGURE 2

Proceeding in this way, I found about 1400 different words ending in *-ment* that were recorded in the dictionary. I did this and as soon as I was finished, I realized that there had been two other studies that had been doing the exact same thing (Anshen and Aronoff 1999, Bauer 2001). I was very nervous to see how their data would compare to mine. You can see in this graph on the slide three frequency curves that show how many new types with *-ment* enter the English language during every half century. What you see is that the three curves are in broad agreement. At least they are not very far away from each other. There are a few discrepancies here and there. For instance, Anshen and Aronoff (1999) find more types around the year 1600. In the very last period, I found more types. That is because the Oxford English Dictionary is continuously updated with new entries. Between 1999 and my study in 2013, there were new types that were added, and that accounts for this difference.

When you look at this curve, you notice that it looks like the back of a camel. Bauer (2001) as argued on the basis of this curve that the productivity of *-ment* has two peaks. It has two periods at which it is a very productive, first in the early 17th century, and then again in the early 19th century. Now, that statement is problematic because the texts that were compiled by the editors of the Oxford English Dictionary are not of the same size across all historical periods.

Quite to the contrary, later periods are represented with lots and lots more text. That means that at the end of the period we are much more likely to find lots of words that end in *-ment*. What I decided to do as a first analytical step was to normalize the type frequencies. By calculating the number of types per 10,000 words I tried to control for the different amounts of text that we find for each historical period.

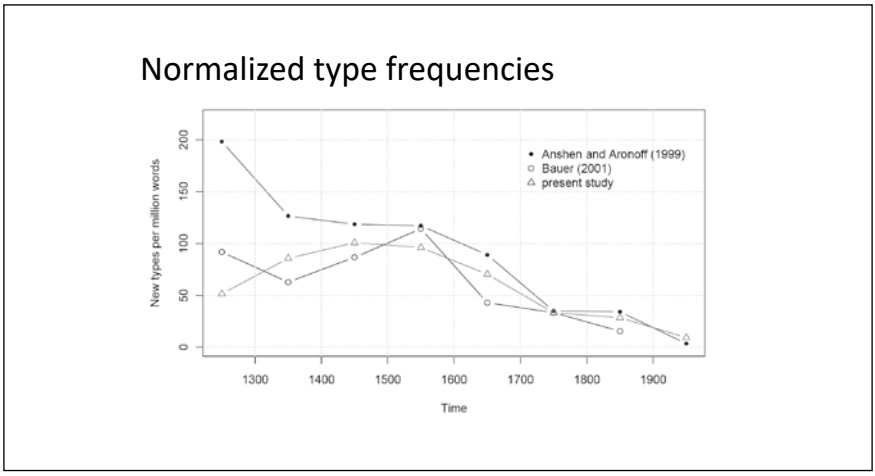


FIGURE 3

Once I did that, the camel shape disappeared. What we see here are the normalized type frequencies. In the beginning, there are some discrepancies, about up to 1550. But after that, all the three curves seem to agree that there is a linear, gradual decrease of new types per time period. That means that from the middle of the 16th century, this construction has been decreasing in productivity more or less steadily. Normalized type frequencies are one way of approximating the concept of productivity, how easily new types are formed, but it is not a very precise measure. It is not a measure that is typically used in current corpus linguistic studies of productivity. Typically, corpus linguistic studies assess productivity through measures that take into account the prevalence of hapax legomena.

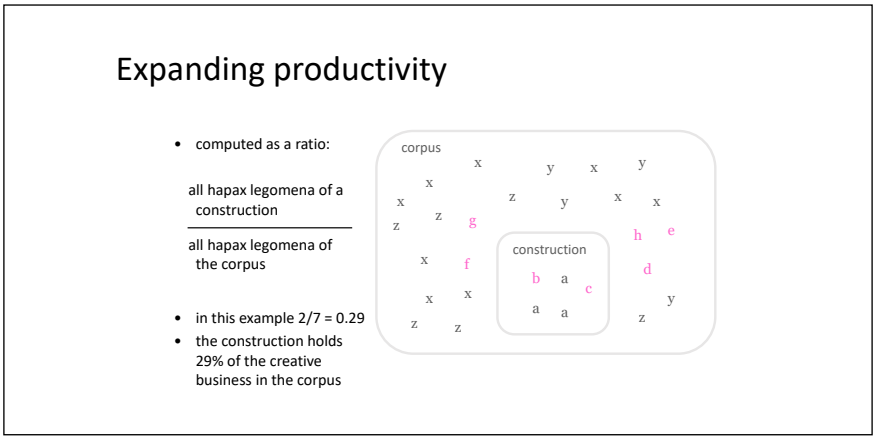


FIGURE 4

I decided to apply a corpus linguistic measure that is labeled expanding productivity. This measure of productivity is calculated as a ratio of hapax legomena in a construction and hapax legomena in the corpus. You see a visualization of the general logic of this measure on this slide. Again, we have a corpus with lots of different words in it. The words are represented by letters like *x*, *y* and *z*. You also see an *f* and a *g*.

I am interested in the productivity of this construction here. The first analytical step is to extract all the types of that construction. In this toy example, we only have three different types, *a*, *b* and *c*. Two of them occur only once. There is only one *b*, there is only one *c*, so they are counted as hapax legomena. The three instances of *a* occur more than once, so they are not counted.

The number of all hapax legomena in the construction has to be divided by all hapax legomena that I find in the entire corpus. In the entire corpus, there are lots of *x*s. They do not count. There are lots of *y*s and *z*s that do not count

either. By contrast, there is only one *f*, only one *g*, only one *h*, one *e*, one *d*, one *b* and one *c*.

Following this logic, we would divide 2 by 7. There are 2 construction hapaxes, and there are 7 corpus hapaxes. That would give us the ratio of all hapaxes that are accounted for by the construction. It gives us the part of the creative language use that the construction represents in the corpus. We could say that the construction accounts for 29% of the creative language use in the corpus. I should maybe say why hapaxes are taken as a representation of linguistic creativity. When I am doing something creative with language, when I am making up a new word, then that word will at first be only very infrequent. The first time I say a new word, it occurs only once. Speakers are creative more or less all the time. They bring new words into their languages all the time. At first, these words are very infrequent. Corpus linguists look to infrequent word types to make inferences about how speakers are creative with their languages.

I calculated this measure of expanding productivity for all time slices in the OED (Oxford English Dictionary) that I had. For each time slice, I counted the *-ment* hapaxes and the overall number of hapaxes in the OED quotations for that time slice. Of course, not all hapaxes are words that are created in that very moment. Some words are just very infrequent, but for those words that are creative, we can be relatively sure that at first they are likely to be used with low frequency.

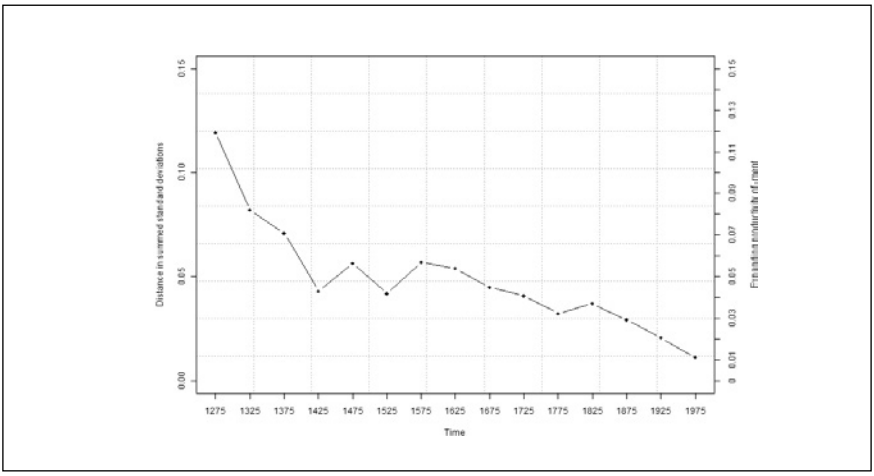


FIGURE 5

This calculation of expanding productivity over time gives me the descending curve that you see on this slide, which looks very similar to the curve of normalized type frequencies that I presented earlier. This suggests that this almost

uniform downward curve represents the decline of productivity that the *-ment* construction undergoes.

To sum this up, how did the *V-ment* construction change in productivity? My conclusion was that there is really just one peak in productivity, and that peak occurs relatively early during the 13th century, which marks the moment of nativization. This was borrowed into the English language and it was nativized. It was popular for a while, but then its popularity faded quickly. I tried to find an non-linguistic analogy for this. You probably know the Rubik's cube, which is a plastic cube with multi-colored squares that you can twist and turn in order to make each side appear in just one color. It was a very popular children's toy at some point in time, everybody needed to have one. The analogy actually goes further than a toy being highly popular at one point and then falling out of fashion. If you go through a box of old toys in your basement, you might still find a Rubik's cube. In similar ways, you still find all the old *-ment* nouns in the English language. They're still there. We still use them. We just do not make new ones.

Moving on to the next question, how did the *V-ment* construction change in form and function across its lifespan?

In order to analyze this, I decided that I wanted to divide the history of the construction into stages, so that I could find out what happens early on and compare that what happens in the middle of the development and at the end. I thus needed a way to determine how I wanted to divide the overall development into stages. In historical linguistics, this is a general question. How do we divide a stretch of time into periods? Do we take one century, then the next century, then the next century? Do we take decades, or even individual years? Or do we go with periods that are motivated on the basis of language-external events? If there are important historical events that affect the language and the culture, that might lead us to distinguish time periods based on those events.

Stefan Gries and I had been developing a method that we called variability-based neighbor clustering or VNC. The general idea of that technique is to arrive at a periodization of historical data that is inductive and data-driven. If the data does not change significantly during one particular time, that time should be recognized as a coherent period. If the data changes suddenly, then we have a reason to posit a boundary between historical periods. Variability-based neighbor clustering is a clustering algorithm that adopts the general principles of other hierarchical clustering methods. Unlike other clustering algorithms however, there is one twist in neighbor clustering, namely that it can only merge temporally adjacent data points. I will explain what this means in just a second.

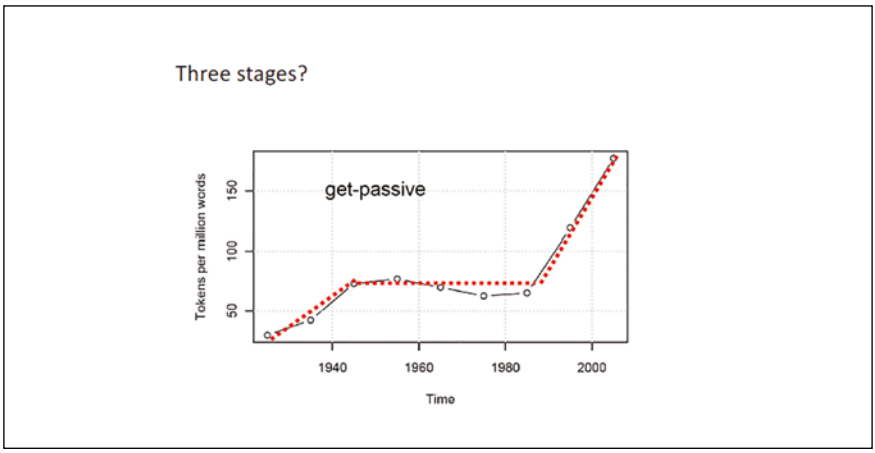


FIGURE 6

In order to explain it, I need to leave the *V-ment* construction for a moment. Let us look at one or two simpler examples. Here we see a frequency curve for the English *get*-passive construction which shows a diachronic increase. This is a data from The TIME Corpus of American English. You see that it starts low, then there is an increase, then there is something of a plateau. After 1980, the construction strongly increases in frequency. How do we partition a frequency curve like this into stages? You could look at this curve, eyeball it, and divide it into an initial period of increase, a second period with a plateau, and a third period with a strong increase.

When we partition diachronic corpus data into periods, there are several potential pitfalls. Typically, diachronic corpus work divides data into sequential periods that are chosen arbitrarily, like centuries or half-centuries or decades. But problem with that is that linguistic change is not always smooth. It can move in fits and bumps, and there can be U-shaped curves. That means that when we measure a linguistic phenomenon over a certain period and take an average value, the results that we get may actually be misleading in some cases. Different time slices, different periodizations yield potentially different results. The ideal way to divide the corpus into time slices would be to take some aspect of the phenomenon that is studied and to have a data-driven way of periodizing the data. How can that problem be addressed? One way of finding structures in large bodies of data is hierarchical clustering.

I think that at least some of you are familiar with the general idea of the method and its applications across different scientific disciplines. Clustering is used, for example, in biology, where researchers use it to investigate similarities

in the DNA of different beetle species. The method allows us to get a sense of how groups of species form larger groups. Clustering is of course also applied in linguistics.

My former officemate, Benedikt Szmrecsanyi, has used clustering to find groups of English dialects on the basis of their morphosyntactic characteristics (Szmrecsanyi 2013). Benedikt didn't measure DNA sequences, but instead took a catalogue of more than 50 morphosyntactic features, and he measured how often these features occurred in corpus data that represents these dialects. Let me say a few words about how this works in principle.

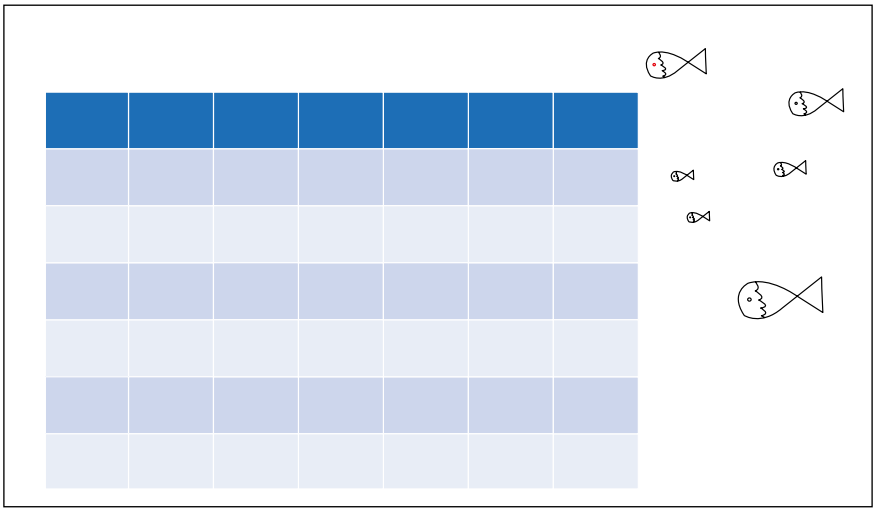
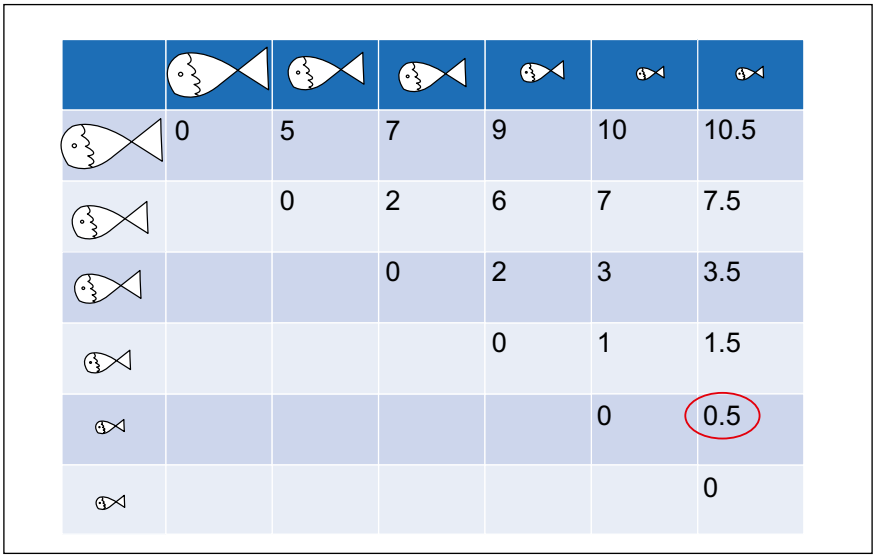


FIGURE 7

We'll take a very simplistic example, with six elements that differ in just one feature, namely physical size. There are six fish that differ in how large they are. How do we divide these six into groups? It would be possible to describe the set as one big fish, two medium-sized fish, and three small fish. It would also be possible to distinguish between three large ones and three small ones. Hierarchical clustering offers an inductive way of deciding how to describe the set. We arrange the fish in the matrix like this. Each fish is compared to each other fish.

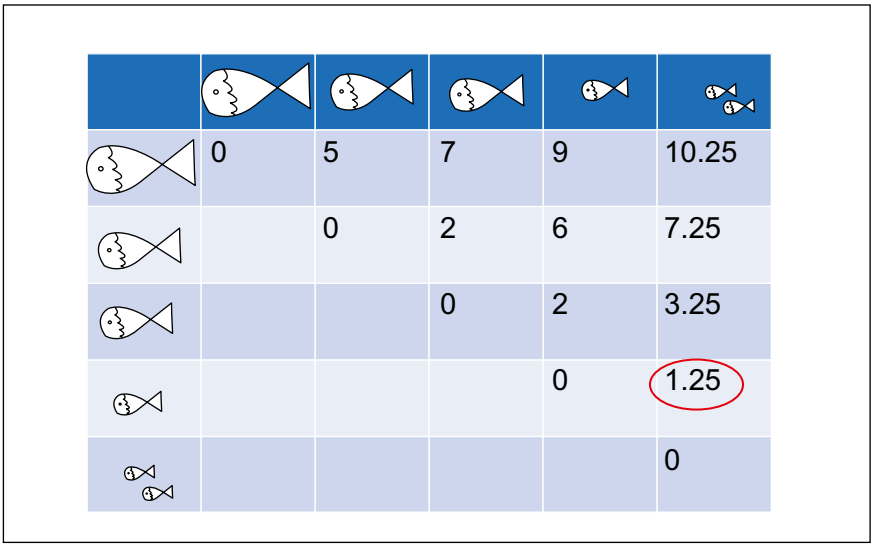
We determine the size difference between one fish and all of the others. Comparing the big fish to itself yields a difference of 0. A comparison with the next one yields a difference of 5. The next one shows a difference of 9, and so on and so forth. For every possible pairing, we determine a measurement of difference. The difference between the two smallest ones is 0.5, which is circled on this slide. When all difference measurements are taken into account, this is the smallest difference between all combinations in the entire set.



	0	5	7	9	10	10.5
		0	2	6	7	7.5
			0	2	3	3.5
				0	1	1.5
					0	0.5
						0

FIGURE 8

The starting point of a clustering algorithm is the establishment of differences on the basis of pairwise measurements of all the data points that you give it. The algorithm determines a measure of difference and looks for the smallest difference in the entire dataset. Once this pairing has been identified, once we have the two fish that are the least different, the algorithm puts these together and calculates an average value, so that we have a table that is a little bit smaller.



	0	5	7	9	10.25
		0	2	6	7.25
			0	2	3.25
				0	1.25
					0

FIGURE 9

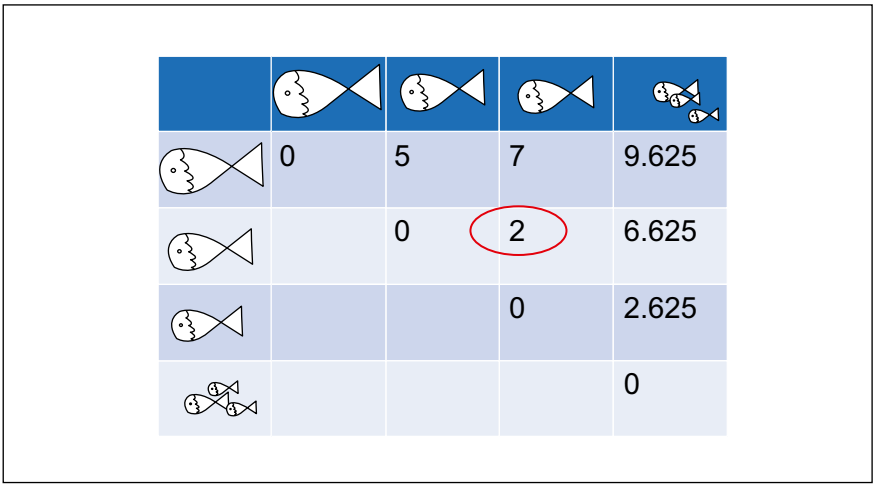


FIGURE 10

Now we have a matrix in which these two have been merged. They form a cluster. The average size is taken to compute a difference between the biggest fish and this cluster, which is 10.25. With the cluster of the two small fish, we now have a slightly smaller table in which some values are recalculated in the second round of the iterative clustering algorithm. The algorithm takes all these numbers and searches for the smallest difference. In this iteration, the smallest difference is between cluster with the two smallest fish and the smallest remaining single fish.

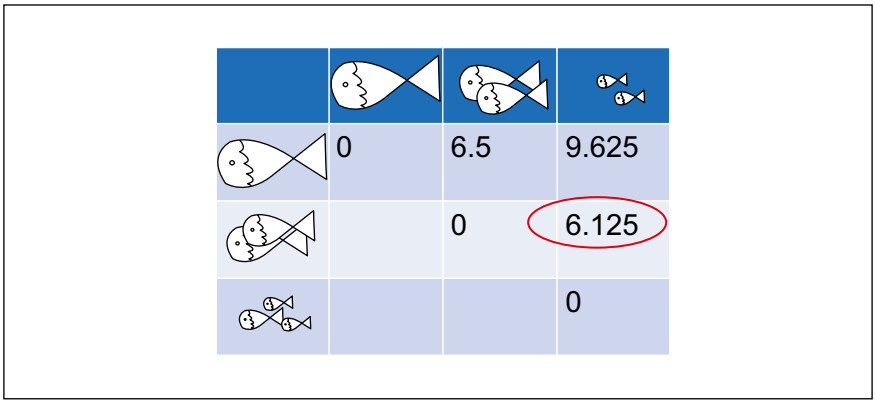


FIGURE 11

In the following iteration, we take the average of that and are left with a table that is even smaller. The algorithm creates the cluster with what is now three

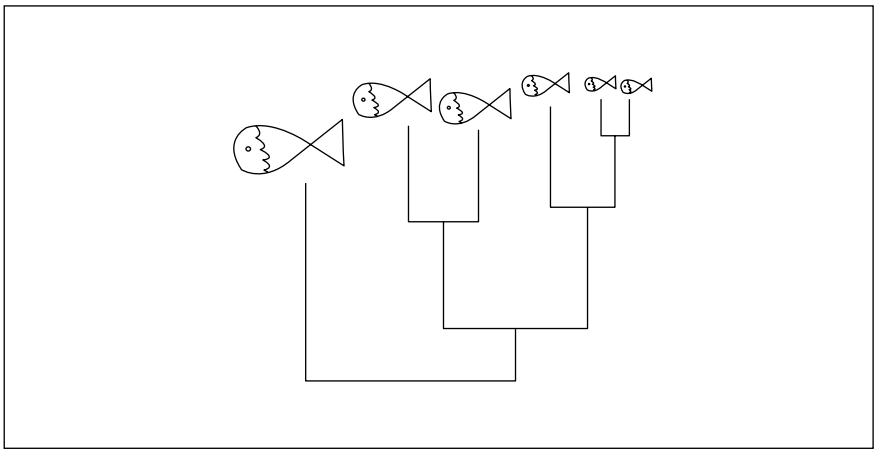


FIGURE 12

fish, recalculates the differences, finds the smallest one, and then the process repeats. This time, the smallest number that we find in this table is the difference between the two medium sized fish.
That means we put these two together and get the smaller table that looks like this. This actually gives you the answer to the question that I asked earlier. Our set contains one big fish, two medium-sized fish and three small fish. The algorithm does not stop there. You can go on with it and the next iteration would group all the smaller fish together.

Overall, we have a clustering dendrogram that looks like this. A useful way to describe this would be one big fish, two medium-sized, and three small ones. But if you wanted to have only two groups, you would differentiate between one a big fish and five small ones.

To return to language, the general idea is that we can use clustering to find how the development of a given linguistic unit can be divided into stages, so that instead of taking fish and measuring their size, we take data from different historical periods and annotate the data for a feature that we are interested in. We could take the frequency of an element. We could take the relative frequency of *dislike* with the *to*-infinitive and *dislike* with the *-ing* form, or we could take the range of collocates, like Michael Israel did with the *way*-construction. Anything that you can count, anything that you can measure and express in a number, you can study in this way. Then, you take those measurements and group them according to their similarity. General hierarchical clustering can be applied to diachronic data in this way. But there is one problem, namely, conventional hierarchical clustering algorithms do not know about the temporal sequence that you have in your data.

If, for instance, we have data that comprises data from 1993, 1994, 1995, 1996, and so on and so forth to 2000, it may happen that the data from 1993 is very similar to the data from the 2000. Conventional hierarchical clustering would group the two years together, meaning that you would end up with clusters that are nonsensical. How do we fix that? That is the problem that variability-based neighbor clustering addresses. The clustering algorithm measures the differences from one period to the next. We only merge those pairings that are temporally adjacent. We can only combine data points that sit next to each other in their temporal sequence.

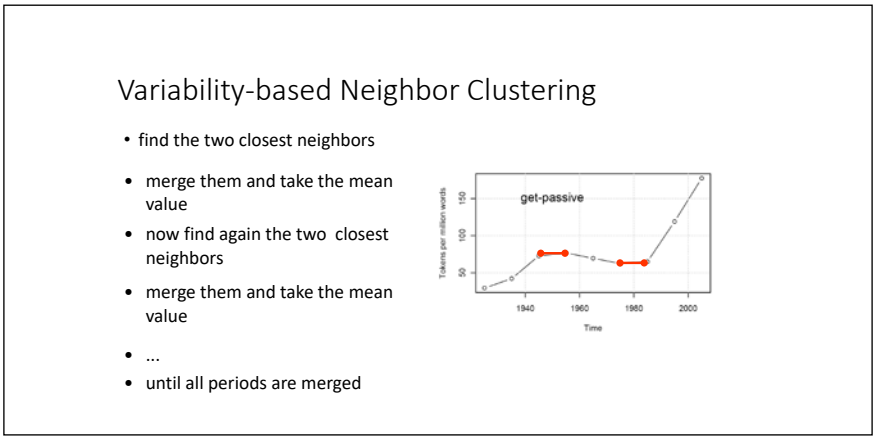
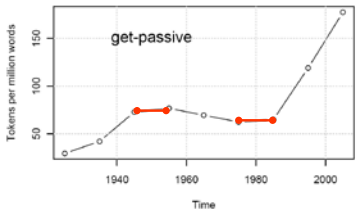


FIGURE 13

Let me come back to the *get*-passive and explain how this works in practice. The clustering algorithm goes through these data points. For each pairing, it calculates a measure of difference. It goes from the first to the second and measures how different they are. It goes from the second to the third and measures the difference. It goes from the third to the fourth and finds a very small difference. Fourth to the fifth, that is also pretty small. Fifth to the sixth, it is kind of small. Sixth to the seventh, this one is really small. Then the last two ones are quite large again. It finds the closest neighbors, which in this case are the sixth and the seven point here. It merges them. It takes the average value, just as with the fish. Then it goes through the entire set again and finds the two closest neighbors. In this case, on the second iteration of the algorithm, the two closest neighbors are the third and the fourth point. It merges them, it takes the mean value, and it does that again and again until all periods are merged together. Clusters of clusters of clusters until you have everything in one tree diagram.



Decade	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
Tokens per MW	29.5	42.2	72.8	76.7	69.6	62.4	64.9	119.0	177.3

Decade	1920s	1930s	1940s	1950s	1960s	1970s and 1980s	1990s	2000s
Tokens per MW	29.5	42.2	72.8	76.7	69.6	63.65	119.0	177.3

Decade	1920s	1930s	1940s and 1950s	1960s	1970s and 1980s	1990s	2000s
Tokens per MW	29.5	42.2	74.75	69.6	63.65	119.0	177.3

FIGURE 14

Just to go over this with concrete numbers once, we start with all data points, and we find the adjacent pair that is the closest match, which in this case are the 1970s and 1980s. Those two are then merged. We take the average. This process repeats so that the next time around, the closest neighbors are the 1940s and 1950s. We take the average and the algorithm goes through the data again and again, until all of these cells are merged into one. We have the overall average, which then gives you a tree diagram that reflects the overall development.

In this case, we learn that in order to divide the development into two parts, we would need to distinguish the last one from everything else. If we divide the development into three parts, distinguish the last one, then the penultimate period, and then all the rest. For four parts, we would have the first two periods,

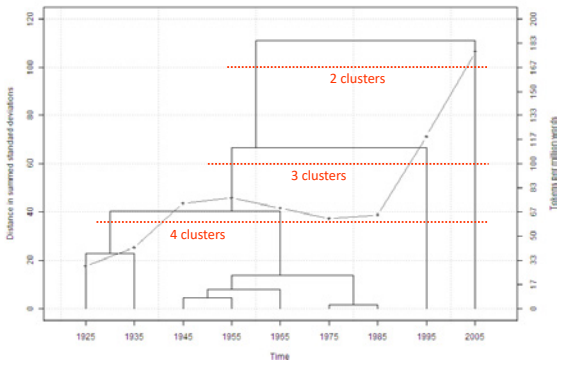


FIGURE 15

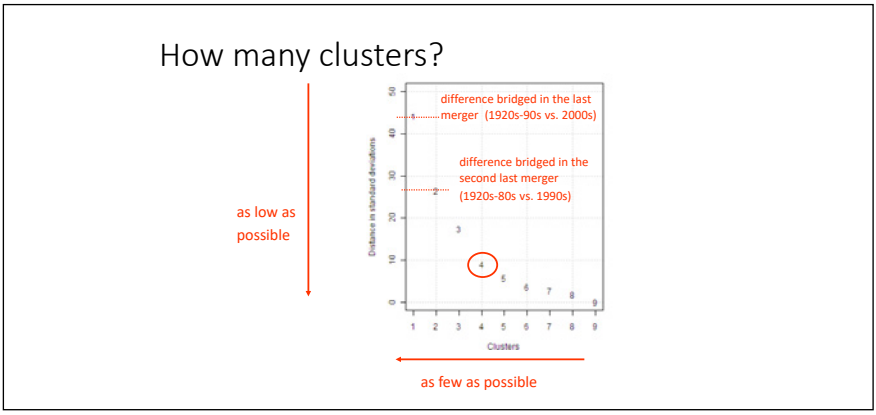


FIGURE 16

then a cluster of five periods that represents the plateau in between, and then the penultimate data point on its own and the last data point on its own. That is how it works. The question is, how many clusters should we assume?

There is no exact science to answer that question. But the difference between each merged cluster gives us a cue. The graph you see on this slide is based on the results of the clustering algorithm, it is called a scree plot. The graph shows us how great the differences are that are bridged with each merger in the clustering algorithm. The distance that is bridged in the last merger is a lot larger than the difference that is bridged in the second last merger, and so on and so forth. In this scree plot, this would be one cluster, two clusters, three clusters, four clusters, and so on and so forth. This graph allows us to identify the ideal number of clusters. We want to be as low as possible in the graph with the lowest number of clusters that is possible. Ideally, we would like to be somewhere in the lower left corner, which would indicate that much variability is explained by a small number of clusters. In our case, a solution with four clusters constitutes a good compromise and the best approximation of being in this corner of the graph.

This leads us to the 4-cluster solution that you see here. If we are interested in the changes that the *get*-passive has undergone, then it makes sense to group these data points together and investigate what makes them differ from each other. We would find out what the early uses of the *get*-passive were like, what happens in this plateau here, and what happens in the later periods, in which many new examples with the construction appear.

To conclude this little clustering interlude, Variability-based Neighbor Clustering shows that in the case of the *get*-passive, the trend has four different temporal stages. The clustering algorithm gives us a periodization of the

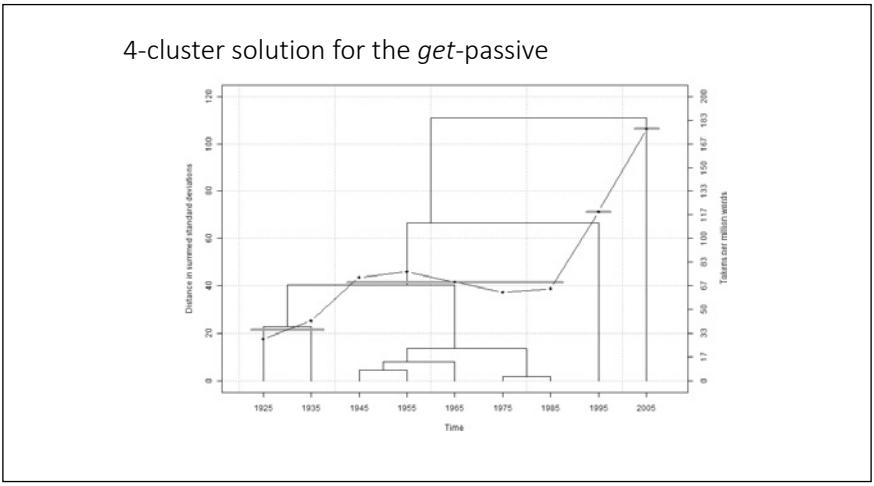


FIGURE 17

development. One thing that is unusual or at least not common in other types of periodization is that the periods are not of equal length. Normally, when linguists divide diachronic data into periods, they will choose equidistant periods such as centuries, half centuries, or blocks of 30 years. Here the periods can have different sizes. The clustering algorithm provides average values. In this case, the average values represent text frequencies. VNC can guide you towards structures that may otherwise go unnoticed or be hard to characterize objectively. I'd like to come back to the *V-ment* construction now. What does this method allow us to do with the changing productivity of the *V-ment* construction?

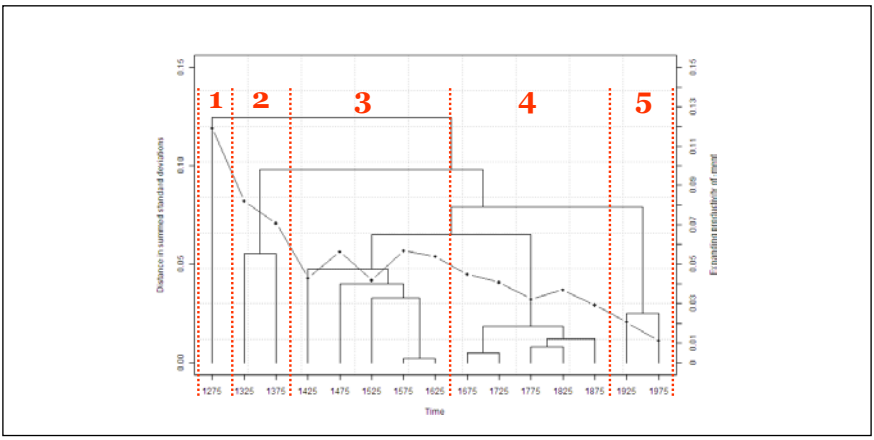


FIGURE 18

This slide shows once more the declining curve of expanding productivity. I took this curve and ran the VNC algorithm over it. The algorithm went through these measurements, compared the differences from one period to the next, and merged those that were the most similar to each other. What you see here is the clustering solution that the VNC algorithm produces. I used a scree plot of the kind that you saw earlier to determine how many clusters I would be justified to assume in this development. The scree plot indicated a distinction between five clusters, which allows us to account for a fair amount of variability without assuming too many clusters.

Overall, I cut up the development into five different time slices that are not all of the same length. The first period is just one half-century. Then we have two half centuries, the third cluster covers a slightly longer period. The fourth period describes a long decline. In the fifth cluster we have the modern period with which the development ends.

In the remaining time that I have, I want to talk about how these different periods can be analyzed, and what we can actually learn about the construction and its development on the basis of this periodization. Using the VNC algorithm to find periods is interesting, but it is really just the preparation for the actual analysis. It is a way of cutting up the data, but what you do with the data is something else.

For the analysis itself, I decided to determine a range of relevant variables that pertain to form and function of the *V-ment* construction. I took all the 1400 types that I had in the database and annotated them for these variables. Then I used a multivariate statistical analysis to explore whether there were patterns of variation that would change over time. Let me start by talking about the different variables that matter to the construction.

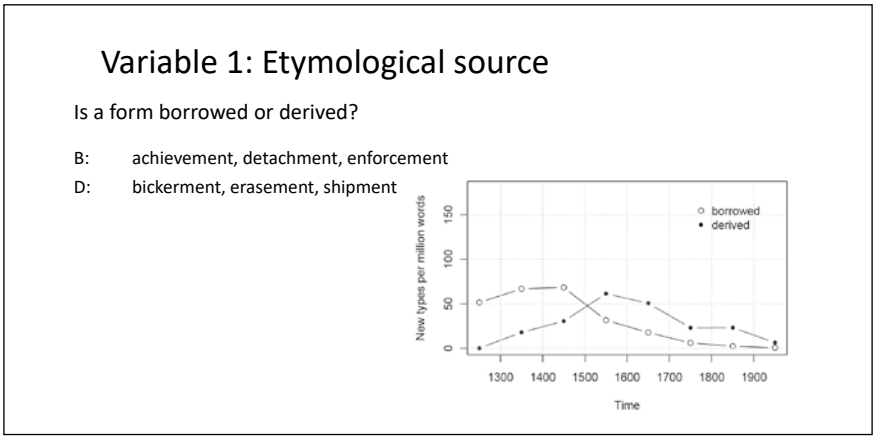


FIGURE 19

The first variable is concerned with the etymological source of the construction. If we have a *V-ment* word, is that word borrowed from French, or is it native to English? Is it derived? There are borrowed forms like *achievement* or *detachment* or *enforcement* that have Romance stems. Then there are forms like *bickerment* or *erasement* or *shipment* that are Germanic in origin. You see in this graph that we have more borrowed types at first, but that they decrease from 1450 onward. The derived forms peak around 1550, and then they decrease too.

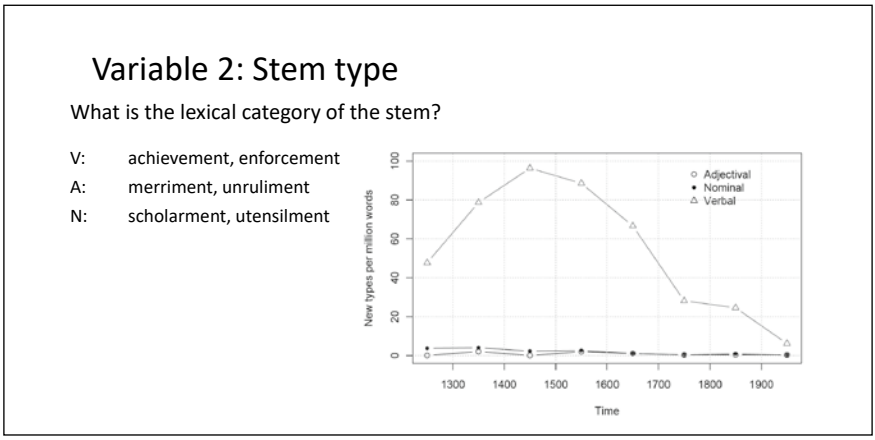


FIGURE 20

The second variable is concerned with the stem type. I mentioned that most types that we find have a verb. *Achievement* has *achieve*, *enforcement* has *enforce*, but there are types with adjectives, such as *merriment* and *unruliment*. There are types with nouns such as *scholarment* or *utensilment*. What you see in this graph is that these adjectival and nominal types are very rare, and that they occur only early in the history of the construction. Most types fall into the verbal category.

The third variable concerns the internal structure of the *V-ment* types. What is the internal hierarchical structure? There are words with a binary morphological structure. *Judge* is a monomorphemic word in English. *Judgment* is a word that has two morphemes. *Treatment* illustrates the same category. Then there are left-branching types, that is, words such as *enrichment*. The verb *enrich* is internally morphologically complex. The verbal stem of *belittlement* has two morphemes, and *-ment* attaches to it. Lastly, there are right-branching types. Those would be words such as *ecomanagement*, which is based on the word *management* that is prefixed by *eco-*.

Variable 3: Branching structure

What is the internal hierarchical structure?

Binary: judgment, treatment
Left-branching: [en+rich]ment, [be+little]ment
Right-branching: eco[manage+ment], non[agree+ment]

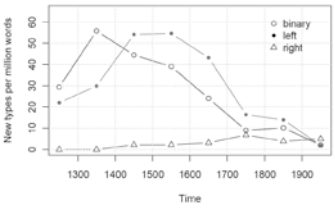


FIGURE 21

Looking at the frequencies, we see that at first the simple binary branching types dominate. They are then superseded by the left-branching types. The right-branching types only come in very slowly and very gradually, but they continue to rise in frequency even through the very latest periods. That already prefigures that they are actually represent something different than the *V-ment* construction itself.

Variable 4: Transitivity

Does the form evoke an entity that is acted upon?

Transitive:	arousement, punishment
Intransitive:	flourishment, merriment

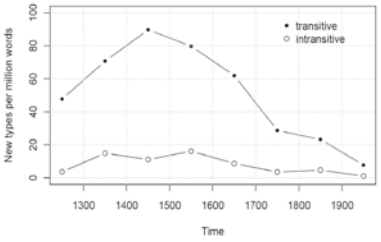


FIGURE 22

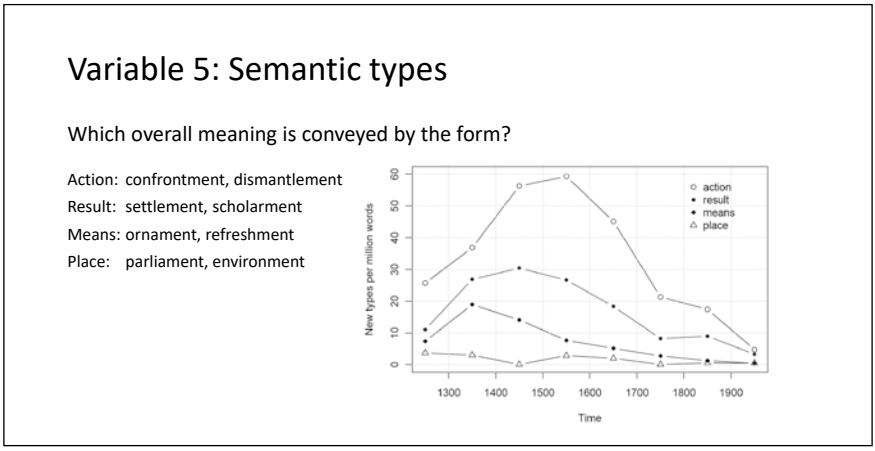


FIGURE 23

The next variable is concerned with meaning, and specifically with transitivity. For this variable, I relied on the argument structure of the verbs that form the stems of many of the *V-ment* types. Does the form evoke an entity that is acted upon? Verbs such as *arouse* or *punish* are transitive verbs that evoke a patient argument. By contrast, verbs such as *flourish* do not take a patient argument. They are intransitive verbs. This graph shows that there are many more transitive types than intransitive types. Both categories show decreases from around 1450.

Variable five is concerned with the different meanings that are conveyed by the *V-ment* construction. I already mentioned that the construction can express an action, as in *confrontment*, the action of confronting someone. It can also express results. A *settlement* is not the act of settling down, but the actual structure that characterizes a human dwelling place. There are types that express means. I have mentioned *refreshment* as an example. There is a fourth category that I assumed, namely, places like *environment*, for instance, or *parliament*. This category encompasses everything that I couldn't characterize as action, result or means.

All the variables that I have presented here are categorical. They represent choices between different categories, not continuous values that you can measure, such as frequency or degrees of productivity. Cross-tabulating categorical variables allows you to explore whether there are any interesting asymmetries in the way the data are distributed.

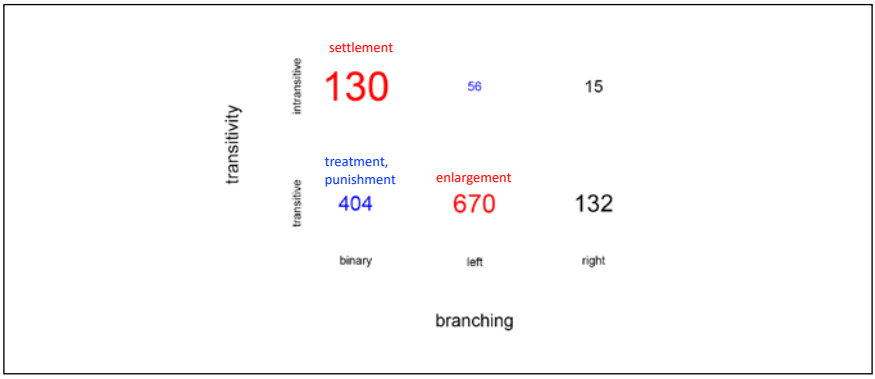


FIGURE 24

Here you see a cross tabulation of the transitivity variable with the variable of morphological branching structure. The rows show how many transitive types and intransitive types are in the database. The columns show how many binary-branching, left-branching and right-branching types are in the database. The observed frequencies that we have here can be compared against the expected frequencies that would be assumed if the distribution is random. A test such as the chi-squared test can determine whether the observed frequencies differ significantly from a random distribution. This kind of logic is visualized on the slide. If a number is shown in red, that means that the number of examples are significantly higher than expected. Blue numbers mean that a cell is significantly underrepresented. There are fewer examples than we would expect by chance. I have tried to visualize that also with the size of the font. We have 130 intransitive binary branching examples, and those are many more than we would have expected by chance. The other two intransitive types are relatively infrequent. To find that many in the upper left corner is something that is detected by the statistical test. You further see that in the upper row, most examples are in the leftmost cell. That is very different from the lower row, where left-branching transitive types, which are situated in the middle, are clearly overrepresented. Underrepresented forms include *treatment* or *punishment*, transitive, binary branching types. Overrepresented forms would be transitive, left-branching types like *enlargement* and also intransitive, binary-branching types like *settlement*.

Cross-tabulating two variables is just the start of the analysis. As you know, I didn't just have two variables. Of course, it is possible to cross-tabulate more variables than just two.

If you add a third variable, you have no longer a table that crosses two variables, but rather you have a cube that cross-tabulates three variables, but the

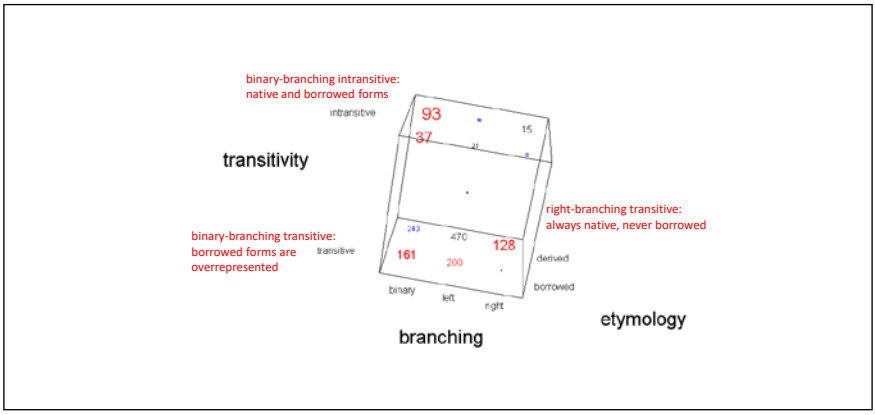


FIGURE 25

logic is essentially the same. For each cell in the cube you examine whether the types in the cell are overrepresented or underrepresented. Again, red numbers indicate that a cell is overrepresented, blue numbers indicate that a cell is underrepresented. Something that is clearly overrepresented in the data is the configuration of right-branching, transitive, and native types: forms like *embodiment*. Binary-branching transitive and borrowed forms like *judgment* are overrepresented as well. When we look at binary-branching intransitive types, both native and borrowed forms and are overrepresented. You can visualize this with up to three variables, but after that, ordinary physical space runs out of dimensions. Luckily, the computer can still handle it.

There is a method that is called configural frequency analysis which you can use for this purpose. The method is described in Stefan Gries' works (2009: 248). The method cross-tabulates a set of categorical variables and examines differences between observed and expected frequencies. I cross tabulated all 1400 types for the variables that I described earlier, taking into account the historical period, the etymological source, the stem type, the branching type, transitivity, and the semantic type. I wanted to find out whether there were configurations of values that would occur with greater than chance frequency during early periods of the data and later periods of the data. I wanted to see if early types would differ from later types.

In the remaining ten or so minutes that I have, I will go through the results very briefly. What happens during the early history of the *V-ment* construction? The first type, the first configuration of features that is overrepresented, is illustrated by forms like *commencement*, which are borrowed types that have a transitive verb stem. Types such as *imprisonment*, *conferment*, *enchantment* or *judgement*, those are words that are overrepresented early on in the history,

and this is consonant with results that earlier studies have obtained. One consensus is that early *-ment* types typically had transitive verbs as hosts. That is confirmed by the data. Let me move on to the second period.

During the second period there are certain oddities that we do not find later. This period sees the emergence of types such as *ointment*, that is, a cream that you put on a wound to make it heal better. These are borrowed types that have a verbal stem. The stem is transitive, the forms are binary branching. What is crucial about these types is that they describe a means, so an *ointment* is a means to healing your wound. The semantic type of means is not very frequent overall, but it rises to a moderate level of frequency during this particular period, and that is what the analysis picks up.

There is another type that also describes a means. That type is illustrated by words like *vesselment*. These are special because they have a nominal stem. There are a few others, including *monument*, but not many. This is a highly infrequent type, but because it is more frequent than expected, it shows up here.

Moving on to the third period, here we finally observe what we can call the prototype of the construction, words like *enlargement*. These are types that are natively formed with a verbal stem that is transitive, that is left-branching and that expresses an action. We have not only *enlargement*, we have *disbursement*, *misusement*, *renewment* and so on and so forth, lots and lots of types. This in fact is the most frequent configuration in the database. There are 174 types in period 3 alone, but there are more than 300 in total. This is very much the conceptual and formal core of the construction. This I found is interesting because Plag (1999: 16) comments on this construction and observes that there are certain forms are acceptable even though they are unattested, they do not exist as words in English. Even though the construction is unproductive, if I use a form that conforms to this prototype, people will accept it as a legitimate word. This concerns unattested forms such as *encodement* and *envisionment*. Those are not real words, but they look close enough to the prototype that we were fooled into thinking that we may have heard this word at some point or other. That is in line with the present analysis. Neologisms may be fine if the host is a prefixed transitive verb.

During this time, another configuration is overrepresented. There is this strange but short-lived fashion of words like *merriment*, which have an adjectival stem. This is only found in natively derived words, not in borrowed words. It is very innovative, but it is really short-lived. All types like *merriment* or *coldment* or *jolliment* or *adjustment* occur in a very short time span of about 60 years.

We are almost at the end. During this period that marks the decline of the construction, we find the overrepresentation of right-branching types, like *disembodiment*. We have something like *embodiment* and prefix it with *dis-*. That means it is not really the productivity of *-ment* that is at stake. Rather, the existing types of *-ment* are cannibalized upon by other prefixes. New forms are coined on the basis of existing types, which yields *maltreatment*, *overenrichment*, *reemplacement*, and so on and so forth. These are typically coined on the basis of the prototype form, like *enlargement*. This is thus an outgrowth of the prototype that is independent of the productivity of the suffix itself. This explains why the right-branching types actually continue to grow, even when everything else is in decline.

This continues through the last period, which gives us another overrepresented type that is also right-branching, but that does not verbalize an action, but rather a result. *Malnourishment* is the result of not having eaten enough for a long time. This type again dodges the strong trend towards the meaning of an action and it represents a metonymic shift from actions to results.

To come to an end, I started with these four questions and I briefly want to summarize the answers. The *V-ment* construction is a combination of a lexical stem and the suffix associated with different meanings. For our purposes, we can think of it as a network of constructions that grows and then fades over time. There are different patterns, different sub-patterns, like the type that leads to the form *merriment*. It is an outgrowth, a little part of the network that flourishes at some point and then contracts again. We have the prototype that is at the center of the network that starts as a cluster of borrowed forms and then expands on the basis of natively formed words.

I further asked how the construction changed in productivity. I said that contrary to proposals that were made earlier, there is really just one peak in productivity, which coincides with the nativization of the construction. My overall interpretation of that was that the construction was successful at first, but then faded very quickly. With regard to changes in form and function, the changes that we can mark here are the construction became nativized and a prototype emerged.

In the following, sub-constructions come and go. Lastly, the construction itself dies, and its types contribute to other developments, such as the emergence of right-branching types. Ultimately, productivity comes to a complete halt.

What about the puzzle that I started out with, i.e. lots of hapaxes but no productivity? My conclusion is that the remaining types and the right-branching words that are formed on the basis of it give us a sense of productivity that is

actually misleading, a kind of false productivity that inflates our expectations of how productive *-ment* really might be, if we are just looking at the suffixes and do not consider the rest of the structure.

What about constructional change? The changes that we observe here do not really fit the mold of common grammaticalization processes that I have been talking about them in earlier lectures. If we are looking for processes such as bleaching or host-class expansion or erosion or decategorialization, we do not really see that. The *V-ment* construction does not undergo grammaticalization. It also does not undergo lexicalization. We observe something else. The *V-ment* construction is created as speakers generalize over lexical borrowings for a brief period, then speakers experiment with this generalization and after a while the construction ceases to exist as a cognitive schema and just its lexical instantiations remain.

For the changes that we see here, we could decompose them into general processes of language change, like semantic change, reanalysis perhaps, but that would fail to explain a number of phenomena that I actually see as crucial to the history of the construction. The emergence of a derivational schema from borrowed forms, prototype effects like Plag's observation that a form such as *envisionment* is acceptable, and then the fading fashions of *merriment* and other forms like that. This means that the generalizations that we see in the history of the *V-ment* construction are local. We are looking at a network of constructions and sub-constructions, and these generalizations will sometimes capture the data more adequately than the broadest possible generalization. I mentioned the cognitive commitment that was proposed by George Lakoff earlier in this lecture series. Lakoff proposed another commitment, namely the generalization commitment. You try to find the broadest generalization wherever this is possible. I agree that the job of scientists is to find generalizations. We are supposed to connect the dots and state our insights in the broadest possible terms. However, sometimes local generalizations are really what is more crucial. The broader generalization might be an elegant theory, but it might not be an accurate reflection of what goes on in speakers' minds. When Ewa Dąbrowska spoke at this forum (2017), she made this point very eloquently, so I will just leave it at that here. The bottom line is that the analysis of language change can profit from adopting a constructional point of view that takes seriously this idea of constructional networks. With that I would like to come to an end and thank you once again for your attention.

LECTURE 6

Competition in Constructional Change

Welcome back to Ten Lectures on Diachronic Construction Grammar. In this lecture, I will continue with the general topic of constructional networks and the nature of the links between constructions and speakers' knowledge of language. In the last two lectures, I have talked about connectivity change. I have discussed the growth and the development of constructional networks, how they expand, how they contract, and how they shrink. I have reviewed different corpus-based methods that allow us to make observations about these processes, and I have explored what general conclusions we can draw from those observations with regard to proposals in grammaticalization theory and other frameworks. There is one kind of relation between constructions that I haven't discussed in depth so far. That relation would capture that two constructions are alternatives to each other. Some constructions have similar functions and they can be seen to be in mutual competition.

In this lecture, I want to explore the topic of competition in constructional change. Let us assume that two or more constructions have emerged as nodes in the constructional network. They have undergone constructionalization in Traugott and Trousdale's (2013) terms. They are connected with each other because they share part of their functional profile, perhaps even part of their form. They would constitute what is often called an alternation. That is a term that goes back to generative syntax and the idea of syntactic transformations. The implication at the time was that two members of an alternation, two member constructions would be seen as instantiating the same underlying structure. This would be illustrated for instance by the active voice, *John drove the car*, and the passive voice, *The car was driven by John*. The passive would be seen as a transformed variant of the active. This particular idea of transformations has fallen out of fashion. I have mentioned Goldberg's surface generalizations hypothesis (2002), which goes in the exact opposite direction by

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691122.

stating that surface form are important, not generalizations across alternating variants. It is however safe to say that despite all this, alternations have made something of a surprising comeback in recent years.

There is something very attractive and interesting about alternations, which are two ways of saying the same thing. You might wonder why languages afford this kind of luxury, to have two things for the same purpose. That is something worth thinking about. Personally, I wasn't drawn to this problem naturally, but I have been convinced that there is something to it and that there is something to be analyzed. I have been strongly influenced, for example, by the work of Benedikt Szmrecsanyi (2006), who I mentioned earlier today. He and his work have opened my eyes towards the intricacies of alternations, especially as we find them in English. We find them presumably in all kinds of languages, but there has been a tremendous amount of work on English.

I am going off script here, but I need to tell this little joke that I stole from Jack Dubois. Jack Dubois once opened a plenary talk that he gave with the sentence, "*The study of language encompasses the fields of phonology, morphology, semantics, and the dative alteration*." That is a quip, but there is some truth to it. The dative alternation is one of those pairs of constructions that have been studied extensively. There are of course other alternations in English. Besides the the dative alternation, there are the two genitives, i.e. the *s*-genitive and the *of*-genitive. There are two comparatives, i.e. *prouder*, the morphological variant and *more proud*, the periphrastic variant. The members of alternations are fruitfully analyzed in terms of competition between constructional variants. That is what I will be talking about in this lecture.

What I will have to say relates to basic idea #1 that I presented in the very first lecture. The idea is that all of linguistic knowledge, according to Construction Grammar, is a network of form-meaning pairs and nothing else in addition. Competition in the network of constructions can be understood in terms of two nodes that are connected in the network. Competition arises if a speaker wants to express a particular meaning, and that meaning is connected to two different forms. Which form will the speaker choose? Given everything I have said so far about how the constructional network is organized, that would depend on the strength of the symbolic link between the meaning and one of the respective forms. The stronger connection wins and the weaker connection loses. We can further imagine a feedback mechanism that is activated when a construction is selected from the alternation. The winning connection might be rewarded and might be made even stronger, and the losing connection could be punished, so that it is even weaker than berfore. The next time the speaker wants to express that idea, any bias that was there in the first place would be even stronger the second time around.

This is a question that has generated quite a bit of discussion. When one construction changes, what happens to other constructions that are connected to it? The common idea is that one construction's success leads more or less directly to the demise of another competing construction. This is very much the heritage of structuralist approaches. The underlying thought is that language change is systemic. We see a lot of evidence for that in sound change. The phonemic inventory of a language can be seen as a system where everything holds everything else in place. Construction Grammar has adopted the idea that constructions can be in competition if their meanings and functions are sufficiently similar.

There is a recent paper by Hendrik De Smet and colleagues (2018) from the University of Leuven that I would like to draw your attention to. It has been published in *Cognitive Linguistics*. In that paper, De Smet and colleagues comment on that particular idea:

> The relation between functionally similar forms is often described in terms of competition. This leads to the expectation that over time only one form can survive (substitution), or each form must find its unique niche in functional space (differentiation).

This captures a broad implicit consensus in the field. However, De Smet and colleagues ultimately question this idea and propose a more nuanced perspective. I will come back to this tomorrow.

For now, let me just say that I have gone on the record with a somewhat non-canonical position with regard to this here. I have already shared that with you. In my book on constructional change, I have advanced the idea that "grammatical change is not a zero-sum game". I thought that slogan would catch on, but it didn't. Anyway, I had my doubts whether the adoption of new functions by one construction would necessarily drive out another construction that serves the same functions. When we look at languages, partial functional overlap is common, not rare. Inversely, we could say that there are phenomena such as polysemy. Polysemy is rampant. If we had perfect structural systems with one-to-one form-meaning correspondences where any expansion of meaning would lead to the retreat of another form, then that really shouldn't be there.

Another qualitative piece of evidence is that grammatical constructions tend to emerge in domains that are already relatively well represented by other constructions. Let me give you an example, English has nine core modal auxiliaries, and that constitutes a working grammatical paradigm. Yet, there are new modals that are coming into the grammar, forms such as *got to* or *have to*, which perform approximately the same function as forms that are already

there. We already have *must*, why do we need *got to*? Why do we need *have to*? I do not think that competition can explain that. You need to make additional assumptions that explain why speakers start to use expressions like *I have to go*, but these assumptions are not predicted by competition.

Bill Croft has made a point that I see as compatible with my views here. He has argued that functional pressure, the kind of pressure that would bring constructions into competition, is not what drives competition between constructions. Functions are crucial, he says, when speakers innovate. Speakers are keen to express a certain idea. They're looking for new ways of expressing a given idea. This is what Croft (2000) calls "altered replication of linguistic forms". For example, they say *I have to* instead of *I must*. They say *I am going to* instead of *I will*. Competition happens in what Croft (2000) calls "propagation". An innovation is already there, propagation describes how it spreads through a community of speakers. Propagation, Croft (2000) argues, is not functionally motivated. It does not have anything to do with how well a form is functionally adapted or how useful its meaning is. Rather, Croft argues that it is socially motivated: "*The basic mechanism for propagation is the speaker identifying with a social group*". What I take away from this insight is that when we study constructional competition, it is very relevant to include social factors into our analysis.

In what follows, I would like to present to you a study of constructional competition from my 2013 book, in which I try to do that. Besides these theoretical considerations, studying competition also has some implications for methodology. I will be turning to a corpus-based method that moves away from collostructional analysis and the other types of analysis that I have been discussing up to this point.

How do we study competition? If we want to understand how competing forms develop over time, it would be good to have analytical methods that study processes of language change more comprehensively than just in terms of text frequency. Most of what I have discussed so far relies crucially on text frequency, but that is not all there is. The frequency of an item may stay constant, but it may still undergo interesting developments that merit our attention. Second, our methods should allow us to identify the explanatory factors that drove a linguistic change and when and why they did so. Also, the method should be able to distinguish between factors that are more important and less important or that exert less or more functional pressure. Lastly, our analysis should facilitate theoretical generalizations rather than just presenting us with single case studies or give us empirical facts.

In this lecture, I would like to present a case study of allomorphic change. This takes us from the verbal grammar that we looked at yesterday and the

nominal grammar that we looked at this morning into a smaller level of linguistic organization. I want to focus on the development that happened to the English possessive determiners. Those are the words *my* and *your* in Present-day English. In the period of English that I am looking at, the Early Modern English period, several variants of these possessive determiners were in use. The first singular form was *myne*. Correspondingly, the second person there was *thine*, which begins with a voiced interdental. These forms are obsolete. If you do not recognize them, that is because you were born too late. These forms dropped their nasal final consonants, so they turned from *myne* and *thine* into *my* and *thy*. *Thy* corresponds what we now express as *your*, which came about through a further change that took place after the period that I am investigating here. I want to exemplify how diachronic corpus studies can usefully adopt methodologies that have been developed in Labovian American Variationist Sociolinguistics, which has a great deal to teach us in this context. I want to explain how change in allomorphy relates to the issue of competition in constructional change.

Let me start by discussing allomorphy. There are a few examples of allomorphy that students of English linguistics have to learn in their first introduction to the field. The phenomenon concerns phonemic variation between two or more realizations of the same morpheme. One of the examples that is often used to illustrate this is English plural allomorphy. We have *ships* with a voiceless /s/, we have *harbors* with a voiced /z/, and we have *cruises* with also a voiced /z/ that has a /ə/ before it. Another type of allomorphy is apparent in the two indefinite determiners of English, so there is *an* /ən/ *apple*, there is *a* /ə/ *banana*.

What conditions this variation? What causes speakers to choose one or the other? The variation between *an* /ən/ *apple* and *a* /ə/ *banana*, you know very well, is conditioned by the sound quality of the next element, i.e. the initial element of the noun. Allomorphy is often conditioned by phonological factors in this way. Typically, allomorphy is discussed in the context of grammatical morphemes. Allomorphy appears in the lexical domain in the phenomenon of stem allomorphy. Typically, speakers do not vary in their choices. We do not find English dialects in which the distribution of a /ə/ and an /ən/ is radically different than in others. That does not mean that there is no variation at all: Alternations such as *dreamed* and *dreamt* show competition between a regular variant and an irregular variant of a past tense form. The forms *cannot* and *can't* are variants that differ in terms of reduction. The variation between tom/eɪ/to and tom/a:/to is due to regional linguistic differences.

What I will be discussing in this lecture is a phenomenon that used to be variable, that used to have a number of different factors governing variation,

but that eventually turned into a fixed state and the variability disappeared. Today, speakers do not say *myne old friend* any longer, instead they say *my old friend*. There was a time when it was possible to be linguistically progressive and say *my old friend*, and when you felt a little bit more conservative, you would say *myne old friend*. The same applies to the second person forms *thine* and *thy*.

Why now would it be advantageous to study allomorphic variation from a constructional perspective? A critic might object that formal variation that is not tied to any tangible meaning variation is in fact irrelevant to Construction Grammar, where we are concerned with form-meaning pairings. When we have two forms that map onto exactly the same meaning, that is not something that many constructional studies focus on. *Myne* and *my* definitely encode the same meaning. Furthermore, if not only the semantics but also the syntactic behavior of the allomorphs is nearly identical, then there would be little left to analyze from a constructional perspective. But contrary to that position, I would like to argue that the issue of allomorphy relates to the question of what a construction actually is, but let's move ahead and get to the actual analysis now.

The starting point of any analysis is of course the data. For this study. I used the Penn Parsed Corpora for Middle English and Early Modern English.

The phenomenon that I am interested in is the one of possessive determiners, the words *myne* and *my*, and *thine* and *thy* in front of a nominal, as in these examples here: *with myne own eyes* or *the daies of my life*. In the Penn Corpora, there is considerable variation with regard to possessive determiners. There is inter-speaker variation, so that different writers use them differently, but also intra-speaker variation, which means that the same writer uses sometimes this form and sometimes another. Many writers use both *myne* and *my*, that is, forms with the final nasal and forms without it. Sometimes they show a certain preference. One writer may have a preference for the conservative variant. Another may show a distribution that is more progressive. Is there an explanation for this variation? In fact, there is not only one explanation. There are many interlocking explanations.

Let me discuss some of the explanatory factors that I took into account. Trivially of course there is time. Later texts are more likely to contain the modern n-less variant, but there is the phonological context too. If the word directly following the possessive determiner begins in a consonant, as *in my life*, that would favor the n-less variant. Stress patterns play a role. When I say "*That was not MY idea, that was HIS idea*", that favors the n-less variant. It is a little tricky to extract stress patterns from writing, of course, at least this

kind of contrastive stress, but there are well-conventionalized stress patterns of words that help us a little bit. There are also language-external factors that condition the variation. Women are known to be progressive across many linguistic changes, and this is no different here. When we look systematically at the gender of the authors, women can be seen to favor the progressive n-less variant. It has been proposed that formality plays a role, so that formal genres introduce a bias towards the n-variant. Then there are frequent collocations such as *myne own son*, which would be *my own son* in modern English. Why should frequent collocations be less progressive? We know from other studies that chunks, frequent collocations, are conservative. They are produced very often together. Within a strong collocation, within a chunk, an old variant has a stronger chance of survival. You see this in some idioms and in some expressions that preserve sort of old syntax or old morphology. Shakespeare could write "*I know not*" instead of "*I do not know*". We still have that syntax in idioms like *I kid you not*, where we have the ancient pattern of negation. You can't do that with other verbs or with expressions that are not idiomatic in any way.

In the following, I want to address three questions. First, which of the proposed factors have a reliable effect on the choice between *myne* or *my*? Second, did the effects stay constant over time or did they change? Third, did *myne* and *thine* change in the same way at the same time, or would we be forced to say that they follow different trajectories? That is important to find out because we want to know whether speakers have formed a single generalization for all of these forms, or whether there is a separate generalization for *myne* and another one for *thine*.

I will start with the first question here, which of these factors have a reliable effect? For that, relied on some help from variationist sociolinguistic approaches, which have long been concerned with exactly this problem. There has been a focus on alternative expressions. Take for instance the quotative *be like* that I have mentioned already a couple of times in this lecture series. When you quote someone else's speech, you could say "*He was like that is great*", or you say "*He said that is great*". Social factors can explain to some extent whether you choose one or the other. The lesson that many sociolinguistic studies have taught us is that speakers' choices tend to be governed by probabilistic explanatory factors. In the case of the choice between the two genitives, i.e. the *s*-genitive and the *of*-genitive, maybe the most important determining factors are not social but rather language-internal, including semantic factors such as the animacy of referents. There are syntactic considerations, so constituent length plays a role, speakers are generally more likely to place long constituents towards the end of an utterance, at least in SVO languages such as

English. There are differences between different text types, and there is a host of other extra-linguistic factors like age, gender, the speech situation and so on and so forth.

In the database of examples that I used for this study, we find examples such as *the daies of my life*. What can we do with that? We can actually determine the relative frequencies of *my* and *myne* in different contexts. Then we ask how likely the use of the n-less variant is in this case, given that we have a text that was written in the year 1564, given that the next word starts with a consonant, *life*, the writer is male and the text is from an informal genre.

Given all of these factors and all these variables and their levels, how likely is this outcome of the progressive n-less variant? This is very much a standard procedure that has been established in sociolinguistic multivariate studies. By now there is also a sizable literature that applies logistic regression methods to morphosyntactic variation.

Turning now to the second question, did the effects of the involved factors stay constant over time, or did they change? We know that the variation between *myne* and *my* was transient. Today, *myne* only survives in deliberate anachronisms, which means that factors that once triggered the use of *myne* are no longer effective today. You no longer know about this form. The system has turned from a network of probabilistic competing factors into a fixed system where everything is discrete. The analysis of changing effects is something that is a little trickier to analyze, and yet it very much describes a well-attested scenario of language change.

For example, historical studies show changing impacts of explanatory factors. Again, I am coming to Benedikt Szmrecsanyi and his work here. Together with other colleagues (Wolk et al. 2013), he has documented changing patterns of varition in the English *s*-genitive, as in *my brother's car*. That construction used to be very strongly restricted to animate possessors, but that restriction has loosened. Today we can say *the company's car* or *the university's policy*. The effect strength of animacy has changed over time. For earlier speakers, animacy was a strongly determining factor, and for present-day speaker, that effect has become weaker.

Then we have apparent-time studies showing different impacts for different age groups. Again, let me take the example of *be like*, which is preferred by female speakers, but most strongly by adolescent female speakers. There is an effect strength of gender that interacts with age. Speakers use this form specifically if they are young and female. We can establish that the processes of language change involve change in the ecology of conditioning factors. Some factors become more important, others may fade away and then cease to be important. Entire domains of variation may eventually become fixed and

fossilized. I would like to argue that diachronic corpus linguistics can offer a contribution to the study of such phenomena.

Before we get to the actual analysis, I still need to briefly talk about the third question, namely, did *myne* and *thine* change in the same way at the same time? Why is this question important? If the two forms change in different ways, that would be evidence for the idea that *myne* and *thine* each formed a generalization of their own. Hence, the two would be different, if of course related, constructions. What the question boils down to is whether we are dealing with one constructional change or in fact with two constructional changes that happen simultaneously.

With all of these in mind, how did *myne* and *thine* change to *my* and *thy*?

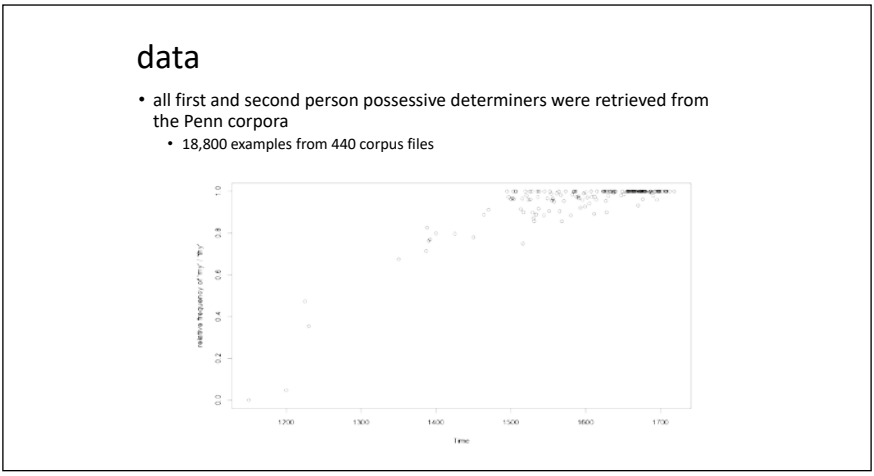


FIGURE 1

Here you see a graph with what sociolinguistics call an s-curve. You see the relative frequencies of the new n-less variant pooled for all corpus files of the same year. The higher a point is on the graph, the closer the data is to present-day usage. You see that the scale on the *y* axis goes from 0 to 1. The value 0 means all writers use the old variant, 1 means all writers use the new variant exclusively. It is plain to see here that there is much more data for the final two centuries than for the previous ones. We know more about the final stages of the development than we know about the earlier stages. But all in all, the *s*-pattern is fairly clear. Of course we know more about the data than just relative frequencies. All examples that I found in the data, all in all some 18,800 examples, were annotated for the explanatory factors that I discussed earlier. The data points you see on this slide come from 440 different corpus files, and every corpus file has the potential to show some intra-speaker variation.

The examples were annotated for eight variables. Let me anticipate a question: Why were these variables chosen and not others? The answer is that I took six of these variables from the existing literature, which had identified them as relevant. To that, I added two other variables that I considered to be important.

The first variable concerns the phonological environment. Does the word after *my* or *thy* or *myne* or *thine* begin with a consonant? Does it begin with a vowel or does it begin with an H? H is used as a third category here because there is variation in the way it is pronounced. Sometimes it is more vocalic, sometimes it is more glottal. We can assume that uses before H will exhibit a mixed profile.

I already mentioned stress, so for the second variable I took the stress pattern of the following word and annotated if the first syllable was stressed or not. English is a stress-timed language, so words have conventional stress patterns. It is *uniVERsity* /ˌjuːnɪˈvɜːsəti/, not *UNIversity* /ˈjuːnɪvɜːsəti/.

Another variable distinguished between first and second grammatical person. Do we have a first-person form like *my*? Or do we have a second person like *thy*?

I distinguished two levels of formality. Formal genres included texts from the Bible and law texts. Informal genres were personal letters or comedy texts.

Then is the gender of the writer. Are we looking at a text written by a male speaker, a female speaker or is the gender of the writer unknown? In historical documents, that is quite often the case. Another complicating factor is that many texts were actually written by scribes. Speakers are often nobility who would dictate their letters to scribes who wrote them down. We know from other research that these scribes didn't impose too much of their own preferences on those documents, but they were fairly faithful to whatever it was that the original writer of the letter said.

Now, on to the variables that I included, but that are new to this particular phenomenon. One variable I included is priming. If we look at the previous left context in the text, the previous 50 words, do we have an instance of *myne* or *thine* or *my* or *thy* in that context? If that is so, then there are reasons to think that the writer will be influenced. If a form is still activated, that would make it more likely that a writer will choose the same variant again.

The next variable concerns the relative frequency of the following element. There are some expressions that occur very often in the data that I analyzed. An expression such as *myne own* for example occurs very frequently. There is abundant evidence that increased frequency of a linguistic item leads to conservative behavior. Since *myne own* is produced and processed as a holistic unit, speakers are reluctant to switch to *my own*, even if they do so with other less frequent collocates.

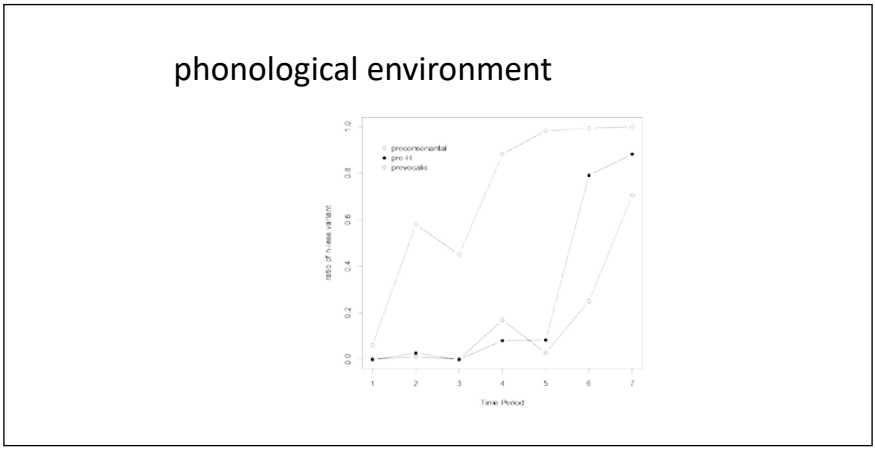


FIGURE 2

Let's now look at all of these variables and let's see how their impact may have changed over time. This graph shows the s-curve that you saw earlier with the raw data, but here we see it split up for the three different levels of the variable of the phonological environment. Three s-curves, one for possessive determiners that appear before a consonant, before an H and before a vowel. Again, the higher the data point, the higher the proportion of *my* instead of *myne*. What you can see in this graph is that the three curves start at different points in time and get to the final destination earlier than others. The effect of the phonological environment thus has a temporal contour. The switch from *myne* to *my* starts in pre-consonantal environments, like *my life*. Possessive determinants before H as in *my head* and before vowels, *my own*, remain on a low level until very late. The curve for the preconsonantal environment starts out in the very beginning, while the others continue to stay low. The difference is largest in period 5, close to the end, and the others catch up after that.

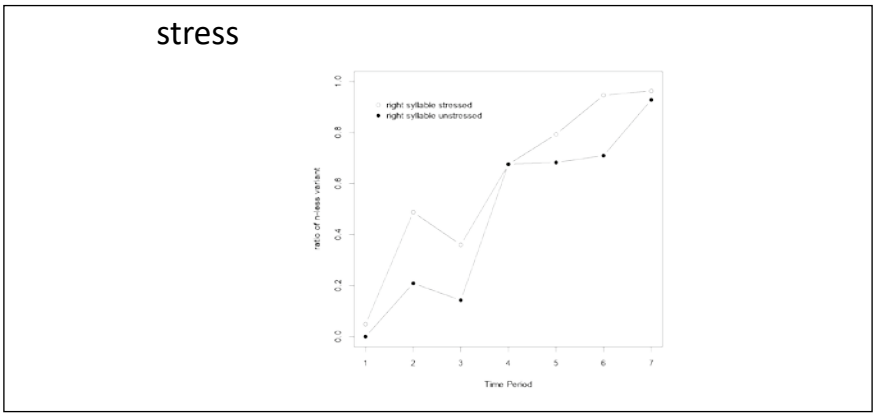


FIGURE 3

Let's move on to stress. In this graph, we see that the factor of stress on the falling syllable has a clear effect, so the n-less variant is more frequent in pre-stress environment, that is with expressions such as *my father* as opposed to *my idea*. But also here, the effect is not equally strong across all periods. In periods 1, 4 and 7, there is practically no difference. The effect is not completely uniform.

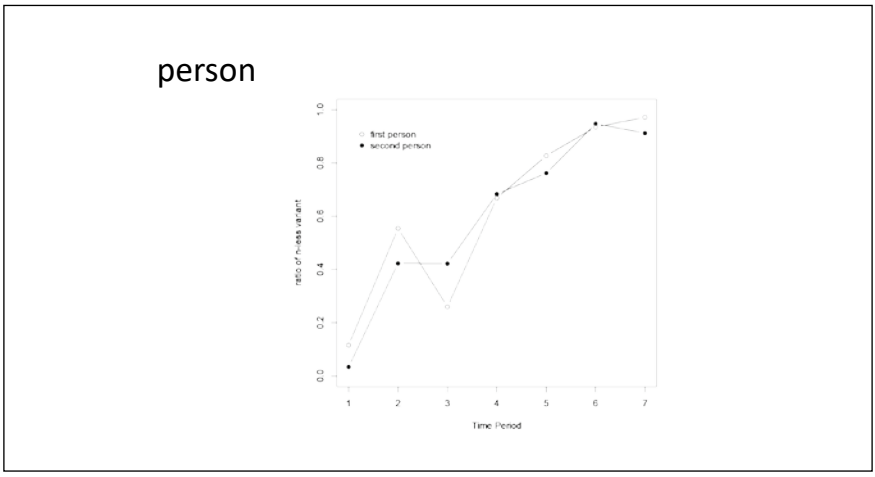


FIGURE 4

With the variable of grammatical person, we hardly see a difference. That is quite noteworthy. If we discard period 3, which shows a little difference, then we could say that on the whole, first person forms are a little bit in the lead. First person forms are the white circles here. *My* is leading the way for *thy*, but yet it looks as though the developments from *myne* to *my* and from *thine* to *thy* proceeded very much in lock step, presumably because they are phonologically very similar. Speakers form a generalization across *myne* and *thine*, so that in contexts appropriate for *my*, they were also deemed to be suitable for *thine*. I will come back to this later in this lecture.

Here's formality. Formality has been argued to be a conditioning factor for the older n-variant. The Penn Corpora actually do not confirm this idea. You see that the two curves are not distinguishable from one another. Very likely, there is no effect here, not even a transient one.

What about gender? Here you see a problem, namely that in historical documents, the gender of the writer often cannot be known with certainty. The only reliable data we have is from the last three corpus periods. Reassuringly, in those three periods, female writers are the black dots, they are highest up. They are in the lead. Males are the most conservative. Males are the empty circles here. We have the unknown gender writers in the middle of them. Even though

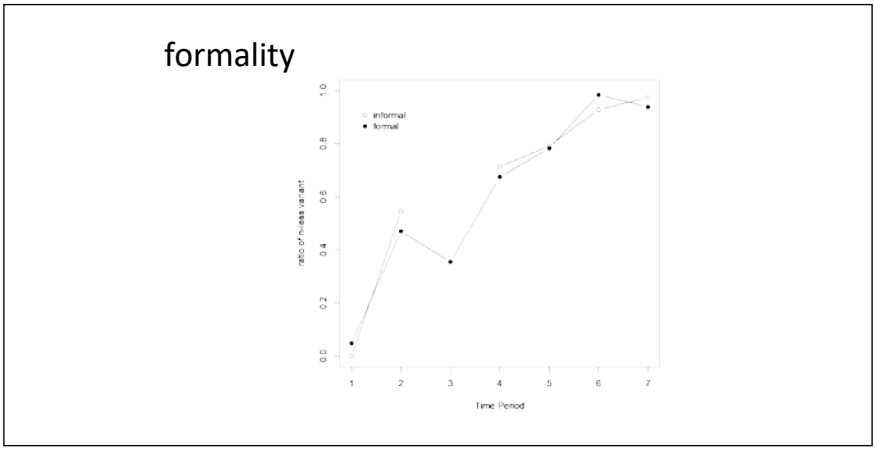


FIGURE 5

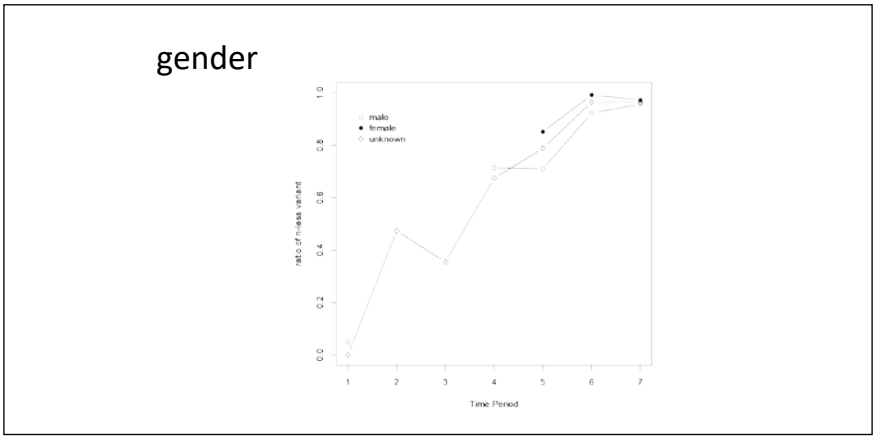


FIGURE 6

the data is patchy, as I am ready to admit, I would go out on a limb and say that there probably is a small effect of gender.

Human language users are creatures of habit. If they use a form once, they are likely to use it again. This graph shows you that the n-less variant is indeed more frequent if it has been preceded by another n-less variant. Again, the only exception to this tendency is period 3. You see here, the lines cross and re-cross again. That might be an oddity in the data. Period 3 is actually something of an outlier here in this data.

So far, we've covered phonology, stress, person formality, gender and priming, we are still missing frequency. For that, I would like to show you a graph that may look a little strange at first.

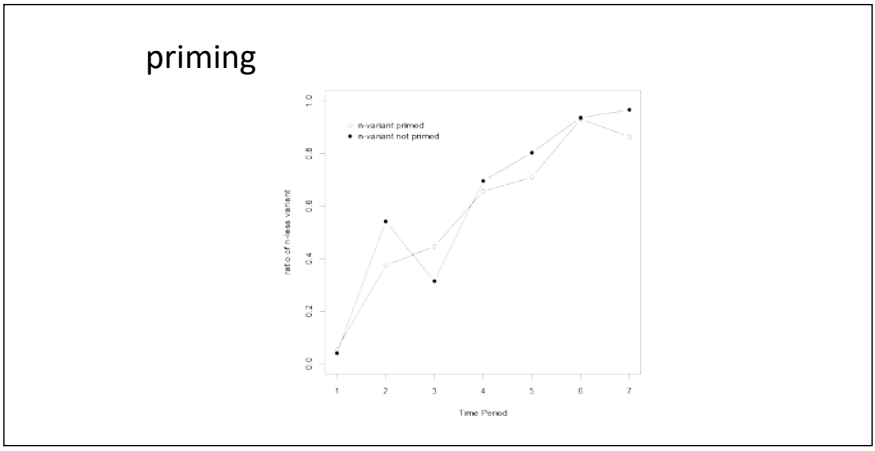


FIGURE 7

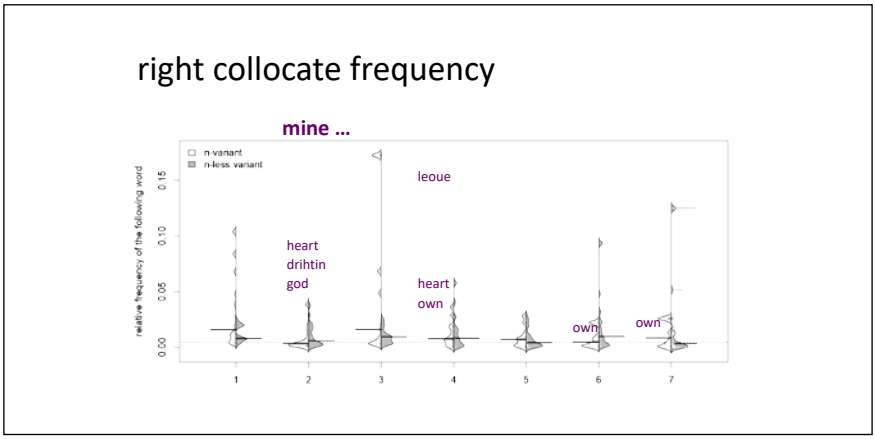


FIGURE 8

What you see in this graph are again the seven corpus periods. For each period, you see uses of the n-variant in white and uses of the n-less variant in grey. There are wavy patterns to the left and right in each section of the plot. Those are frequency distributions of the right side collocate. The interesting pieces of information are the little bumps that you see high up on the graph. Those are collocates that appear with high relative frequency in the respective periods with *myne* and *my*. My hypothesis was that the n-variant would show a preference for high frequency items, because high frequency chunks would be stored as single units, so they would be conservative. I expected *myne* to appear a lot with very frequent items. A pairing like *myne own* illustrates this. The little bumps in white, on the left-hand side of each line are in line with

that prediction. However, high-frequency collocates are not restricted to the n-variant. Frequent items with *my* appear as well, which is the opposite of what I predicted. Let me show you the actual words that correspond to the bumps. In the early periods, high-frequency collocations include *myne heart* or *myne god*. In period 3, we have *myne love*, which translates as *my dear*, a term of address. In later periods, we have *mine own* in period 6 and 7.

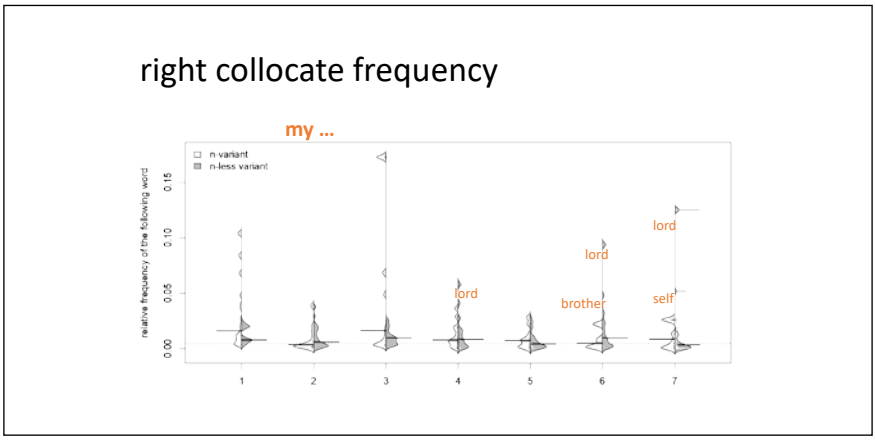


FIGURE 9

Conversely, if we look at the frequent collocations with *my*, we see that the n-less variant has also a good number of chunks and most prominently so in the last period. Here we would have expected greater numbers of low frequency items. Again, frequency has an effect, but the effect is inconsistent. This suggests that my initial hypothesis was too naive.

Let me summarize what I have said so far. The n-less variant is increasingly more frequent in later periods. With regard to the following segment, the n-less variant originates in pre-consonantal environments and then spreads to pre-vowel and pre-H. As for stress, the n-less old variant is more frequent before stressed syllables. With regard to person, the n-less variant is a tiny bit more frequent in the first person. With regard to gender, the n-less variant is a little more frequent in texts of female writers. I didn't find any consistent effect of formality. There is a small effect of priming. Collocate frequency does not show a consistent effect. There is something going on, but it is not clear what. Knowing all of this is a good start, but the impact of different variables across time is only one part of the story.

The piece that is still missing concerns the third question. Did *myne* and *thine* change in the same way? We've seen that they change roughly at the same time, but how similar or how different are they with regard to their

conditioning variables? If there are differences, this would suggest that there are two changes taking place and not just one. Let's take a look.

On the following slides, you will see a number of so-called mosaic arrangements. One for first-person, forms on the left, and one for second-person, forms on the right. For the n-less variant forms *MY* and *THY*, they are shown in the lower half. The n-variant forms are shown at the top. Overall you see that the modern n-less forms are overall more frequent in the corpus than the old n-variant forms. You also see that in the second-person mosaic, there are relatively more n-variant forms. There is a difference between *MYNE* and *THINE*. Overall the percentage of *THINE*, relatively speaking, is greater. For demonstration purposes here, the areas of first and second person forms are shown as equally large. There are of course many more first-person forms than second person forms. I just want you to be aware of that.

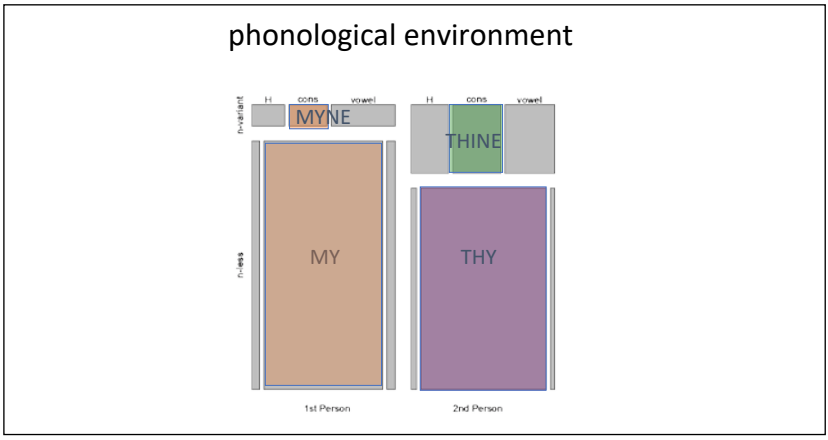


FIGURE 10

The crucial information in this graph concerns the differences in the way in which the four areas are internally partitioned. What you see is that *my* and *thy* mostly occur before words that start in a consonant. This is simply because adjectives and nouns, the elements that follow possessive determiners, have a tendency to start in a consonant. With *myne* and *thine*, consonants are underrepresented while vowels and H'es are overrepresented. The question now is, are vowels and H'es equally overrepresented across first and second forms, or is there a difference?

When we look at the upper parts of the graph, it appears that first person forms have a slightly bigger preference for prevocalic environments. If you compare the brown block against the green block, then the brown block is wider, but the effect is not very pronounced. Likely, there is no difference

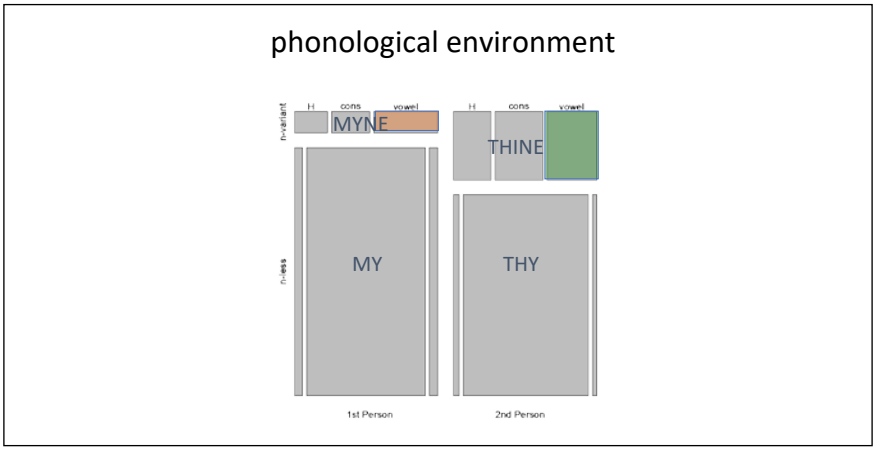


FIGURE 11

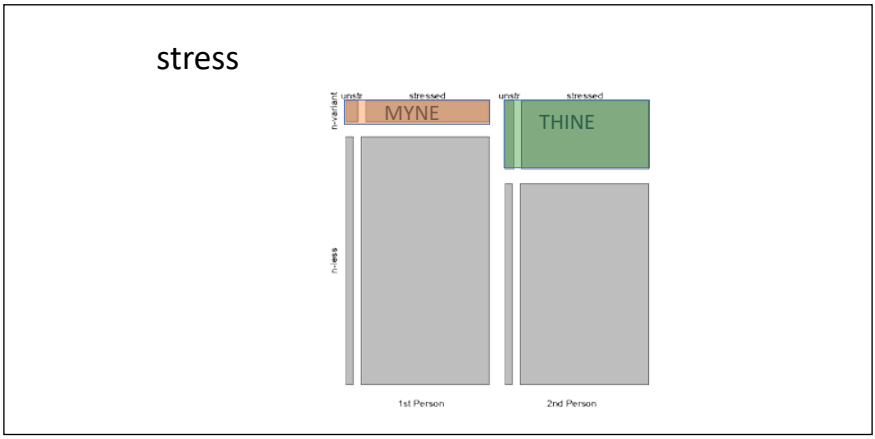


FIGURE 12

between *MYNE* and *THINE* with regard to the first variable of the phonological environment.

What about stress? Here, we have the same kind of mosaic design. This time, the partitioning contrasts unstressed and stressed beginnings of the following word. You see that in most cases, the following syllable is actually stressed. If we look at the upper parts of the graph, both *MYNE* and *THINE* show a slightly larger proportion of unstressed following syllables. With *MYNE*, this preference might be just a tiny bit stronger, but again, there is definitely not a strong effect.

Things look different when we come to formality. For the mosaic for formality, what is striking is that the n-variant shows a greater ratio of examples from

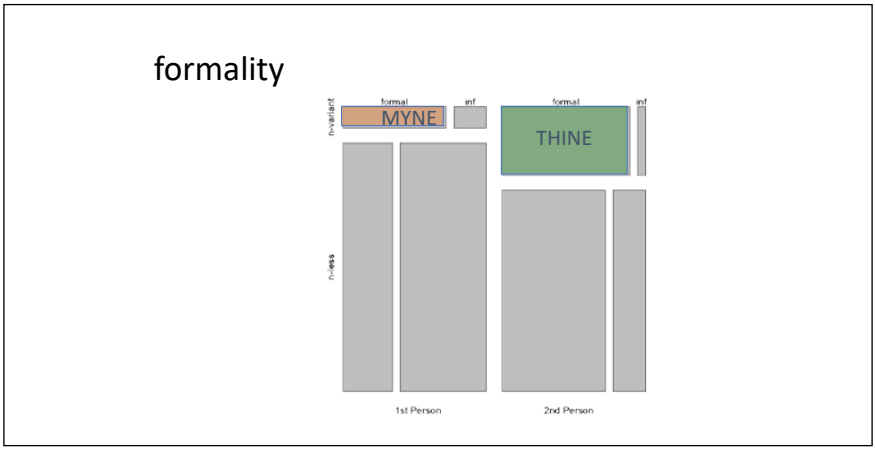


FIGURE 13

formal genres. With *MYNE*, formal examples are colored in brown. For *THINE*, which are shown in green, there are many more examples in formal genres. There is a straightforward explanation for this effect. *THINE* hardly ever occurs in informal texts, because the alternative form *YOUR* develops in this period, from which these data are taken. The rise of *YOUR* definitely interferes with the usage of *THINE* and *THY*. But it does not interfere with *MYNE* and *MY*. Here we have suggestive evidence to the effect that first and second person forms might in fact have developed in different ways. According to the variable of formality, *MYNE* and *THINE* do likely not develop in identical ways.

I would like to come to gender. This mosaic plot shows an interesting but somewhat minor difference. Note that female users are relatively less likely to use *THINE* in their writings. You see that females have this very thin strip of *THINE*, shown in green. They use *MYNE* a little more in comparison. It is true that most female writers that are in the data come from later periods of the corpus. This could be an artifact, but that kind of artefact can be addressed in the statistical analysis that I am about to talk about.

Lastly, it remains to be discussed whether the effect of priming is different across first and second person forms. This is the last graph of this kind. Here, it is evident that the n-variant benefits from priming. You see that the areas for n primed are wider in the upper parts of the graph. In the lower parts, in the strips that we see for the n-less variant that are colored in brown and purple, they are more narrow. Primed examples account for a bigger share of the overall number of examples than in the n-less examples. The lesson from that is clear actually. You do not need priming to get *MY*, but priming may help you to get *MYNE*. When there is a form that it is on its way out, you are generally less

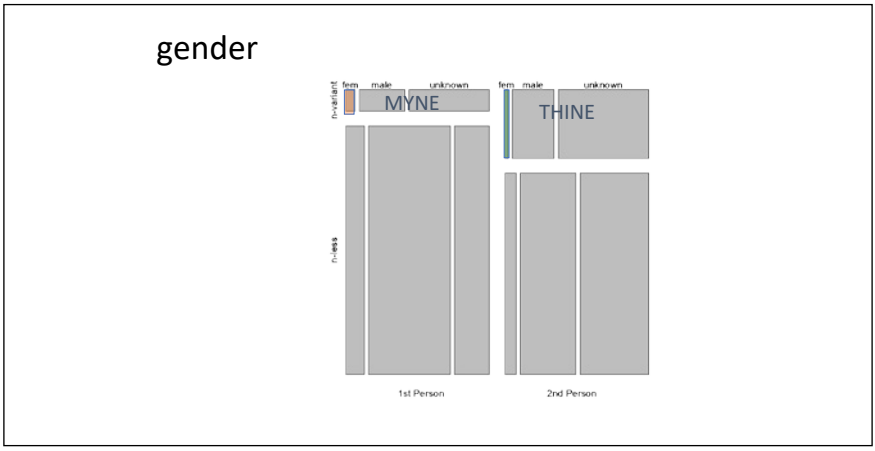


FIGURE 14

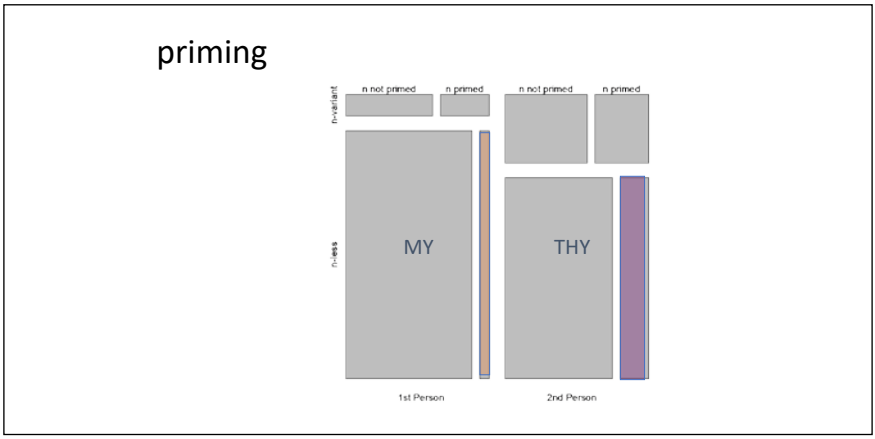


FIGURE 15

likely to use it except when you've recently heard it. That is an opportunity to remember it and then come back to use that old variant. Is this effect different across the two persons? Probably not. It looks very similar, but again, this is something for the regression analysis to figure out.

If we now come back to the three questions one more time, we can try to formulate a few conjectures. Which of the factors have a reliable effect? We're fairly certain that time, the following segments, stress, person, gender, priming, and the collocates that we see would have an effect.

Did their effect stay constant over time or did they change? We hypothesize that there would be changes with regard to the following segment. That effect looked like it would be time-sensitive. Also, the effects of stress and collocate

frequency, which were non-uniform across time, would lead us to expect this kind of interaction effect.

Did *myne* and *thine* change in the same way, at the same time? There are some minor differences between *myne* and *thine* with regard to formality, gender and priming. I fitted a logistic regression model to the data that tests for the respective main effects and interactions.

The statistical modeling allows us to determine which factors make a difference at what time. The method I used is known as binary logistic regression. That is a technique that can be used to investigate the variables that influence a binary choice either in an experiment or in natural language use from corpora. In language, binary choices can be seen as two ways of saying the same thing. Let me take an example that I haven't used so far. A phonological binary choice in English is exemplified by yod-dropping. Speakers can say "*She is such a good* /stju:dənt/" with a /j:/ glide in it, or say "*She is such a good* /stu:dənt/" without that glide. Another binary choice would be the alternation between *be going to* and *will*, for example *I do not know what I am going to do* vs. *I do not know what I will do*. There is the dative alternation, i.e. *John gave Mary the book* and *John gave the book to Mary*.

What do you analyze in such a design? The basic question, in all of these cases, would be whether we can explain why a speaker sometimes does one thing and sometimes another. What are the variables that can explain this behavior? The explanatory variables can be linguistic (language-internal) or they can be non-linguistic (language-external).

With regard to yod-dropping, your region of residence plays a role. Your age might play a role. Your gender or your level of education could come in. All of these are language-external factors. For the dative alternation, i.e. *John gave Mary the book* and *John gave the book to Mary*, there is a host of factors that have been studied that contribute to the alternation. Pronominality is one such factor. Whether the recipient is expressed with a pronoun, *John gave her the book*, or a full noun phrase, *John gave the woman the book*, plays an important role. If the recipient is pronominal, then there is a preference for the ditransitive, as opposed to the prepositional dative. There are even some hard constraints, such as the animacy of the recipient. I can say "*John threw his keys to the floor*". I cannot say "*John threw the floor his keys*". That would imply that the floor is animate and is waiting to catch the keys that John is throwing. Inanimate recipients bias speakers very strongly and consistently towards the prepositional dative. Finally, the definiteness of the theme is important. Whether I have an indefinite noun phrase like "*John gave Mary a letter*" or a definite noun phrase "*John gave Mary the letter*", that makes a difference, not that big of a difference, but a measurable difference. Indefinite noun phrases are preferably placed at

the end. That is why there is a preference for the ditransitive. With regard to the dative alternation, we have these language-internal factors of indefiniteness, animacy, and pronominality . All of those are language-internal, concerning morphology, semantics, and pragmatics, respectively.

Similar to the alternations I just presented, the case between *myne* and *my* features one linguistic variable which we call the dependent variable, which is binary and categorical, that is, a choice between A and B. A and B are called the "levels" of the dependent variable. Then there are several explanatory factors, language-internal and language-external. They can be categorical, such as accusative vs. dative, or pronominal vs. nominal. They can also be continuous, as in length in words or word frequency. All of these variables can be integrated into the analysis.

You employ a binary logistic regression generally in order to obtain answers to the following questions: Which explanatory factors have an influence on the dependent variable? How strong are the respective variables in their influence? How well do the variables explain the variation that we observe?

What kind of data can enter the analysis? In corpus-based research, you would have to retrieve all occurrences of language use in a given corpus where speakers actually have the choice between A and B. For the case of the /stju:dənt/ and /stu:dənt/ alternation, you would have to find all words where /u:/ can be preceded by a glide /j/. With the dative alternation, you would have to find all instances of the ditransitive and all instances of the prepositional dative.

Coming back to *my* and *myne*, here we would collect all instances in which writers use a possessive determiner and have the choice between using either the conservative n-variant or the progressive n-less variant. What does the regression do with that data? It calculates a formula that allows you to predict for new examples, from a different corpus, how likely it will be that speakers choose the n-less variant rather than the n-variant. If in a large database of examples, we find an example such as *the daies of my life*, then how likely is the n-less variant, given its date of production, given its phonological environment, given the gender of the writer and given the formality or informality of the genre?

There is one complication that I need to discuss. Sometimes variables may interact. Several explanatory factors can conspire with each other, which yields an effect that is different from what we would call a simple main effect. A main effect obtains when an explanatory factor always has the same effect on the choice between A and B. In an interaction effect, by contrast, an explanatory factor has an effect on the choice between A and B, but this effect depends on another variable.

Let me give you an example. We're interested in speakers' use of polite or direct pronouns in a corpus. The dependent variable would be a binary choice: Does the speaker use of a polite pronoun or a direct pronoun? The explanatory factors that we take into account are age, gender and formality. A main effect of gender would be that female speakers have a bias towards the use of polite pronouns. An interaction effect would be that age and gender conspire. If gender interacts with age, that means that old female speakers are biased towards polite pronouns, but young female speakers are not affected in the same way.

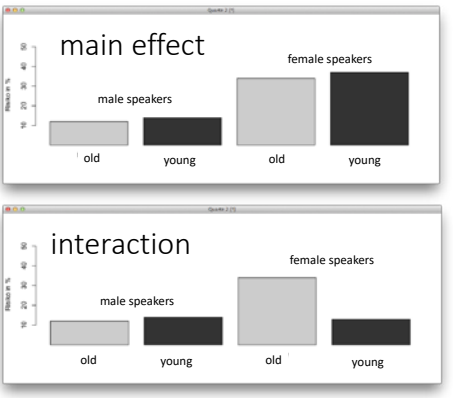
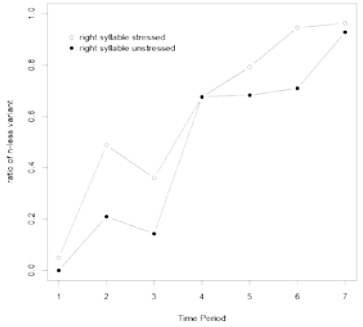


FIGURE 16

I try to visualize that on this slide. The upper graph shows a main effect. You compare male and female speakers across two different age groups. It turns out that male speakers pattern together. Age does not play a role, as young and old males are alike, and young and old females are alike. This would be a main effect of gender. An interaction effect would be if old female speakers have a preference for polite pronouns, but young female speakers do not, and they pattern exactly like the old and young male speakers.

This graph shows a potential interaction effect in the use of *myne* and *my*. The interacting variables are stress and time. When we look at this graph, right stress has an effect, but sometimes it is strong, sometimes it is not. We should therefore test whether there is an interaction between time and stress.

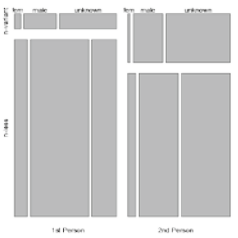
Here I am coming back to one of the mosaic graphs with gender. Gender has an effect, so that females use more n-less forms. They are more progressive, but we might ask whether this effect is equally strong for first person and second person forms. We could test whether there is an interaction between gender and person.



Right stress obviously has an effect, but sometimes it is strong, sometimes not.

It should be tested whether the interaction between **Time** and **Stress** is significant.

FIGURE 17



Gender probably has an effect (females use more n-less forms), but is this effect equally strong for 1st and 2nd person forms?

It should be tested whether the interaction between **Gender** and **Person** is significant.

FIGURE 18

Which factors make a difference at what time? I fitted a regression model in which I tested for main effects of all of the explanatory factors that I mentioned before. I tested for a selected range of interaction effects. For time and the following segment, we can be fairly sure that there is an interaction, because we observed that the effect of the phonological environment was time-sensitive. For time and stress, we saw that the effect of stress is not uniform across time. The same goes for time and collocate frequency. I also tested for interaction effects between person and formality, person and priming, and person and gender. Since I used mixed-effects regression modeling, I was able to include the authors and the respective corpus files as random factors.

What came out? There are main effects for time, for the following segment, for stress, for priming, for gender and for collocate frequency, meaning that only two variables that were included originally turned out to be non-significant, and those are person and formality. The non-significance of person means that there is no evidence for a separate generalization for second and

first person possessive determiners. The non-significance of formality means that the use of possessive determiners is not reliably different across formal and informal texts. I ran another analysis in which I excluded those variables that did not have a significant effect.

The revised analysis obtains significant effects for the variables that mattered already the first time around. There are several interaction effects that the model included as significant. There is an interaction between time and stress, so that the effect of stress becomes weaker over time. There is an interaction of time and collocate frequency. The effect of collocate frequency is significant but unstable. For the interaction between time and the following segment, a following consonant always increases the likelihood of the n-less variant, but the effect is strongest in the early periods. The technical indicators tell us that model provides a useful summary of the data.

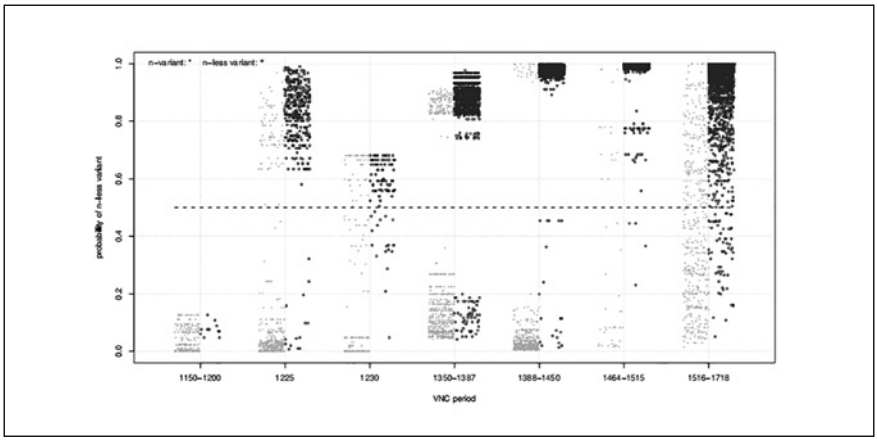


FIGURE 19

I am sorry that I am bombarding you with strange-looking graphs in this lecture, but I would like to give you another quick look under the hood of the regression model. What you see in this graph shows you how well or how poorly the model discriminates between examples of the n-variant and the n-less variant.

The grey little specks you can see on this slide are examples of the n-variant. The graph also shows examples of the n-less variant, the new variant. Those are shown as black little circles. The graph has a *y*-axis. If a data point appears up high on the y-axis, that means that the model returns a strong prediction that we are looking at a modern n-less variant. If the model places a data point further down, near zero, that means that the model predicts an n-variant. Ideally, we would have all the grey spots down in the lower half, and we would have all

the black circles in the upper half. Ideally, we would want a very crisp and clear distribution that this half down here is all grey, that half up there is all black. What you see is that this is not 100% the case.

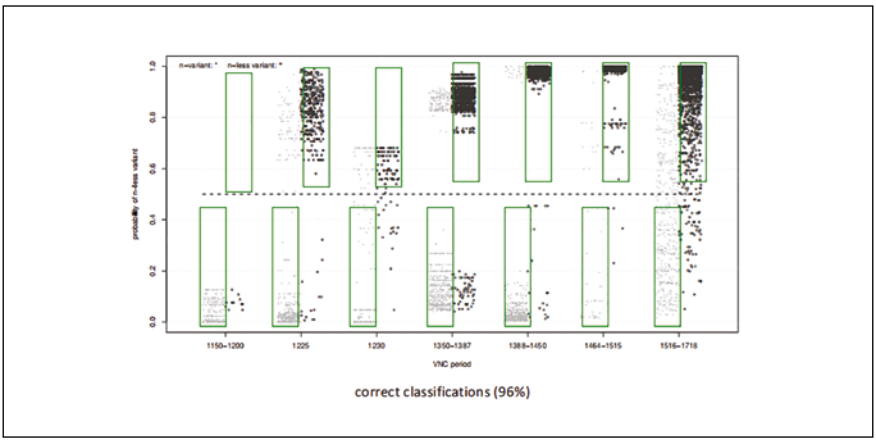


FIGURE 20

To rephrase this, in this version of the graph, I have put all of the correct predictions that the model makes into green boxes. If we have dots in a green box, that means the model has made the right assessment. For example, in the last period, the model predicts for the overwhelming majority of n-less forms correctly that they are n-less. There are only a few in the lower half that are misclassifications. That is actually the reason why I want to show you this graph. We want to figure out where and why the model fails. Models are imperfect, and we want to learn a little bit about why and when they fail.

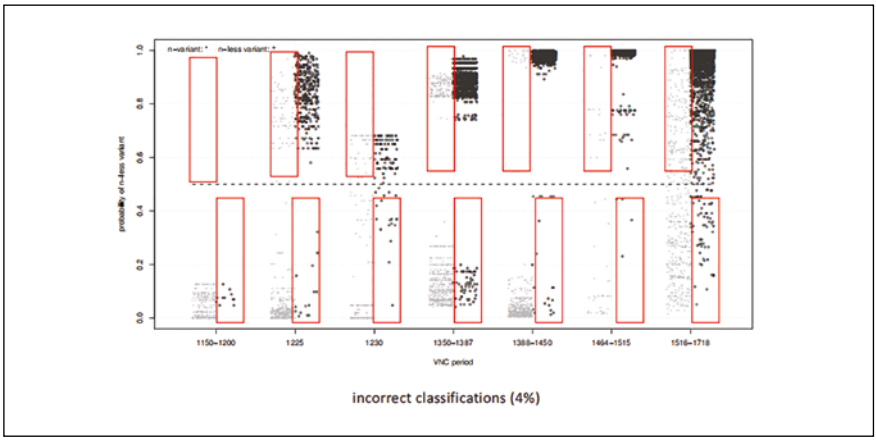


FIGURE 21

This slide shows the misclassifications in red boxes. In the first period, there are a few modern forms, but the model misclassifies them and gives very confident predictions that these are n-forms. In the second period, we observe the same thing, but notice that there are also many n-forms which the model classifies mistakenly as n-less, that is what you see in the red box in the upper half. Going further, we see that there are lots of forms that are misclassified. But globally, the model makes the right predictions most of the time. Let's look at some of the misclassifications in more detail.

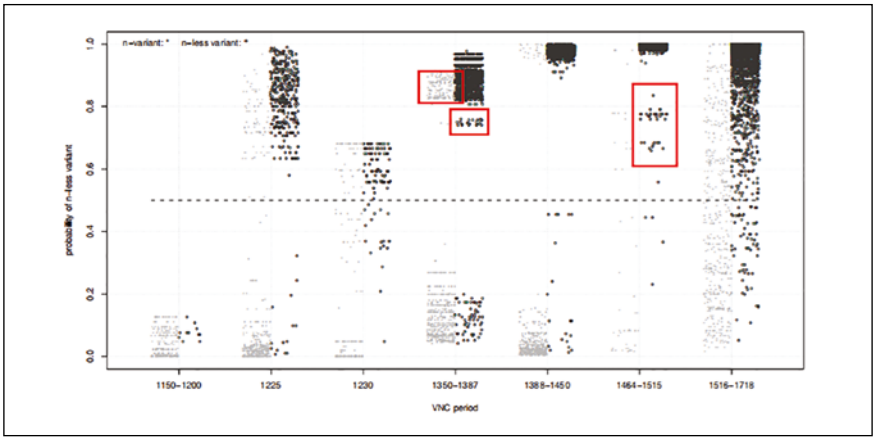


FIGURE 22

For example, what is going on here in the fourth period? There is a set of examples classified as n-less, but they do in fact have an n. It turns out that in this period, the model overestimates the discriminative power of the following segment. If a form is followed by a word beginning in a consonant, such as *my brother* or *my friend*, then the form is given a high likelihood estimate for the modern variant. This can be considered as an honest mistake that the model makes. The following segment used to be a very powerful discriminant variable in the earlier periods, but in the fourth period, that importance is decreasing. The factor is becoming less predictive, less strong.

If we are looking at the sixth period, there are a number of examples that the model correctly classifies as n-less, but not very confidently so. The red box in the sixth period shows correct classifications, but the model is less confident of that than with other data points that are further up. What characterizes the examples in the red box? Those are examples before vowels and H'es. These phonological environments are predictive of an n-form, so the model classifies them as n-less with less confidence. Similarly, the black circles in the fourth

period that are in the red box, they are not misclassifications, but again the model's predictions are less confident. These examples are not followed by a stressed syllable, as for instance *my confession*.

To make a long story short, this graph allows you to look qualitatively at the statistical analysis, which can otherwise be a black box. Statistics can and should be a tool to guide you towards the examples that are relevant and that you want to look at more closely. The method further allows you to perform cross-validation. The main point of that is to verify how good the analysis is by confronting it with unseen data from a different corpus. In this case, the model actually succeeds in classifying that data rather well. The resulting analysis yields the same significant predictors and interactions.

Summing up, I think that there are issues of theoretical interest in Construction Grammar that can be analyzed in a very satisfying way with methodologies that we borrow from frameworks such as variationist sociolinguistics. In this particular case, I was after the question of whether two forms develop in identical or different ways. I hope to have shown that the similarities between *myne* and *thine* greatly outweigh the differences. All in all, that would suggest that their development constitutes one single constructional change.

The change from *myne* to *my* is a case of constructional competition that has gone to completion. The propagation of the new variant has led to a complete substitution of the old variant. This kind of propagation can be shown to relate to different factors, both language-internal and language-external. We have a very good idea of what caused competition at various points in time. The quantitative techniques that I have applied here can help us uncover how propagation actually proceeds. Different factors influence competition at different times with different strengths. That is the conclusion that I would like to leave you with today. Thank you for your attention.

LECTURE 7

Differentiation and Attraction in Constructional Change

Welcome back to these Ten Lectures on Diachronic Construction Grammar. In my last lecture yesterday, I have gone into the topic of what happens diachronically with constructions that can be seen as alternatives to each other. In certain cases, constructions are in mutual competition, and over time, one member of an alternating pair of constructions can replace another one.

I have said that in cases of this kind, we typically have variation that includes social factors. I have mentioned that Bill Croft (2000: 166) has argued for a view of propagation that is socially motivated, so that language-external factors play an important role. Today, I will be looking at another phenomenon in which several constructions act as alternatives to each other. The topic for this lecture is "differentiation and attraction in constructional change". We'll be looking at a set of constructions that stand in a paradigmatic relation and that speakers can choose from for the expression of a given idea. We will see how these constructions develop over time.

In contrast to yesterday, this will not be a story of one construction winning out over another one. Rather, we will see that some constructions become more different from one another, and others are attracted to one another and become more similar. As in the previous studies before, it is very useful to think about these developments in terms of links between constructions. On the node-centered view of constructions that still prevails in the field, we would simply say that over time there is change in the features of the constructions that are inscribed in the nodes. As two constructions are differentiating more and more, their features become maximally different over time. That is a valid way of thinking about it, but I would like to suggest that there is a much more natural way to think about it, namely in terms of links.

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691182

Let's say that there are two forms, A and B, and each of them is connected with a range of meanings. Some of these meanings overlap. There are some meanings that are specific to A, and some meanings that are specific to B, and some overlapping meanings can be served by both constructions. When the speaker thinks of a meaning that lies in a non-overlapping area, there is no choice. The speaker has to pick the construction that expresses that non-overlapping meaning. Every time they do that, the connection between the form and the non-overlapping meaning is strengthened, and the connection between the form and the overlapping meaning is punished, so that it is weakened a little bit. This is a natural explanation for why over time we see cases of semantic differentiation.

A constructional, link-based explanation thus supports the view that linguistic units can be in competition, and that this competition either leads to substitution or differentiation. This is the idea that De Smet and colleagues (2018: 197) have examined in their paper that I mentioned yesterday. Across different theoretical frameworks, differentiation can be viewed as the default consequence of constructions or linguistic forms being in competition. However, De Smet and colleagues notice that competition can also have the opposite effect. There can be attraction. That process is driven by analogy. Let me read a quote from De Smet et al. (2018: 197): "*As a result of analogy, competing forms often show attraction, becoming functionally more* (*instead of less*) *alike*". That is a somewhat paradoxical situation. When we have multiple constructions that are paradigmatically related, and they can become either more different or more similar. *Attraction can maintain and increase functional overlap in language*.

This relates to an idea that I mentioned earlier in Lecture #2, namely the grammaticalization process of paradigmatization. When a new grammatical construction emerges, it tends to join an existing paradigm, and it tends to adapt its behavior to that paradigm. Let me give just two examples from the history of English.

Indefinite determiners like *a* and *an* derive from a numeral word, *one*. The numeral *one* did not always belong to the paradigm of English articles. English did not always have an indefinite article, but in synchronic language use, the indefinite article has become a part of the article paradigm of English, contrasting with the definite determiner *the*. It has formed a little group of elements that serve the same grammatical function and do so by contrasting with each other.

The second example is more recent, i.e. the discourse marker *now*. Often when you will hear me say "*Now, there is this problem*", that *now* does not mean literally *at this moment*, but it indicates a shift of topic. The discourse marker

now developed out of an adverb meaning *at this time*, and in present-day language use, the discourse marker *now* behaves like other discourse markers like *well*, or *you know*, or *okay*, that I could use alternatively when I am coming to a new topic. I could say, "*Okay, now let's move on to this*".

The underlying question is why constructions arrange themselves in paradigms. One explanation that has been proposed is that paradigms can be viewed as constructions themselves, as generalizations over generalizations.

This brings me back to Controversy #2 from the first lecture. There are researchers that argue for constructions at a very high level of abstraction, so called higher-order schemas, in which speakers represent generalizations across abstract patterns. One example of this would be the dative alternation. If speakers form a generalization across *John gave Mary the book* and *John gave the book to Mary*, then they have a generalization that essentially constitutes a paradigm, a small paradigm in this case, as there are only two members. My main question for this lecture is the following. How do differentiation and attraction work in cases where we have paradigms of constructions?

In order to investigate that question, I brought another case study from my 2013 book, in which I look at a syntactic clausal pattern. The pattern involves a paradigm of forms that I call concessive parentheticals. They are illustrated by examples such as the ones that you see here, (1) *Power, while important, is not everything*, (2) *It is an earnest, if unsophisticated, film*, and (3) *Although a Democrat, he has strong Republican support*.

In this lecture, I want to introduce you to these constructions. I want to show how they have developed over the past 150 years, and I want to demonstrate how the analysis of these constructions speaks to the issue of these abstract representations of groups of constructions as a higher-order schema.

Concessive parentheticals can be defined in terms of several features. First of all, they contain a concessive subordination conjunction, which can be instantiated by elements such as *while* in "*Power, while important, is not everything*", *if* in "*It is an earnest, if unsophisticated, film*", and *although* in "*Although a Democrat, he has strong Republican support*". Semantically, concessive clause relations cancel or reject a potential implicature or a potential conclusion that someone might draw. The example "*Although a Democrat, he has strong Republican support*" expresses that one might think that a Democrat will not be supported by Republicans, but that this conclusion is not valid in this case.

Second, concessive parentheticals have a predicative element. There are these words like *important*, *unsophisticated*, or in the third example, *a Democrat*, which are qualities that are predicated over an entity. The phrase *although a Democrat* conveys that he is a democrat. That is a predication.

The third feature is syntactic. Concessive parentheticals are embedded into a superordinate syntactic matrix structure. They have a host clause in which they appear. For example, the concessive parenthetical *while important* occurs inside the sentence *Power is not everything*.

That embedding explains why these constructions are parentheticals. A parenthetical is a linguistic structure that is put into brackets, that functions as an afterthought, as something that you could also prefix with *besides*. You'll often see me doing a gesture that indicates that I'm putting my words into brackets. In writing and orthography, there are often commas, and in fact this is sometimes called comma intonation. "*Power, while important, is not everything*". You hear me saying this with little pauses. In the last two lectures, we have seen that variation is crucial for constructional change, and that is no different in this lecture.

Concessive parentheticals exhibit variation with regard to the features I just mentioned. There is variation in the conjunctions. There are four different ones that can be used and that I will be studying, *although*, *though*, *while* and *if*.

There is the further variation in the positions in which these parentheticals can appear. They are not always in the middle. They can appear at the very front of an utterance. *Although unorthodox, the logic here is simple*. If you put it in the middle, it goes *The logic here, although an orthodox, is simple*. It can even appear at the end, *The logic here is simple, although unorthodox*. With regard to their position, these constructions are rather flexible.

There is a massive variation with regard to the syntactic categories of this predicative element that I have mentioned, ranging over adjectives as *unorthodox*, adverbs as *reluctantly*, noun phrases like *a Democrat*, prepositional phrases, past participles, and even entire clauses. There is a wide range of syntactic variation.

Finally, there is another type of syntactic variation that concerns the embedding of the concessive parentheticals in their matrix structure. Concessive parentheticals can be embedded at the level of the clause. In the example "*Power, while important, is not everything*", *while important* is embedded in the clause *Power is not everything*. But there is a second way in which concessive parentheticals can be embedded, which would be at the level of a phrase, for example, a noun phrase. In "*It is an earnest, if unsophisticated, film*", the parenthetical appears inside a noun phrase: *an earnest, if unsophisticated, film* is a heavy noun phrase.

In summary, concessive parentheticals are headed by a concessive subordinating conjunction, they involve a predicative element, they are hosted by a matrix clause, and they vary with respect to their conjunction, their relative position in the clause and their internal syntactic structure.

My first empirical question concerns the emergence of concessive parentheticals. How did concessive parentheticals constructionalize in the sense of Traugott and Trousdale (2013)? There are two possible hypotheses that I would like to explore. Hypothesis A would be that concessive parentheticals came about through a process I have mentioned before, namely reduction. Reduction commonly occurs in grammaticalization. Phonological reduction is a frequent phenomenon, but reduction can also be observed on the syntactic level. Hypothesis A would state that concessive parentheticals have come about through reduction of a full concessive clause. Some of the words are left out, and we are left with a more compact structure.

Hypothesis B would posit that concessive parentheticals came about through analogy from other clause types. For example, if we take temporal subordinate clauses, we know that they have reduced variants as well. In English, you can say things like "*While young, swans are actually grey*". I do not know if you've ever seen baby swans. When they hatch, they look like grey ducks, and then they grow and shed their grey feathers, and they end up white. The phrase "*while young*" thus means "during the time that they are young". Since this exists as a temporal subordinate clause structure, why not simply analogize that and form concessive clause that work more or less in the same way? *While young, Reed is rated as a top lawyer*, that does not mean "While he is young, Reed is a good lawyer", but rather "although he is young". You expect young lawyers to be less experienced and not as good. The example counters this conclusion.

Both reduction and analogy are recognized as forces that shape language change. With regard to parenthetical structures, reduction is more commonly invoked than analogy. If we have an elliptical phrase pattern, it is quite natural to think that it is the result of reduction, where speakers cut corners and shave down an expression to a more economic form.

Reduction is of course ubiquitous in language change. This has been expressed, for example in Givón's (1971) slogan "Today's morphology is yesterday's syntax". But crucially, reduction is not the only game in town. Fischer (2007) has shown that the default assumption of reduction processes in syntax can lead to faulty conclusions. Brinton (2008) comes to the same conclusion. Brinton (2008) has investigated parentheticals in English, structures such as *I mean*, *I find*, or *you see*, and she finds that diachronic corpus data lends little to no empirical support to the reduction hypothesis. That means with regard to concessive parentheticals, we should also be very careful before we assume reduction as the most plausible explanation.

There are further problems for reduction, because there are some examples of concessive parentheticals that you simply cannot expand into a full clause. Let me just present one example here. *Hood is a seasoned though disillusioned*

diplomat. It is not possible to expand that example into **Hood is a seasoned, though he is a disillusioned diplomat*. That yields an ungrammatical sentence. Concessive parentheticals that are embedded at the phrase level do not allow an expansion of this kind.

How can we find out whether reduction or analogy is more likely? Again, I would like to use corpus data to examine that question. I want to operationalize the structural features of these constructions in such a way that their corpus frequencies can inform the controversy between reduction and analogy, and I would like to weigh the relative likelihoods of the two scenarios.

For the analysis, I used the TIME corpus of American English, which consists of journalistic writing. Why did I do that? Concessive parenthetical constructions are at home in elegant journalistic writing. It is a very compact structure. It is one that allows you to convey a lot of content in relatively little space, and that is what journalists need to do. My data comprises two concessive conjunctions, namely *although* and *though*. I also investigated two polysemous conjunctions namely, *if* and *while*. The main function of *if* is its use as a conditional conjunction. Its concessive function is merely secondary. The same is true for *while*, which is mainly a temporal conjunction, not a concessive one. I took random samples of 5000 concordance lines for each conjunction and then manually identified target examples. Most of the examples for *if* and *while* were of course conditional and temporal, not concessive.

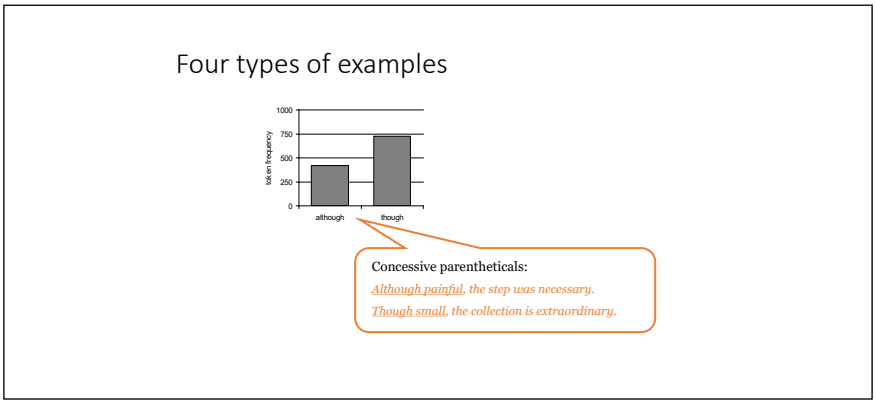


FIGURE 1

Let me talk about the kind of data I was working with. Overall there are four different types of example. The first type includes concessive parentheticals with the conjunctions *although* and *though*. In this graph that you see that I obtained about 400 examples of concessive parentheticals with *although*, such as, "*Although painful, the step was necessary*", or "*Though small, the collection is extraordinary*". You see the token frequencies of those constructions.

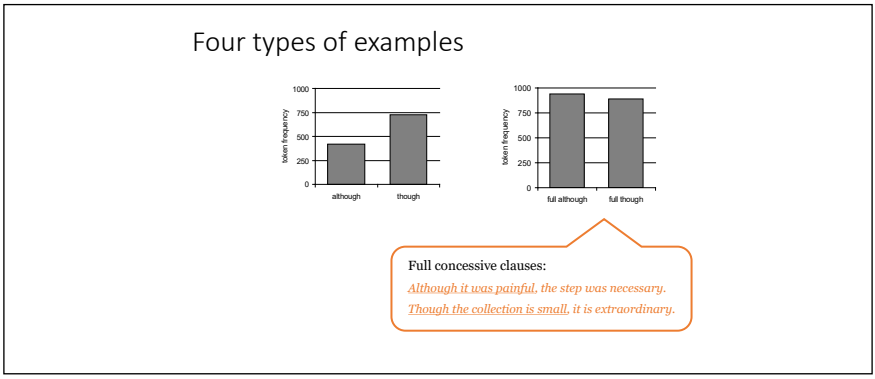


FIGURE 2

The second type are full concessive clauses with *although* and *though*. I obtained about 800 examples for each category with *although* being illustrated by examples such as "*Although it was painful, the step was necessary*", or "*Though the collection is small, it is extraordinary*".

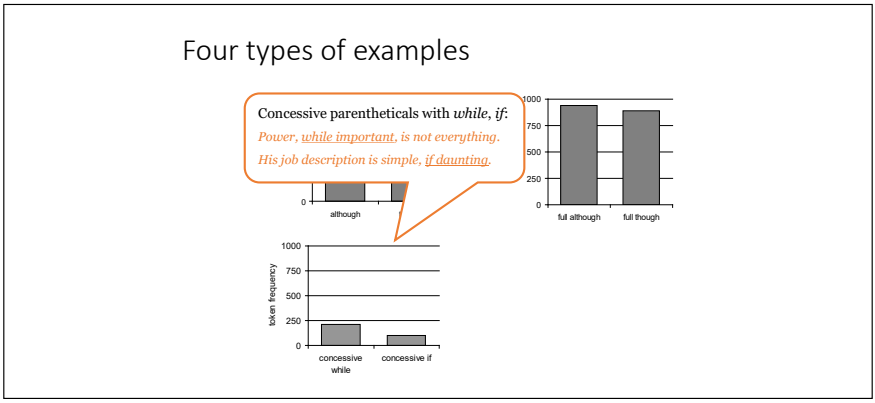


FIGURE 3

Moving on to the third type, here we have concessive parentheticals with two other conjunctions, namely *while* and *if*. You see that these types are less frequent. They are illustrated by "*Power, while important, is not everything*" or "*His job description is simple, if daunting*".

The fourth and final type are temporal and conditional parentheticals with *while* and *if*, which are again quite frequent. Temporal *while*-parentheticals are illustrated by "*He was driving, while under the influence*". A conditional example with *if* would be "*If possible, patients should be treated at home*". These are the four sources of data that I used. Why these four types? That relates to the two hypotheses that I wanted to test.

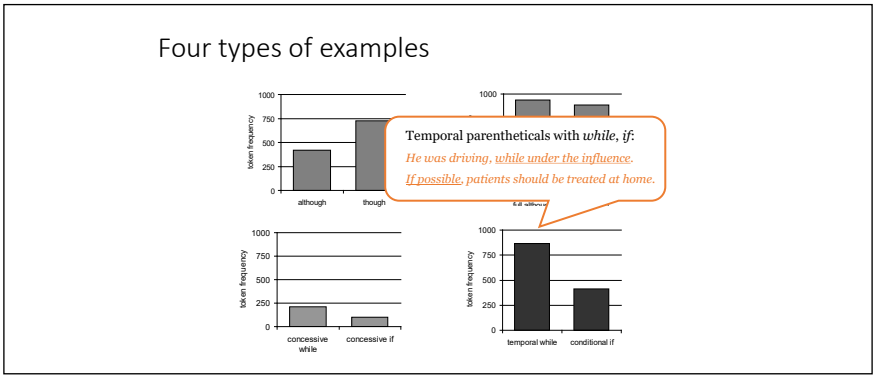


FIGURE 4

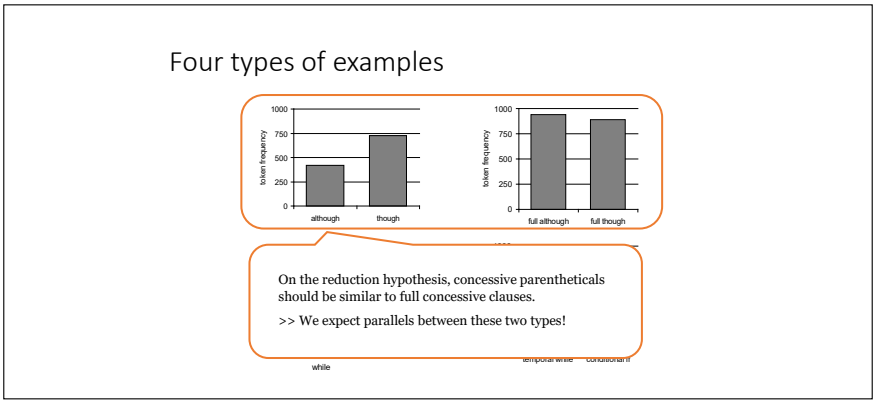


FIGURE 5

On the reduction hypothesis, concessive parentheticals derive from full concessive clauses, so these two types should be similar to each other. We expect parallels between these two in terms of their structural features and in terms of the lexical items that occur within them.

By contrast, on the analogy hypothesis, concessive parentheticals should be similar to temporal or conditional parentheticals. We would expect parallels between these two types. Concessive parentheticals with *while* should be similar to temporal parentheticals with *while*. Concessive parentheticals with *if* should be similar to conditional parentheticals with *if*.

I annotated all examples in the database for the relative position of the subordinate clause. Is the subordinate clause initial, medial, or final? I also annotated the syntactic structure of the predicative element in the subordinate clause. Is that predicative element an adjective, a noun or something else? I further annotated the examples for their lexical types. This means that I noted the adjective *small* in *although small*, or the adjective *necessary* in *if necessary*.

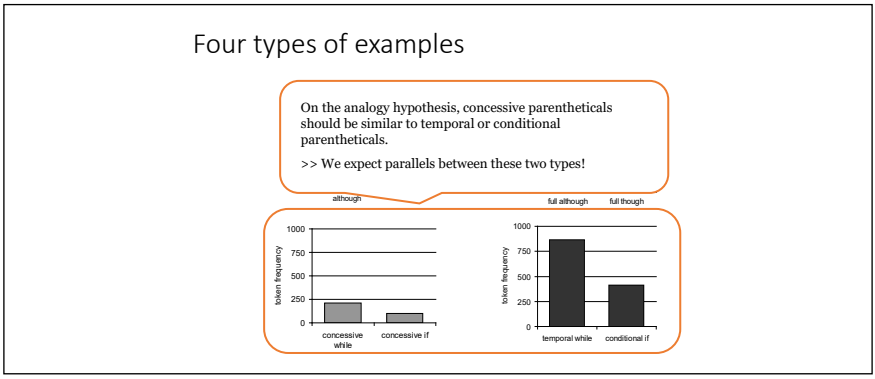


FIGURE 6

In order to perform quantitative comparisons, I used chi-squared tests of independence. Let's talk about *while* first.

On the analogy hypothesis, temporal parentheticals give rise to concessive parentheticals, and I mentioned the parallelism between "*While young, swans are actually grey*" and "*While young, Reed is rated as a top lawyer*". This would predict that temporal and concessive examples with *while* would be similar in terms of their preferences for their position, syntactic structures and lexical collocates.

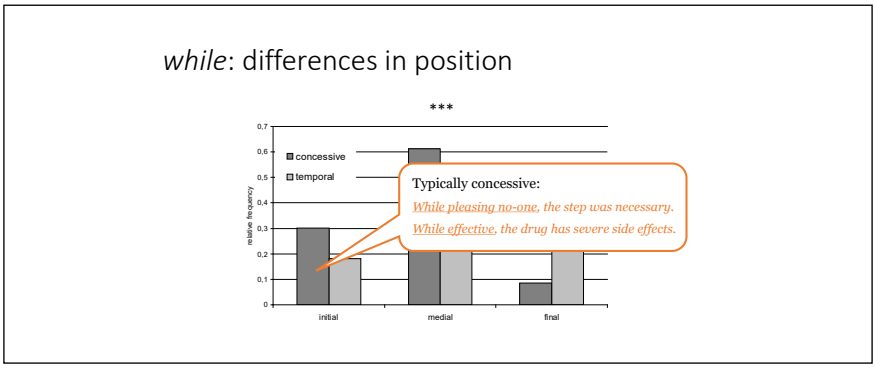


FIGURE 7

Is this the case? The short answer is no. When we compare temporal and concessive uses of *while*, there are significant differences with regard to their syntactic positions. Parentheticals with initial *while* are typically concessive, as in *While pleasing no-one, the step was necessary* or *While effective, the drug has severe side effects*.

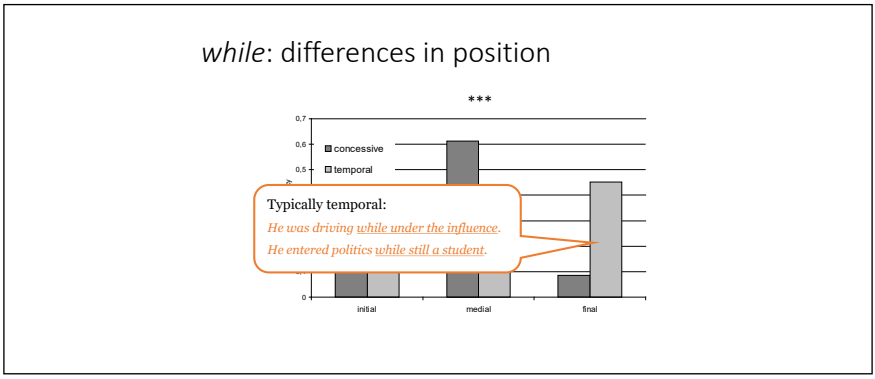


FIGURE 8

If we look at examples with final *while*, the relative preferences are very different. Final parentheticals typically have temporal meaning, as in *He was driving, while under the influence* or *He entered politics while still a student*.

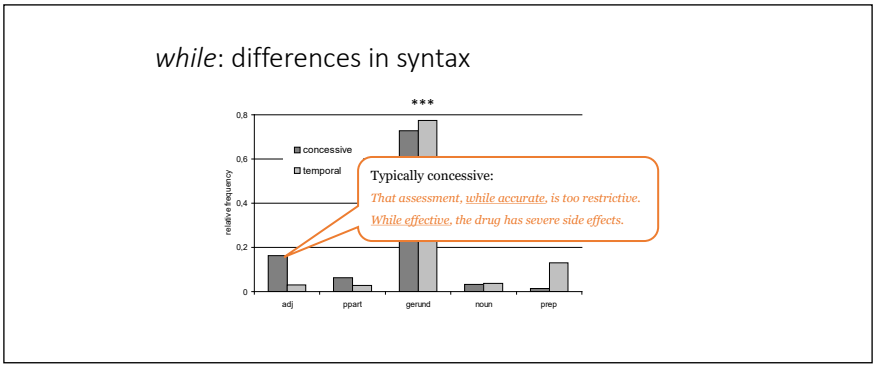


FIGURE 9

I observed further differences with regard to the syntax of the predicative element. In the category of adjectives, there is a relative preference for concessive meaning. This is illustrated by examples such as *That assessment, while accurate, is too restrictive* or *While effective, the drug has severe side effects*.

At the other end of the graph, there are prepositional phrases, which are preferentially used with temporal meaning, as in *He was driving while under the influence* or *He met her while at Harvard*.

I further examined the lexical words that appear in these constructions. For temporal and concessive *while*, there is practically no collocational overlap. They do not occur with the same word types. Concessive *while* occurs with adjectives such as *accurate*, or *agreeable*, or *beautiful*. Temporal *while* is used

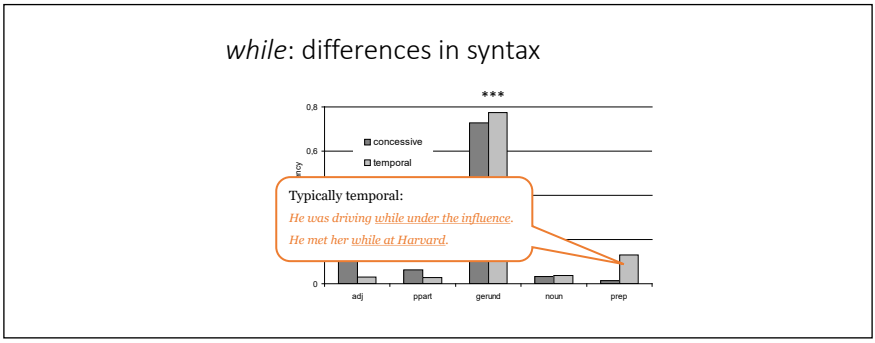


FIGURE 10

with adjectives such as *alive*, *asleep*, or *drunk*. There is no overlap. That suggests that it is rather unlikely that one developed out of the other. If we had this natural relation between them, we would expect some remnants of overlap.

In summary, it is unlikely that concessive parentheticals with *while* emerged as an analogy to temporal *while*-clauses, because they differ in terms of their syntactic placement, the structure of the predicative element, and their collocates. That is a negative result for the analogy hypothesis.

Let's look at *if*. On the analogy hypothesis, concessive parentheticals with *if* should derive from conditional parentheticals. Conditional parentheticals such as "*If successful in the semi-final, they will play against the Netherlands*", should give rise to concessive parentheticals, such as "*If successful in the semi-final, they still lost against the Netherlands*". That last example means that although they were successful in the semi-final, they still lost against Netherlands. As before, with *while* we observe differences rather than similarities with regard to the same parameters.

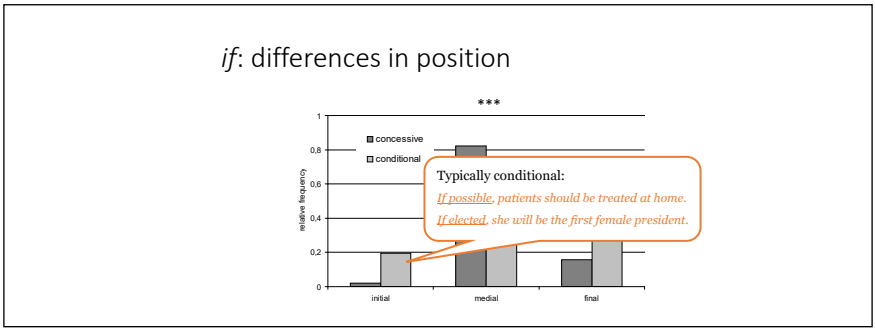


FIGURE 11

Here you see the differences in syntactic position. The graph shows differences in each category, initial, medial, and final. Let me just talk about the initial position here. Initial *if* is typically used with conditional meaning, as in "*If possible, patients should be treated at home*" and "*If elected, she will be the first female president*".

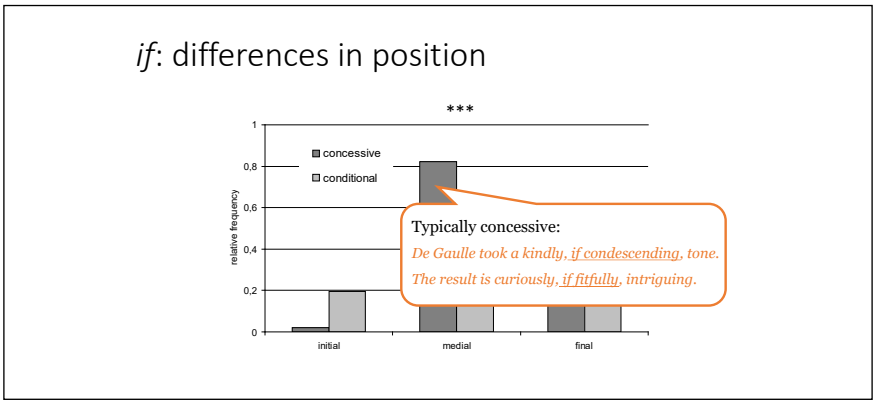


FIGURE 12

By contrast, medial parentheticals are typically concessive, at least they have a much higher preference for concessive meaning. Let me just take the second example, "*The result is curiously, if fitfully, intriguing*".

Let's look at the predicative elements. Also here, the pairs of bars have very different heights. Especially past participles are a strong cue for conditional meaning, as in *He will face a life sentence, if convicted*, or *If elected, she will be the first female president*.

By contrast, *-ing* clauses are typically concessive. An example would be *His job description is simple, if daunting*.

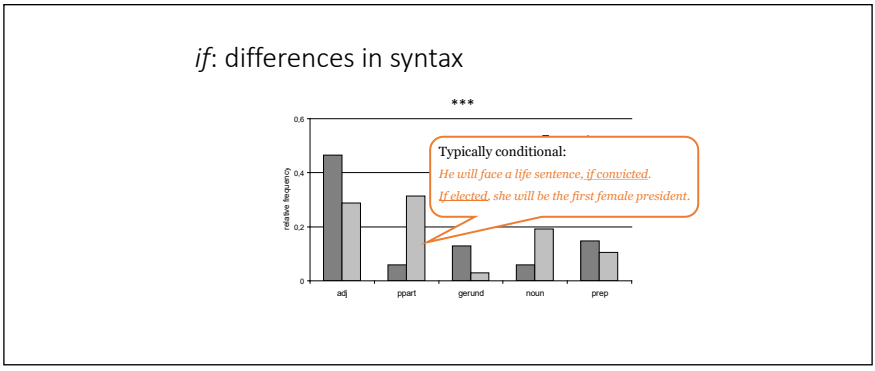


FIGURE 13

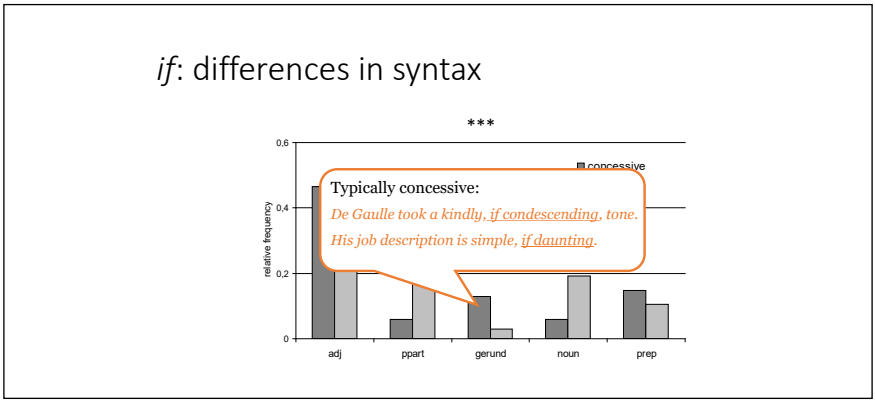


FIGURE 14

In terms of collocates, we see almost no overlap between conditional *if* and concessive *if*. There is one exception to that tendency, namely the adjective *accurate*. In *If accurate, the results are sensational*, the meaning is conditional. On the condition that his results are accurate, then this is great news. But in the second example, we have concessive meaning: *It was an amusing, if not fully accurate, report of his activities*. Besides that, concessive *if* and conditional *if* occur with very different sets of adjectives.

In conclusion, the evidence suggests that concessive parentheticals with *if* did not emerge as an analogy to conditional *if*-clauses, because there are simply too many differences across the structural and collocational parameters.

That leaves *although* and *though*. Let's move on to them. Here the data allows us to test the reduction hypothesis, which would predict that concessive parentheticals derive from full concessive clauses. Examples such as *Although he is a Democrat, he has strong Republican support* should, over time, give rise to *Although a Democrat, he has strong Republican support*. More specifically, full and parenthetical examples with *although* and *though* should be similar in terms of their syntactic structures and the lexical collocates. That is what we are going to look at next.

This slide presents the syntactic properties of full clauses in light grey and parentheticals in dark grey. You see that there are some differences, but overall the pairs of bars are more in line with each other than in the graphs that we saw with *if* and *while*.

This graph shows the syntactic categories for *though*. Again, the results are fairly similar. A chi-squared test yields the result that the differences are significant, but in comparison with *if* and *while*, these differences are minor.

In terms of collocates, there is much more overlap between full clauses and parentheticals than what we observed earlier with *if* and *while*, namely

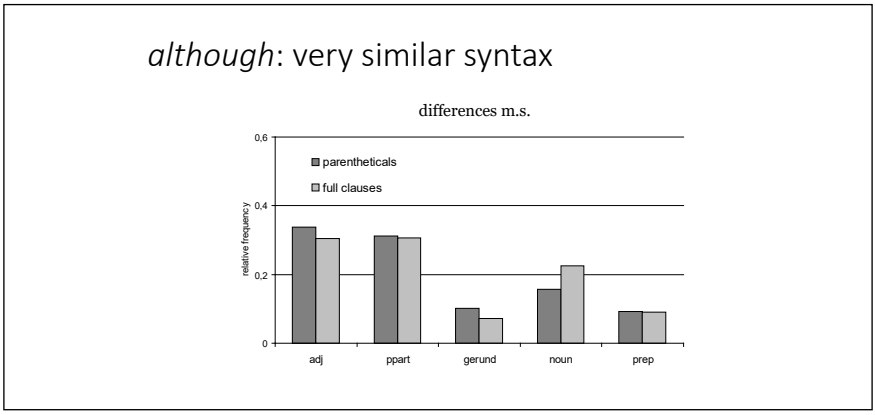


FIGURE 15

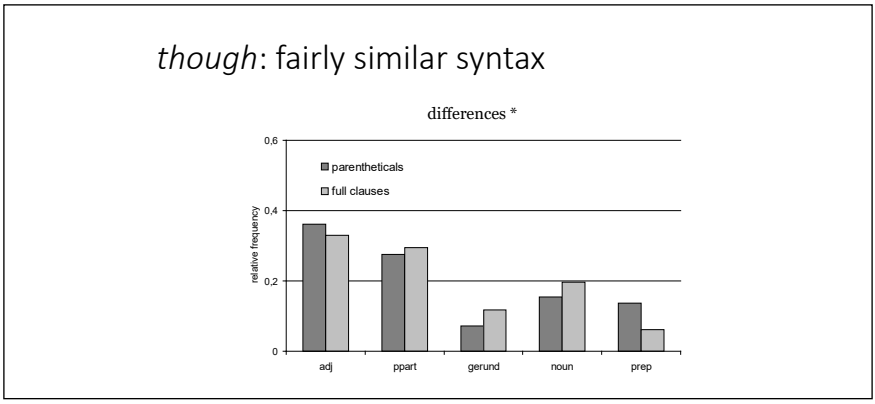


FIGURE 16

about 10.5%. You might say that this is not very much. That is true, but it is definitely more than what we saw earlier. In conclusion, the corpus data are largely congruent with the reduction hypothesis. Concessive parentheticals and full concessive clauses have relatively similar syntactic structures and they have some collocational overlap. When we compare *though* and *although*, it seems that *though*-parentheticals are a little bit more different from their full counterparts than this is the case for *although*. They seem to lead the way in this differentiation process. If you imagine that we have a paradigm that is emerging, with every construction variant trying to find its niche where it serves a particular function, then *though* is in the vanguard, and it has emancipated itself to a stronger degree from the full clauses than its counterpart with *although*.

Let me come to a close with regard to the issue of reduction or analogy. In the analysis of parenthetical, elliptical, or otherwise apparently reduced

structures, the process of syntactic reduction should not be assumed as a default. I take that idea from Fischer (2007) and Brinton (2008). However, in the case of concessive parentheticals with *though* and *although*, the reduction hypothesis holds up rather well. We have some structural evidence, and we have collocational evidence. All of that makes me confident to posit a relation between full and parenthetical clauses. The alternative hypothesis of analogical change from temporal and conditional parentheticals is not at all supported by the corpus evidence.

I'd like to come back to De Smet and colleagues (2018) and their observation that analogy has an important part to play in the mutual attraction of constructions. I want to do so by focusing on the question of whether concessive parentheticals become more similar or more differentiated over time. I have mentioned the idea that we observe a paradigm that comes into being, and it merits some consideration to examine what analogy may have to do with it.

I have mentioned that concessive parentheticals show a fair amount of variation across a range of several variables. I have also suggested that concessive parentheticals can be seen as a paradigm of constructions. Paradigms of constructions, as researchers like Florent Perek have recently argued, could be regarded as higher order schemas and as meta-generalizations that are represented at a very abstract level in the network of constructions. My question for the remaining time that I have here would be whether there is a general concessive parenthetical construction? Is there a construction that represents that generalization, that paradigm?

Let me specify that question a little bit more. How do we find out whether there is such a generalization? The generalization should correspond to a schema in speakers' minds that allows them to combine any kind of concessive conjunction, like *although*, *though*, *if* and *while*, with some kind of predicative phrase structure that can be an adjective, an adverb, a noun, a prepositional phrase, a past participle, *-ing* clause, and that can be embedded in matrix structures that are either a sentence or a noun phrase. If you think of that as a very general syntactic rule, a very productive schema that can generate many different concessive parentheticals, that would be the kind of generalization that speakers would have to have in their minds.

What would the evidence for such a high level generalization look like? You know that I like to work with diachronic data. I think that diachrony holds a cue with regard to this question here for us. If we see that the constructions of a paradigm structurally assimilate over time, we would have a reason to say that a meta-generalization is forming. If the conjunctions, for example, combine more and more freely with different types of predicative elements, and if the relative frequencies of structural variants becomes more and more

homogeneous over time, that would be evidence for an increasingly paradigmatic structure.

Let me give you a hypothetical example. Let's say that adverbial parentheticals are at first primarily attested with *although* as in *John apologized, although reluctantly*. Let us further say that eventually such structures also appear with *if* and with *while*, so that speakers start saying "*John apologized, if reluctantly*" or "*He apologized honestly, while grudgingly*". Those examples are not corpus examples, but let's say that the structural profiles of these alternatives become more similar over time.

Whether different concessive parentheticals become more or less similar over time can of course be measured. Relative similarities between concessive parentheticals with *although*, *though*, *if* and *while* can be operationalized in terms of the relative frequencies of their structural variants. How often is *although* used with adjectival structures in the matrix clause? How frequent are structures that are embedded in a noun phrase? Those are these examples like *an interesting*, *if problematic, idea*. How do those frequencies compare to the other conjunctions?

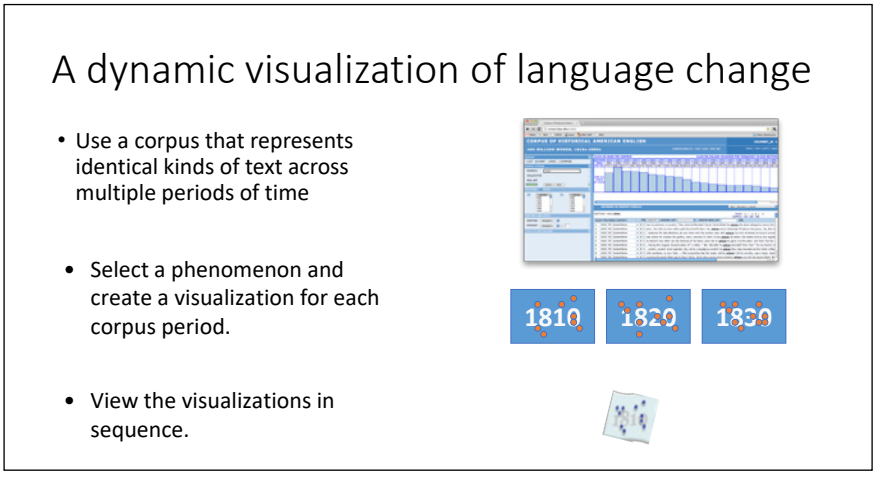


FIGURE 17

At this point, I want to take a small step back to explain to you the methodology that I applied in this study. In order to see whether concessive parentheticals become more similar or more differentiated, I used a dynamic visualization of language change. The idea with that is that you use a corpus that represents identical kinds of text across periods of time. Then, you select a phenomenon and create a visualization for each corpus period. Then you view the visualization in sequence. Let me give an example of how this works in practice. Let's

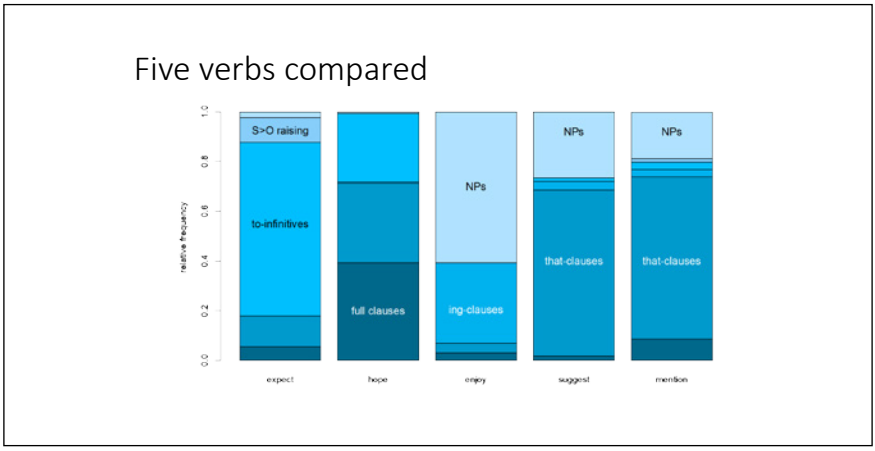


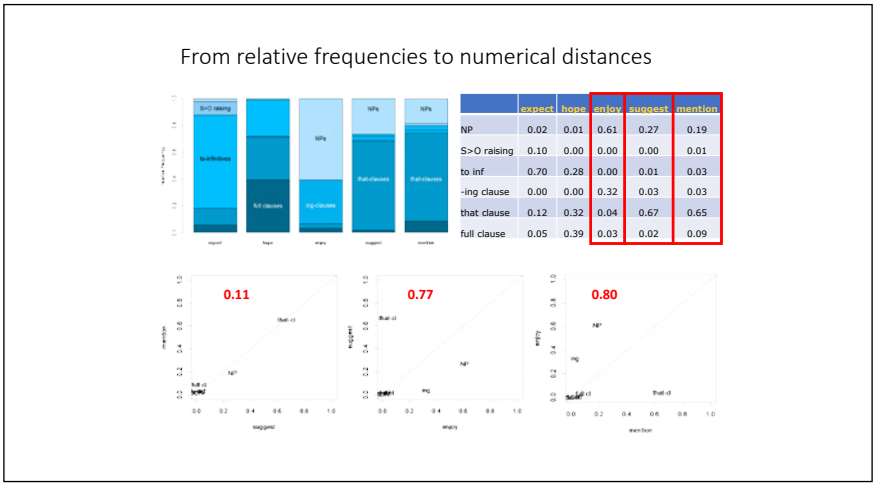
FIGURE 18

take complement-taking verbs in English. Verbs such as *expect*, *like*, or *imagine* project a syntactic complement structure that can take different shapes. I can start a sentence with *I expect*, and I can finish it with, for example, a noun phrase, as in *I expect a visitor*. I can finish it with a more complex infinitive structure , as in *I expect to hear from John*. I can have a *that*-clause, as in *I expect that John will win*. Other options include raising constructions, as in *I expect John to do well on the exam*.

The key point here is that these different subcategorization frames differ in their relative frequency. Every complement-taking verb has a certain unique profile.

In this graph, I have compared five different complement-taking verbs, namely *expect*, *hope*, *enjoy*, *suggest* and *mention*. The *y*-axis shows the relative frequencies of the different complementation patterns that I mentioned. The relative frequencies add up to a hundred percent. The verb *expect* largely occurs with *to*-infinitives. It sometimes occurs with raising constructions, as in *I expect John to win*. The verb *hope* often occurs with full clauses, as in *I hope I can see you again very soon*. The verb *enjoy* frequently occurs with noun phrases, as in *I enjoyed the food* or I enjoy *talking to you*. *Suggest* and *mention*, frequently occur with *that*-clauses and noun phrases, as in *I suggest that we meet this afternoon*, or *I suggested the fish*. The overall profiles are different, but some verbs are closer to each other than others. The complementation profiles of *suggest* and *mention* are fairly similar. The other verbs have more individual properties.

The data in the bar graph is the basis for another visualization that we are going to create. The relative frequencies from the bar graph are shown as

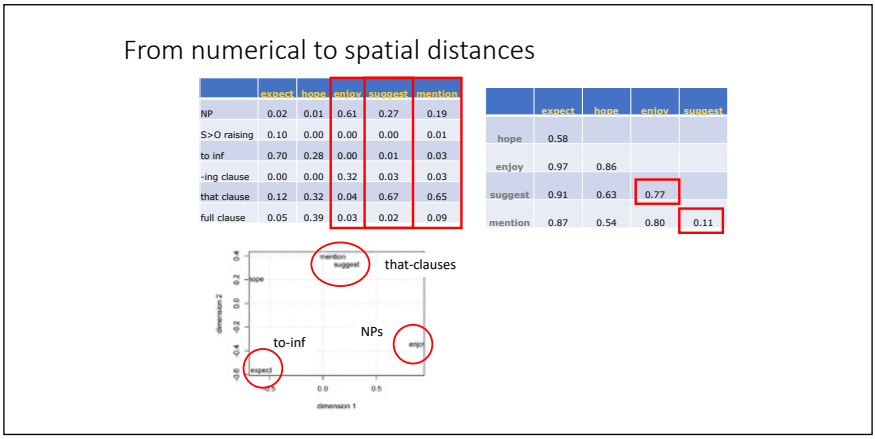


	expect	hope	enjoy	suggest	mention
NP	0.02	0.01	0.61	0.27	0.19
S>O raising	0.10	0.00	0.00	0.00	0.01
to inf	0.70	0.28	0.00	0.01	0.03
-ing clause	0.00	0.00	0.32	0.03	0.03
that clause	0.12	0.32	0.04	0.67	0.65
full clause	0.05	0.39	0.03	0.02	0.09

FIGURE 19

numbers in the table on the upper right. They can be used to calculate pairwise measures of similarity. For example, if you compare the relative frequencies of *suggest* and *mention*, you get a score of 0.11 which indicates a high degree of similarity. When you line up the relative frequencies of *suggest* and *mention* and put them into a two-dimensional coordinate system, they form a nearly perfect diagonal.

By contrast, when you compare *enjoy* and *suggest*, you see that the relative frequencies are far away from the diagonal. *Suggest* frequently occurs with *that*-clauses. *Enjoy* frequently occurs with NPs. The two are not at all alike, and this results in a larger score of 0.77. The same is true for *enjoy* and *mention*, which have very different preferences.



	expect	hope	enjoy	suggest	mention
NP	0.02	0.01	0.61	0.27	0.19
S>O raising	0.10	0.00	0.00	0.00	0.01
to inf	0.70	0.28	0.00	0.01	0.03
-ing clause	0.00	0.00	0.32	0.03	0.03
that clause	0.12	0.32	0.04	0.67	0.65
full clause	0.05	0.39	0.03	0.02	0.09

	expect	hope	enjoy	suggest
hope	0.58			
enjoy	0.97	0.86		
suggest	0.91	0.63	0.77	
mention	0.87	0.54	0.80	0.11

FIGURE 20

Calculating measures of similarity for all possible pairs yields a distance matrix, which is what you see on the upper right of this slide. That kind of distance matrix can be transformed into a two-dimensional map that visualizes the distances. The verbs *suggest* and *mention* have a score of 0.11. As a result, they pattern together in the map that you can see on this slide.

The verbs *enjoy* and *suggest* have a larger score of 0.77, and they are farther away from each other in this representation. The different positions in the graph reflect the preferences of these verbs for different complementation patterns. *Mention* and *suggest* have a strong preference for *that*-clauses. *Enjoy* has a strong preference for NPs. *Expect* frequently occurs with *to*-infinitives.

We are now going from toy data to real data. In doing that, we are just expanding this kind of data set that you've just seen from 5 verbs to 45 verbs. I have taken them from the literature on English complement-taking verbs, including *expect*, *like*, *try*, *suggest*, *mention*, and so on and so forth. For each of these I retrieved frequencies from the Corpus of Historical American English, using data from the 1860s to the 2000s. I extracted frequencies of the subcategorization frames, i.e. the syntactic options that I have shown you for the periods.

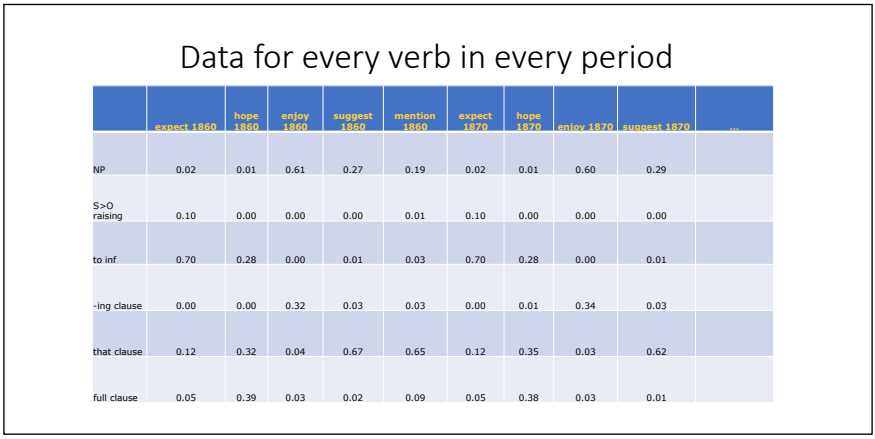


FIGURE 21

Proceeding in this way, I obtained data for every verb in every period. That means that the table that we are looking at now is a lot larger than what we had before. I am comparing *expect* in the 1860s to *hope* in the 1860s, to *enjoy* in the 1860s, to *suggest* in the 1860s, and so on and so forth. *Expect* from the 1860s is also compared against *expect* from the 1870s, where the relative frequencies might have changed a little bit. Every verb is thus compared against all the others, and it is compared against itself in different time periods.

We end up with a distance matrix that has 675 × 675 cells. We have 45 different verbs times 15 sequential periods. The distance matrix forms the basis for an explorative statistical analysis. One method that allows you to do that is called multi-dimensional scaling. That technique takes this distance matrix and aims to place all data points on a two-dimensional map, in such a way that the mutual distances are preserved as accurately as possible. This means that the method is going to cut corners. The representation is never perfect, but there are quality measures that are in place. The maps that you will see here meet common quality standards. Instead of viewing all data points at the same time in a single graph, we view them time-slice by time-slice in a movie-like fashion.

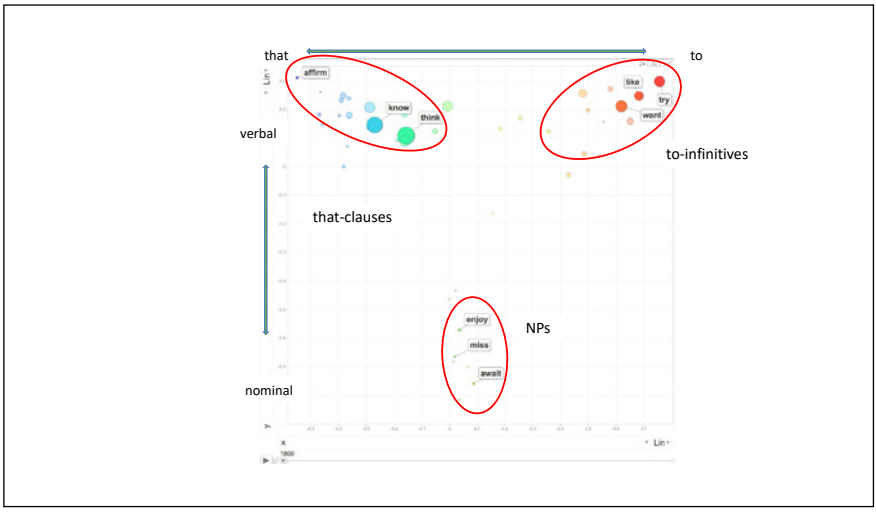


FIGURE 22

On this slide you see a map just like the one that I showed you earlier with *expect*, *enjoy*, and *suggest*, except that now we have all 45 verbs and we have analyzed them across the 15 different time slices of the COHA data. On the basis of diachronic data, we can see how the configuration of verbs develops over time.

Over time, the complementation preferences of the different verbs undergo changes. The overall system in its triangular shape stays in place. That system shows clusters of verbs with different preferences. They are arranged in what we could call a paradigm. You see little points moving about, those are infrequent verbs for which the relative frequency measurements are not particularly reliable. The larger bubbles tend to stay in in their places, which shows that their profiles are stable.

Let me give you a guide on how to interpret this map. This cluster in the upper left, with verbs like *affirm*, or *know*, or *think*, they frequently occur with *that*-clauses. The verbs in the upper rght corner, *try*, *like*, and *want*, they occur with *to*-infinitives. The verbs at the bottom of the graph tend to occur with NPs, i.e. *enjoy*, *miss*, and *await*. That does not mean that their preferences are absolute. Those are just preferences. Every single one of these verbs can potentially at least occur with all of the complementation patterns.

At the top, we have a continuum of preference from finite *that*-clauses on the left to non-finite *to*-infinitives on the right. With regard to the *y*-axis, we have verbal complex structures at the top. We have nominal and more compact structures at the bottom.

Throughout the entire time that is analyzed, the two most important dimensions of verbal complementation are these two distinctions, the distinction between *that*-clauses and *to*-infinitives and the distinction between nominal and verbal complements.

Since this is just an illustration of the method, I will now move on and get back to concessive parentheticals. But before I do, there is one thing that I would like to show you, and that is the history of the verb *confirm*. Given all I have just said, *confirm* starts out in the 19th century as a verb that prefers noun phrases. Over time, *confirm* has drastically changed its complementation behavior. In the earlier decades it occurs with noun phrases, as in *They confirmed the rumor*, or *They confirmed his position*. By the 2000s, *confirm* is centrally in the cluster of verbs that tend to occur with *that*-clauses cluster. That is how the verb is used now, as in *They confirmed that the story was accurate*. Another interesting development concerns the verb *dislike*, which I have mentioned a couple of times in earlier lectures. You remember the frequency charts with *dislike to do something* and *dislike doing something*. These data actually reflect that same development. *Dislike* starts out in the *to*-infinitive cluster. As time goes on, it steadily but surely develops a preference for *ing*-clauses, which instantiate a nominal structure, i.e. gerunds, but which are nonetheless more verbal than just straight nominal phrases. That is why verbs that prefer *ing*-clauses are situated in the middle of this whole configuration.

This was a fairly lengthy illustration of what this method does. I want us to get back to concessive parentheticals.

Let's look at the data for the concessive parentheticals, which look very similar to what we have just seen with the complement-taking verbs. Here we see historical data from the COHA that shows the syntactic variants of concessive parentheticals with *although*, *though*, *if* and *while*. Something that you see here is that all four conjunctions appear in parentheticals with adjectives. All of them have sizable black bars and the ratios are about the same. These are examples such as "*Power, while important, is not everything*". The one thing that

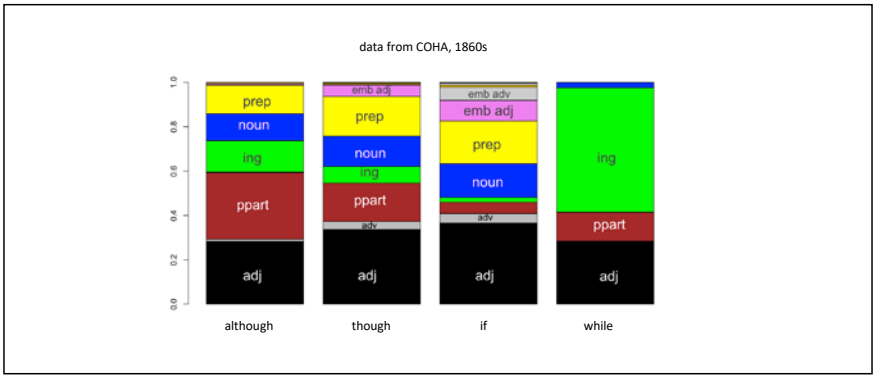


FIGURE 23

stands out in this graph is this large portion of *-ing* for *while*. The conjunction *while* has this strong preference for *ing*-clauses because of its temporal heritage. *If*, on the other hand, has the strongest preference for phrase-embedded examples. The part of the bar that is colored in pink represents examples such as "*That is an interesting, if complicated, idea*". You see that *although* and *though* are relatively similar in their distribution. *Though* has a little more of the phrase-embedded structures, and it has fewer examples of *-ing* than *although*. *Although* has fewer past participles and a few more adverbs.

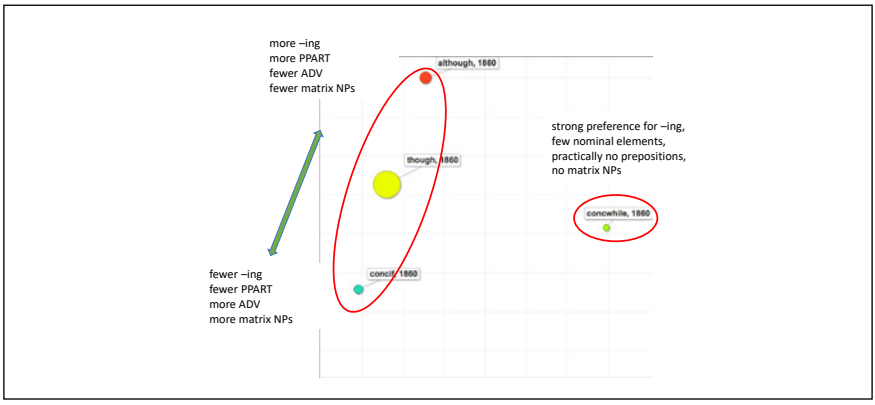


FIGURE 24

I transformed this data with multi-dimensional scaling in order to create the graph you see on this slide. The distances between *although*, *though*, *if*, and *while* reflect mutual similarity. This graph is based data from the 1860s.

Let's examine it in detail. *While* has a strong preference for *-ing*, it has few nominal elements, practically no prepositional phrases and no matrix NPs. The left side of the graph shows *although*, *though*, and *if* arranged on a cline that we

can interpret. *Although* has more *ing*-forms than *if*, more past participles, fewer adverbs and fewer matrix NPs. Conversely, *if* has the fewest *ing*-forms in the set, it has the fewest past participles, it has more adverbs and more matrix NPs.

How has all of this has changed over time? Have concessive parentheticals become more similar over time, or have they become more different? What I hypothesized when I started this research was that concessive parentheticals would indeed converge on a common behavioral profile. I thought that over time, they would become more similar and gravitate towards the middle of the graph. Let's see what happens in reality.

If we look at the diachronic developments, we see that *though* and *although* converge on a similar pattern. By contrast, *if* and *while* move towards the outer margins of the graph. We see assimilation for *although* and *though* and differentiation for *if* and *while*. That means that a developing paradigm may incorporate conflicting processes. We can find two pairs of the paradigm matching up and functioning as near-synonyms that you can analogize to each other. And there are others that specialize in different functions.

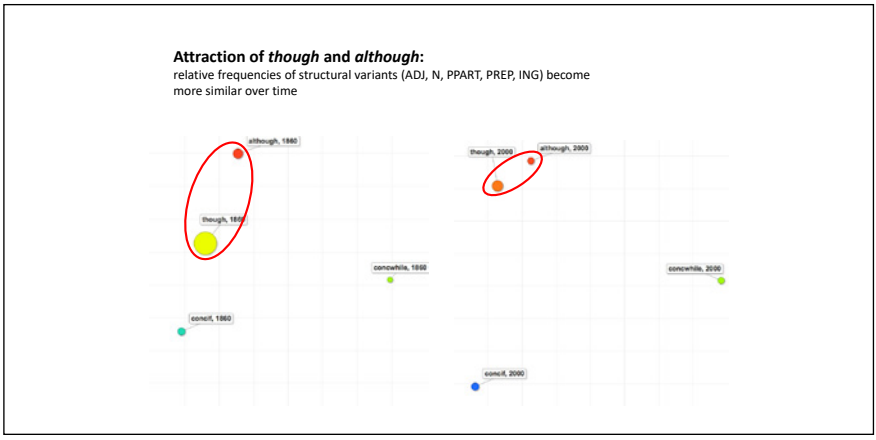


FIGURE 25

This slide shows the attraction of *though* and *although*. The left panel shows the graph from the 1860s, and on the right you see the 2000s, where *though* and *although* have clearly become more closely associated.

At the same time, *though* and *if* have been dissimilating. They are rather close to each other in the 1860s, but they become more different over time. There is an outward movement of *while*. This slide shows where it started out in the 1860s and where it ended up in the 2000s. The same is true of *if*, which moves further towards the bottom of the graph.

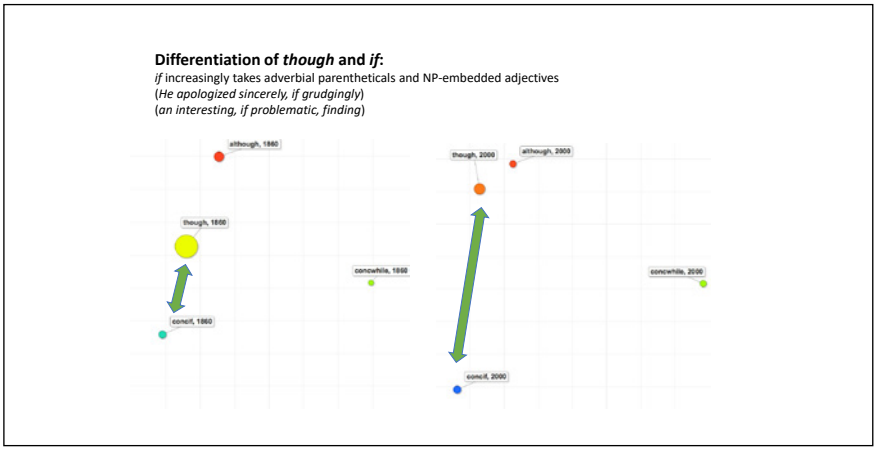


FIGURE 26

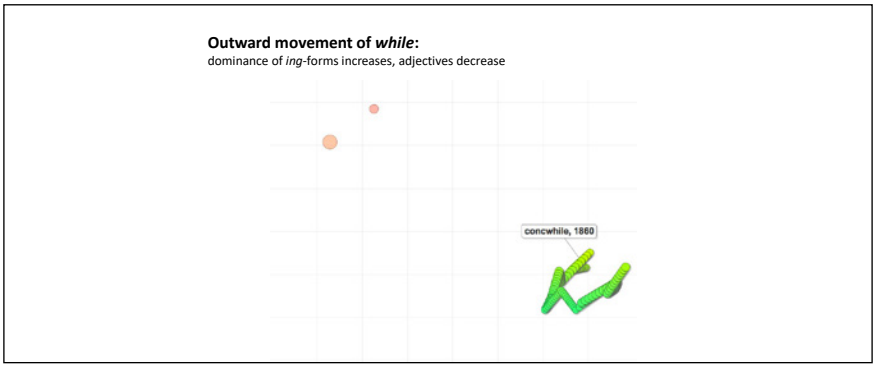


FIGURE 27

To come to an interpretation of this development, between 1860 and 2008, there is no evidence to suggest that a general concessive parenthetical construction emerges. I would claim that for this construction family, it is unlikely that there is an overarching, higher order schema that would allow speakers to produce all kinds of variations of that construction. Rather, I think the local generalizations have become increasingly more important.

We observe processes of structural dissimilation with *if* and *while*, but within the overall process, *although* and *though* converge on a mutual constructional schema. I have been using the term paradigm for that idea. There is one other term that I think is useful in this context, and that would be the term of a construction family. A construction family is a set of constructions that share some resemblances, but that are not necessarily all connected to each other in the same way.

To sum up, developments such as the ones that can be observed with concessive parentheticals illustrate what I wanted to express with the slogan "grammatical change is not a zero-sum game".

Paradigmatization is what we could call the constructionalization of a higher order schema. Within paradigms, partial functional overlap is actually common, not rare. Grammatical constructions may of course arrange themselves in a complementary distribution in functional space. That is something that is true not only of the data that you have seen here, but also of grammatical paradigms in general, across a wide range of languages. But that said, grammatical constructions may also inhabit the same functional space where they would then serve as variants of each other, and presumably stand in mutual competition, as I discussed yesterday with the case of *my* and *myne*.

I hope I have given you some ideas on how competition, differentiation and attraction are interrelated and how they are all necessary to understand what happens when groups of constructions change in the constructional network. I have also tried to argue that it is important to try to understand all of this in terms of links rather than in terms of constructional properties that are described directly into the nodes. With that I would like to end, and I thank you once more for your attention.

LECTURE 8

The Asymmetric Priming Hypothesis

Welcome back to Ten Lectures on Diachronic Construction Grammar. The title of this lecture is "The asymmetric priming hypothesis". In the past four lectures, I have presented several corpus linguistic studies of constructions and how they changed over time. I have explained how I see the relation of grammaticalization theory and Diachronic Construction Grammar. I have discussed corpus linguistic methods that can be applied for the study of these processes. This lecture will be something of a change of pace. I won't be talking about diachronic corpus data in this lecture and I will present instead an analysis of synchronic corpus data and experimental behavioral data. You might ask yourself what psycholinguistic experiments can contribute to the study of language change.

Let me give you a preliminary answer to that question. The answer to that has to do with one of the central questions of usage-based linguistics, namely, how cognitive processes that operate in the here and now affect language in such a way that it changes in regular ways over time. In the very first lecture I have discussed ten basic ideas of Construction Grammar. The tenth one was the idea that language draws on domain-general socio-cognitive processes that include categorization, association, routinization, generalization, schematization, joint attention, statistical learning, analogy, metaphor and several others. We as human beings are equipped with a set of social cognitive skills that conspire to let us learn and use language.

Michael Tomasello expresses this idea in a way that is really simple and to the point (2005: 193): "*Children acquire all linguistic symbols of whatever type with one set of general cognitive processes*". As someone who is interested in language change, I want to find out how these processes operate in actual speech situations that influence how language changes over the long term.

In this talk, I want to take a closer look at one such cognitive process, namely a specific subtype of priming that has been called "asymmetric priming", which has been suggested as a force that shapes language change.

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691224.

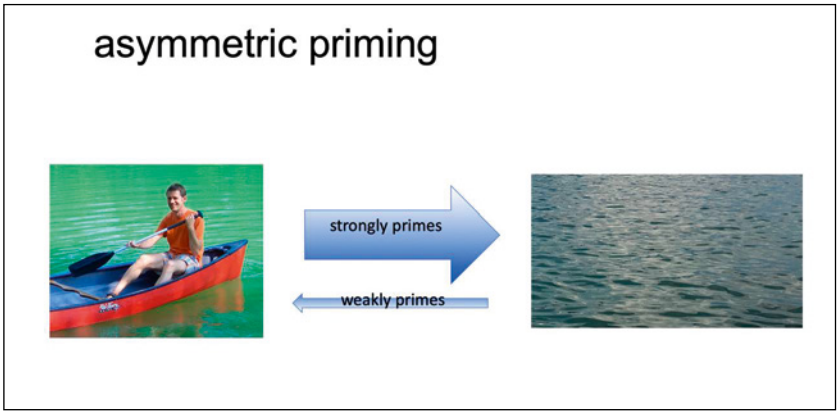


FIGURE 1

To get us started, take a piece of paper, take a pen and write down three words that come to mind when you hear the word *paddle*. One very frequent response that people write down is the word *water*. In other words, the word *paddle* primes you for the idea of *water*. It evokes the idea of *water* or at least it makes it easier to process the idea of *water*. By contrast, what do you imagine people will write down as their first associations when I give them the word *water*? Starting with *water*, you will think of *drink*, *rain*, *sea*, *river* and maybe *flower*, because flowers need water. But one item that you surely won't find among the first three words is the word *paddle*.

Paddle strongly primes *water*, but *water* only weakly, if at all, evokes the idea of a *paddle*. Now you know what asymmetric priming is, but you are perhaps wondering what it has to do with language change.

I will get to that in just a minute. First, let me give you an overview of this lecture. In the first step, I will introduce what's called the asymmetric priming hypothesis. I will outline what this hypothesis predicts for behavioral data and for corpus data. Then I will present an experimental study that addresses these predictions. After that, I will continue with the second study that tests the asymmetric priming hypothesis on the basis of corpus data. In the fourth part, I will offer some tentative conclusions. Without further ado, what is the asymmetric priming hypothesis?

The hypothesis has been proposed in a programmatic paper written by Jäger and Rosenbach (2008). In this paper they state the following:

> We argue that the psycholinguistic mechanisms of PRIMING may account for the empirical observation that grammaticalization processes typically proceed in one direction only.

As I have discussed in previous lectures, grammaticalization is concerned with the emergence of grammatical forms, and it describes a number of different interlocking types of change that all proceed in one direction only. Full forms with lots of phonetic substance develop into forms with less phonetic substance. Forms that are only loosely integrated in discourse develop into forms that are syntacticized, into forms that are further compressed into morphological constructions. Concrete meanings are becoming more abstract and schematic, but not the other way around. The changes in grammaticalization are asymmetric. There is one direction and we see the changes moving in that direction only. Jäger and Rosenbach argue that they have an explanation for unidirectionality. They claim that there is a cognitive process, something that happens in conversation and acts synchronically on the minds of speakers, that could explain why in the long run changes happen in languages the way they do.

Jäger and Rosenbach elaborate on this first statement:

> Very generally, the prediction is that in any reported case of change, where the development goes unidirectionally from A to B, A should prime B but not vice versa.

When we have a grammaticalization process in which A turns into B, then there should be an asymmetric priming relation, so that A primes B, but not the other way around. That is a claim that can be tested empirically in an experiment. Let me show you a concrete example of what Jäger and Rosenbach are proposing. Let's look at a case where we have A changing into B, and let us consider the predicted priming relation between the two.

Imagine that you are out on the road. That stick figure here on the slide, that is you, and you meet a friend of yours who tells you "*I am going to the station*". The asymmetric priming hypothesis would predict that even though your friend just talked about the act of walking, you would be likely to think about your friend's future and his actions in the future. What is your friend going to do? Is he going to catch his train? Where will he be going and what will he be doing once at his destination? The utterance "*I am going to the station*" triggers all these temporal associations. Conversely, if you're talking on the phone and the person you're talking to tells you "*Oh, I gotta go. I am going to miss my train*", your friend's utterance encodes a future event. According to the asymmetric priming hypothesis, talking about time should not make you think about physical motion, at least not as much. You're talking about a future event and you can think about that in its own right, without thinking about motion. This, in a nutshell, translates into the following research question: Does *go* prime *be*

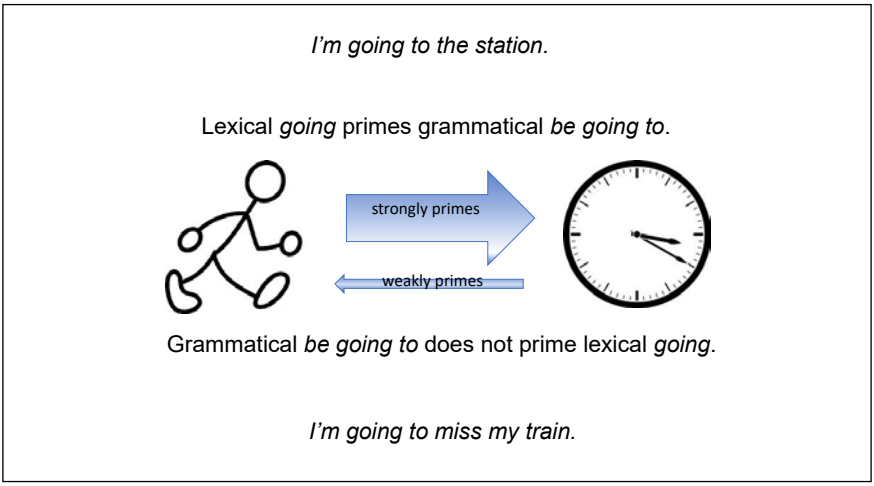


FIGURE 2

going to, but not vice versa? According to what I have just told you, we have lexical *going*, which should prime grammatical *be going to*, but grammatical *be going to* should not prime lexical *going*. Now, we are talking about actual linguistic forms that we can integrate into experimental stimuli that we can present to subjects. Doing so allows us to check whether the predictions of the asymmetric priming hypothesis actually hold up.

Does *go* prime *be going to*, but not vice versa? If it does, asymmetric priming might be an explanation for one aspect of unidirectionality in grammaticalization, namely, unidirectional semantic change. That is just one aspect of unidirectionality, but it is an important aspect. It would be one big step forward if we could explain this unidirectional aspect of grammaticalization in terms of a psychological process that operates in the here and now.

In order to find out whether this is actually the case, David Correia Saavedra and I designed an experiment with which we wanted to assess whether or not the asymmetric priming hypothesis makes the right predictions about language processing in real-life speakers. What did we do?

We constructed a database of 20 elements that in English have both lexical uses and grammatical uses. Basically any grammaticalized form that you can find has a lexical counterpart of some sort. There are exceptions like demonstratives for instance, which often do not have lexical counterparts. But for most of grammaticalized forms, there are correspondences such as the ones that you have on this slide here. The English verb *have* has grammatical uses as in "*I have solved the problem*", and it has lexical uses such as "*I have a problem*". Lexical *have* encodes possession, and grammatical *have* encodes the perfect.

Grammatical *have* has other functions as well, as in "*I have someone service the car*", where it has causative meaning. Another pair in our database consists of lexical *keep* and grammatical *keep*. Lexical *keep* means "*Keep your jacket on*", maintain it in a certain position. "*Keep walking*" means carry on walking. It aspectually modifies the meaning of the lexical verb.

Besides verbs, our database contains other parts of speech, including forms that start out as lexical adjectives and then become a part of a grammatical element. The adjective *long* is an adjective, and you can use it lexically in expressions such as "*as long as a python*". It also appears as part of the clause connector *as long as*: "*As long as you do not get caught, you can do anything*". We can reconstruct the semantic development that went from extension in space, i.e. how long something is, to a temporal extension, as in "*That was a long story*" to conditional meaning, as in "*as long as you do not get caught*". The grammaticalization path of *long* is thus a semantic development from spatial meaning to temporal and conditional meaning.

We created a database with 20 such pairs, and we designed a test of the asymmetric priming hypothesis in order to find out whether there are any asymmetries in the ways lexical elements prime their grammaticalized counterparts and vice versa. I mentioned earlier the example of *going* and *be going to*, saying that *going* should prime *be going to*, but not vice versa. The same phenomenon should be observable across all pairs in our database, so that *have a problem* should prime *have solved the problem*, but not vice versa. *A long python* should prime *as long as*, but *as long as* should not prime *a long python*.

We used an experimental method that is known as the maze task. Since this procedure is not so common, I brought you a little illustration. In the experiment, participants have two response buttons, one on the left and one on the right. They are given the following instructions. They are told that they will see two words displayed on the screen, and that their task is to select one of the words that they see. The words that they select should combine to form a meaningful sentence. The next slide shows an example, so you can actually experience directly what the participants saw.

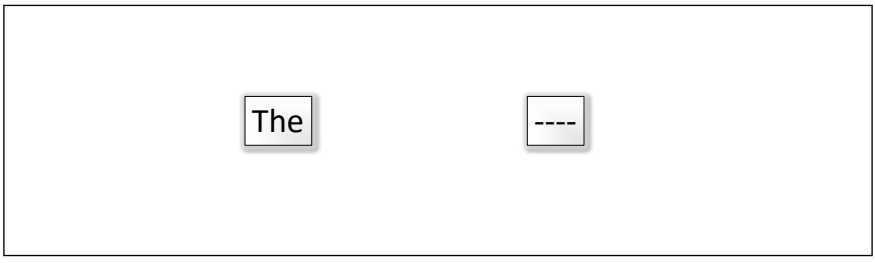


FIGURE 3

Here's the first screen. There are two fields and because this is the first pair, there is only one word and the other box shows only three hyphens. Here our participants have to press the button for *the* on the left.

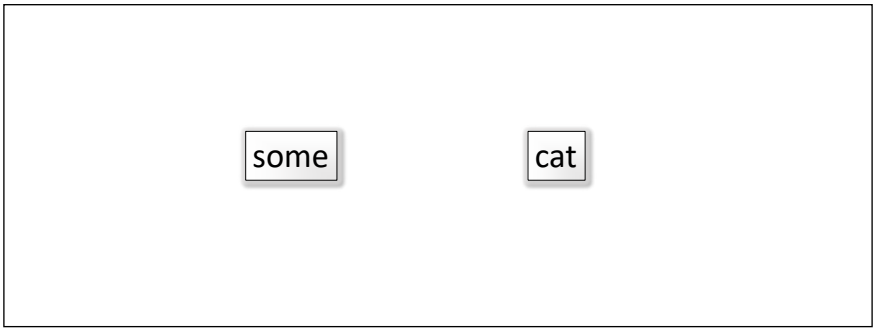


FIGURE 4

The next screen shows the word *some* and the word *cat*. After *the*, what word do you have to choose to continue a grammatically functioning sentence? You have to choose *cat*, so *the cat*.

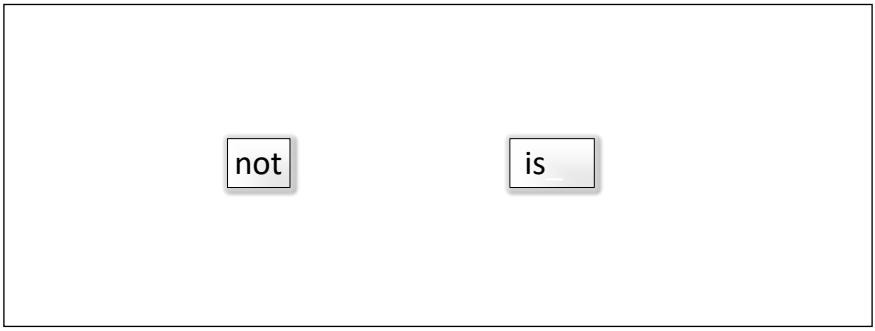


FIGURE 5

Here we have *not* and *is*. Which one do you like better? Is? Ok.

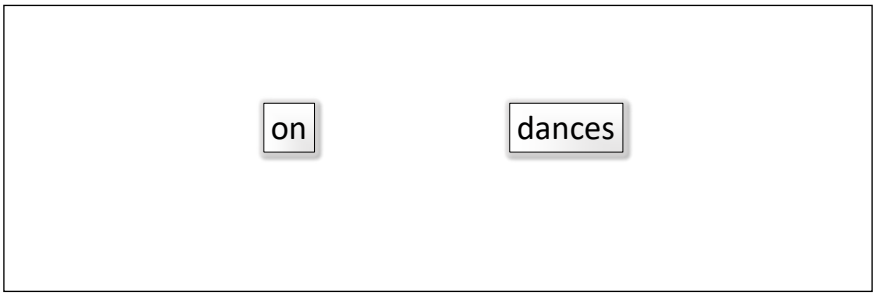


FIGURE 6

We have *on* and *dances*, so *The cat is on*.

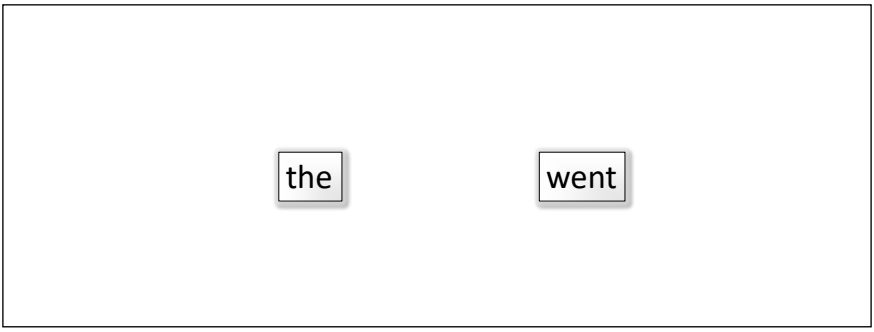


FIGURE 7

The. Ok.

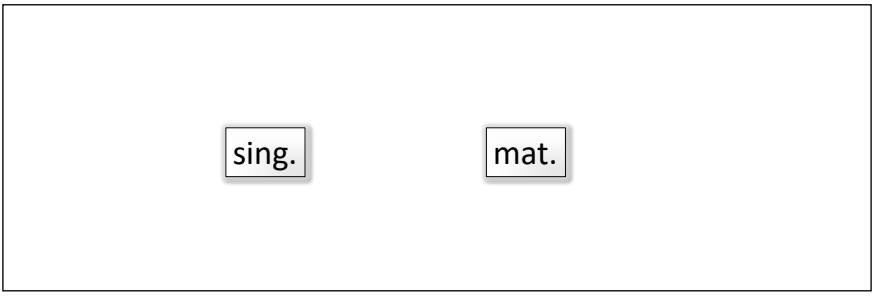


FIGURE 8

Mat. This gives us one of philosophy's most famous sentences, *The cat is on the mat*. What you have just done was exactly what our participants did, except they had to do it for a much longer time, and of course with other sentences. The maze task is a mixture between a self-paced reading design, where you read and you click as soon as you have read something, and a forced-choice task, i.e. there are two options, and you have to pick one.

Subjects see screens that present two options, but only one of them makes sense, given the prior context. Every time they press a button, we measure reaction times, so we know exactly how long it takes them to pick the right word. This can be exploited for the analysis of asymmetric priming effects.

For our experiment we recruited 200 speakers of American English via an online platform. These people could sign up online and do this task in their own homes. That comes with advantages and disadvantages. The disadvantages are that we do not know who they are, we do not know if they were checking their social media while they were doing the experiment, and we do

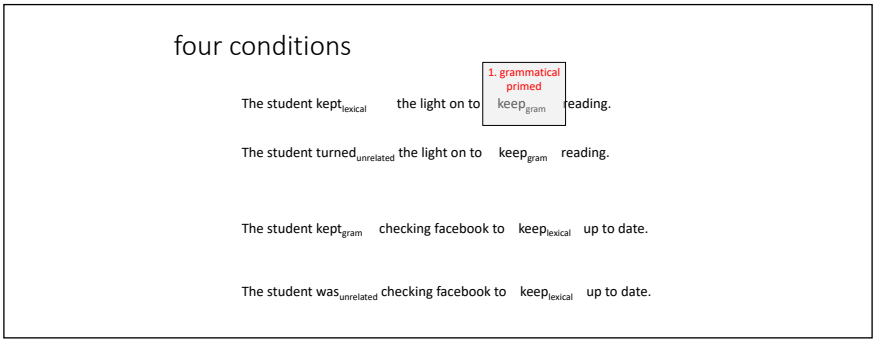


FIGURE 9

not know how much attention they were paying. However, the advantage is that you can recruit many participants in a very short time.

The 200 speakers of American English clicked their way through 40 stimuli sentences, 20 of which contained a primed word. Those were the critical stimuli, the rest were fillers. Only fully correct responses entered the analysis. If someone made a mistake and selected the wrong word, we threw away the data point. We only took sentences in which the participants got every single word right. The overall experiment took them about 12 minutes. We measured reaction times, i.e. how fast they selected the correct element across four different conditions, that is four different versions of the stimuli. This is something that I need to explain in detail.

In the first condition, we collected responses to a grammatical element that had been primed by its lexical counterparts. Let me give you an example for that kind of sentence. If we have a sentence like "*The student kept the light on to keep reading*", we have two instances of the word *keep* in the sentence. The first is lexical, "*the student kept the light on*" and then there is "*to keep reading*". We have the grammatical version of *keep* in the second half of the sentence. This would be what we call the grammatical primed condition, that is, we are measuring how quickly participants respond to grammatical *keep* at the end of the sentence, and that grammatical *keep* has been primed by lexical *keep*. According to Jäger and Rosenbach (2008), this should be easy because there is a positive priming effect from the lexical source word, which should give people an advantage.

The second condition is similar, except here the grammatical word is not preceded by its lexical source, but rather by an unrelated word. The sentence would simply be "*The student turned the light on to keep reading*". In this context, grammatical *keep* does not have the advantage of being primed by its lexical source.

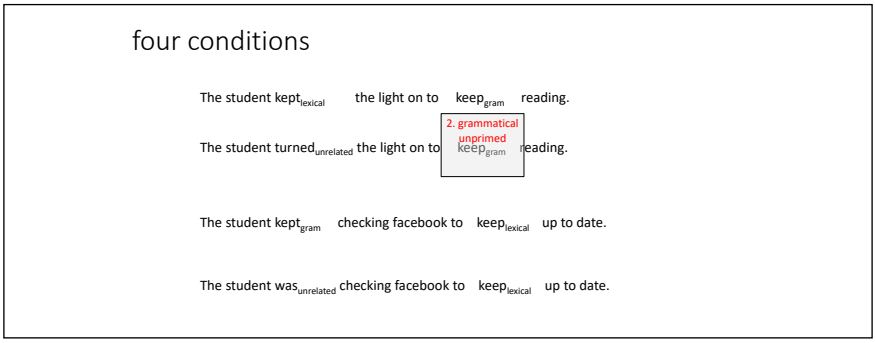


FIGURE 10

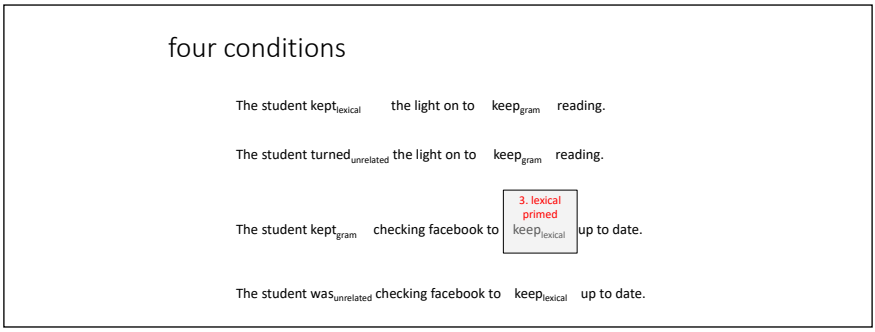


FIGURE 11

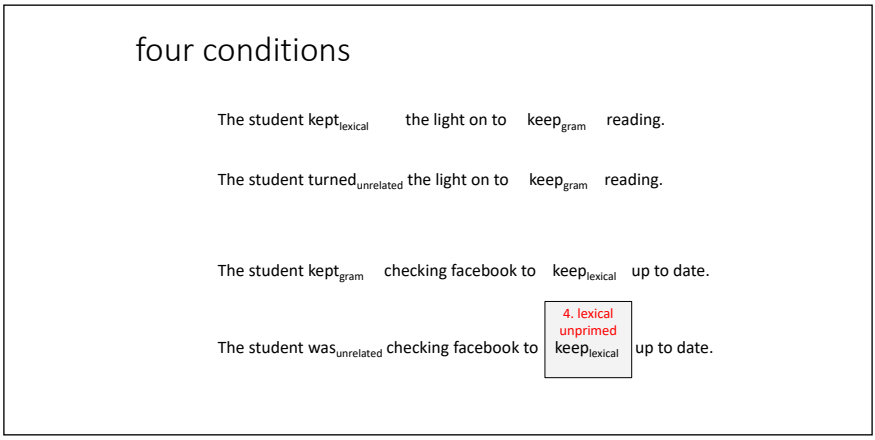


FIGURE 12

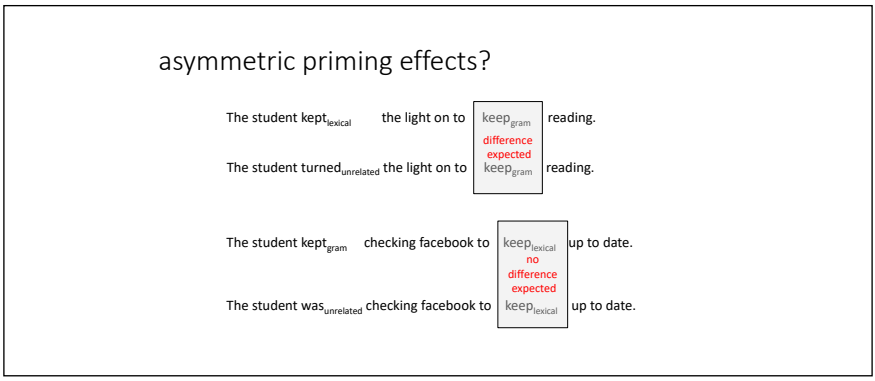


FIGURE 13

In conditions three and four, we measured responses to lexical variants. In "*The student kept checking facebook to keep up to date*", we have the grammatical variant of *keep* first, *the student kept checking*, and *keep up to date* in the second half of the sentence. We thus measured responses to lexical *keep* that had been primed by its grammatical counterpart. According to Jäger and Rosenbach (2008), that type of priming should not yield an advantage. There is no priming predicted from grammatical *keep* to lexical *keep*.

The last condition is exemplified by sentences like "*The student was checking facebook to keep up to date*". This is the lexical unprimed condition, in which lexical *keep* is not primed by a previous element in any way.

What priming effects are we predicting? On the asymetric priming hypothesis, *keep* in the first sentence should have a processing advantage over *keep* in the second one. In the second pair of conditions, a different prediction holds, namely, since grammatical *keep* is not expected to prime lexical *keep*, condition three and four should yield the exact same results.

For the analysis, we used a regression analysis design. Our dependent variable in this case was a continuous variable, namely, the reaction time. How quickly did people press the right button? The analysis includes several predictor variables. The first of course is the presence of priming. We also controlled for the fact whether a stimulus was grammatical or lexical. Since we are corpus linguists, we also included the text frequency of the items as a variable, because it is known that higher frequency items are more easily processed. We further included several control variables, namely gender, age and handedness of our participants.

We used a mixed-effects regression design and included random intercepts for the participants. Some people are just quicker than others. Some people just

had their coffee. Others had a rough night. We control for that with that factor. We also included a random factor for the test item, which allows the same kind of variation for all the different elements in the database. We included an interaction effect between the priming variable and the variable that distinguishes between lexical and grammatical. This is of course the heart of the asymmetric priming hypothesis: Does priming work differently, depending on whether we are measuring the response time for a grammatical element that has been primed by its lexical counterpart, or whether we are measuring the response time for a lexical element that has been primed by its grammatical counterpart? That is the prediction that the asymmetric priming hypothesis makes.

We also tested for an interaction effect between priming and frequency, because frequent items may not need priming to be processed very quickly. When I give you a very rare word, for instance, *procrastinate*, the next time you hear *procrastinate*, this memory of hearing this word just recently pops into your mind and you can process it rather quickly. Contrast that with when I pronounce the word *the*, which you hear so often. Hearing it one more time won't change your response to it the next time you hear it. That is captured by an interaction term between priming and frequency. High frequency elements do not profit from priming as much as low frequency items.

What came out of the analysis? We computed a full regression model that includes all the variables that I have mentioned. None of the control variables that we included, i.e. age and gender and handedness, change the results in any way. The analysis further does not show evidence for an interaction between priming and frequency, so we remove that from the model as well.

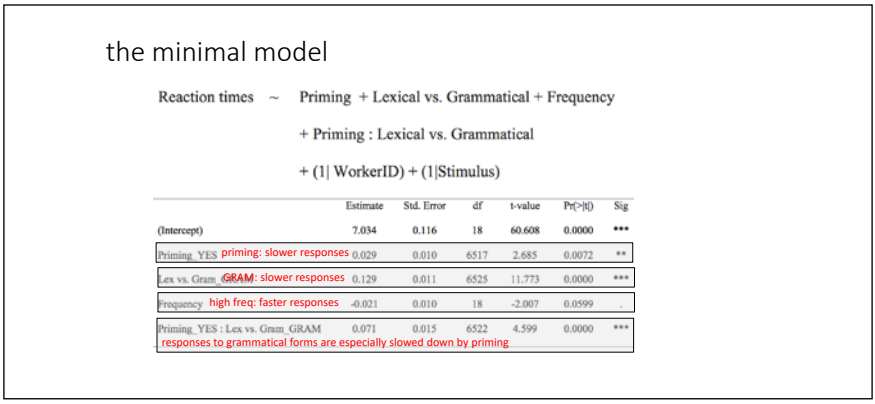


FIGURE 14

The exclusion of those variables results in a minimal model that only retains the variables that yield a significant effect. Those variables are the priming variable, the lexical versus grammatical variable, text frequency, and the interaction of priming, and lexical and grammatical. Let me walk you through the effects that we see in the table on this slide.

The first effect that I want to talk about concerns the priming variable. It turns out that forms that have been primed by their lexical or grammatical counterpart are not verified faster, but slower. That is the opposite of what one might predict. Primed forms have generally slower response times.

Another effect concerns the lexical versus grammatical variable. It turns out that grammatical forms in general have slower responses. That is equally surprising, because grammatical forms tend to be highly frequent. It should be very easy for the human processor to recognize these forms. This means that we have two rather unexpected results.

The next effect concerns frequency. High frequency items yield moderately faster responses. It is an established finding that high frequency forms are processed more easily. We find that in our data as well.

Then is the asymmetric priming effect. Priming interacts with the lexical versus grammatical variable in such a way that responses to grammatical forms are especially slowed down when those grammatical forms are primed. In other words, in "*The student kept the light on to keep reading*", that grammatical form of *keep* at the end of the sentence is the worst condition in terms of how quickly participants can process it.

How does that line up with our expectations? As you remember, we expected no difference between lexical primed and lexical unprimed. That prediction turned out to be correct, but as far as the predicted difference between grammatical primed and grammatical unprimed is concerned, we expected an advantage of the grammatical primed condition. Not only did we not find that, we found that responses in the grammatical primed condition are slower than responses in the grammatical unprimed condition. The difference is statistically highly significant.

On this slide you see a visual representation of the response times. We expected no difference between lexical primed and lexical unprimed. Indeed, if you look at the box plots there side by side, you can see that they are very similar. The average is around 900 milliseconds, and there is no recognizable difference. There is, however, a difference between the primed and unprimed grammatical conditions. The difference is significant, but it is in the direction that was not predicted by the asymmetric priming hypothesis.

So, does *go* prime *be going to*, but not vice versa? We observed asymmetric priming effects, but they do not work in the way we expected. Priming between

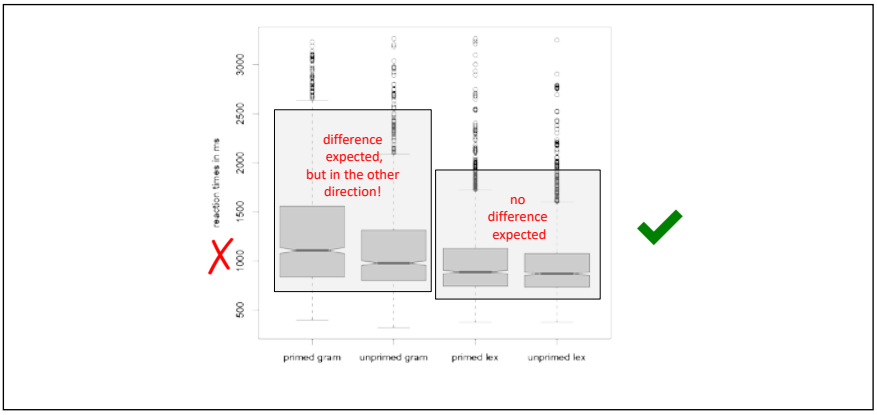


FIGURE 15

lexical forms and their grammaticalized counterparts is negative. It takes longer to process a given form when it has been primed. This negative priming particularly affects the processing of grammatical elements. To put it into a formula, *go* slows down *be going to* more than vice versa. Experimental evidence thus seems to contradict the asymmetric hypothesis.

Before giving up on it, we wanted to consider another methodological perspective and for that I turn back to corpus-based evidence. When we examine the way people actually use language, the way they write and the way they talk, are grammatical forms in corpus data preceded by the lexical sources more often than one might expect? Do people construct sentences like the one that we did? Or do they just not do that? We investigated these questions with the help of methods from distributional semantics.

At this point I am afraid I need to take a step back and say a few words about distributional semantics in general. I will eventually come back to asymmetric priming, but there are first some issues that I need to get out of the way. Please bear with me.

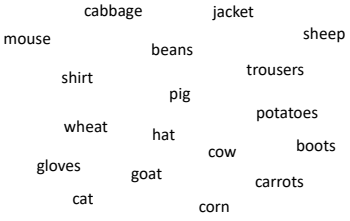


FIGURE 16

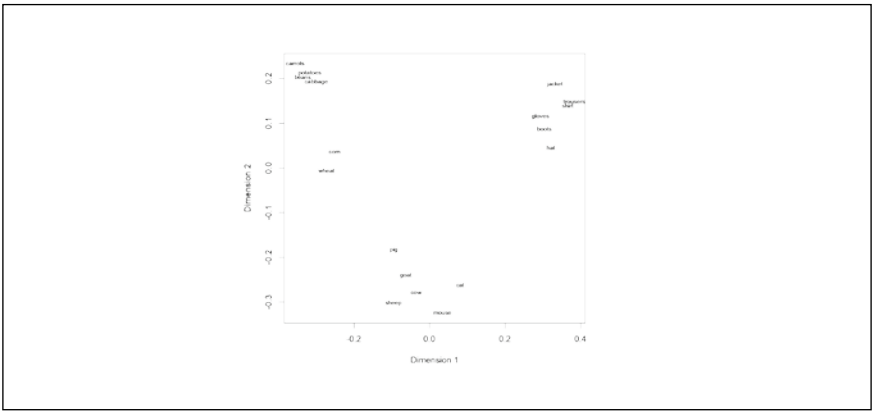


FIGURE 17

On this slide, you see different words that describe things that could be found on a farm. We have *cabbage*, *beans*, *pig*, *cow*, *potatoes*. We have *wheat*. We have clothing items like a *hat*, *gloves* or *boots*, things that a farmer might wear. These words are semantically related to each other, but of course they are not all equally similar to each other. In fact, if you were to give these words to a human observer, they would probably tell you that the words fall into three different categories, namely animals, vegetables and clothing items. The words *carrots* and *potatoes* instantiate vegetables. The words *gloves*, *hat* and *shirt* are clothing items. We have animals such as *pig*, *cow* and *sheep*. Distributional semantics is a computational way of studying semantic similarity across different words.
On this slide here, you see a visualization that looks a lot like the grouping of words that I showed on the previous one, except this one has not been made by a human being. Here we have the result of a computational analysis, in which a computer has categorized our farm words in a way that is quite similar to what a human being would have done. We have the vegetables in one corner, the upper left, we have the clothing items to the right, and we have the animals further down in the middle.

One thing you see is that in some way the computer here has been a little more insightful even than my own intuitive analysis, because you see that the computer thinks that *corn* and *wheat* aren't really prototypical vegetables, so *corn* and *wheat* form their own grain kind of category that is a little apart from the other vegetables. This is evidence that this method actually works.

But how does this work? How does the computer know that some words are related and others aren't? And what work steps are involved? I would like to walk you through the steps that are involved, so that you have a sense of what lies behind this.

The first work step that is involved is that we choose a vocabulary of key words that we are interested in. We retrieve concordances of these words from a corpus, which gives us access to frequencies of items in their context. Let me show this in practice.

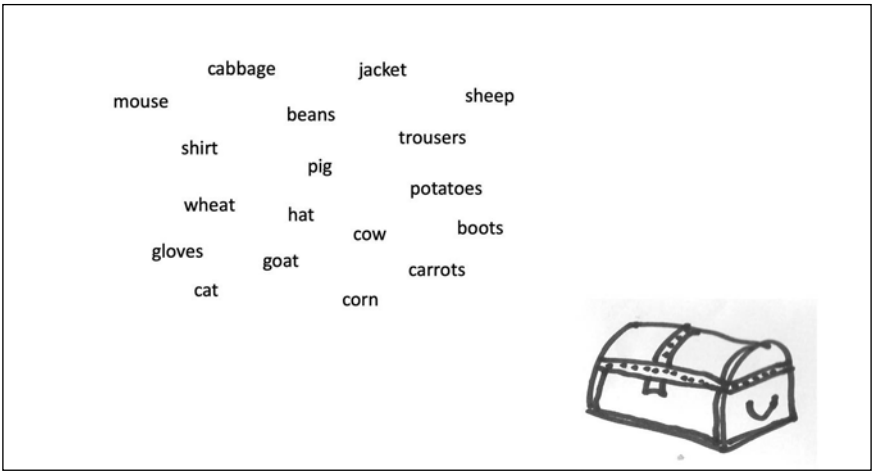


FIGURE 18

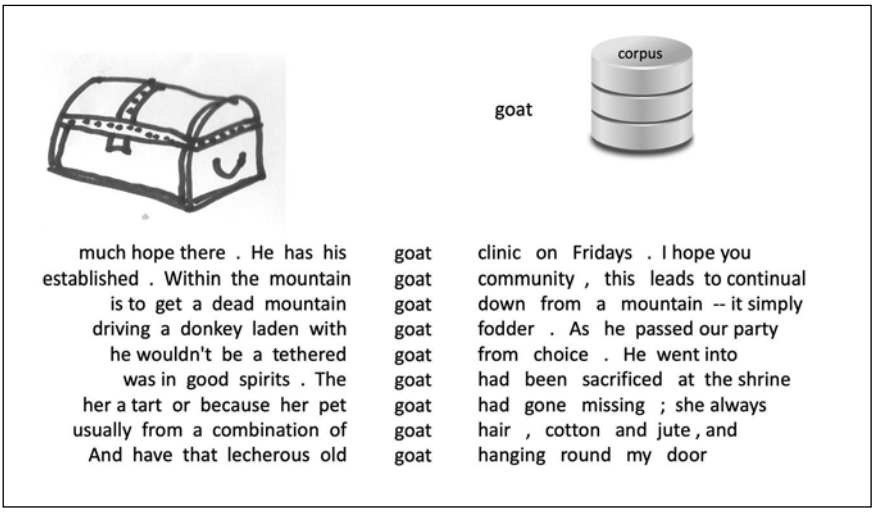


FIGURE 19

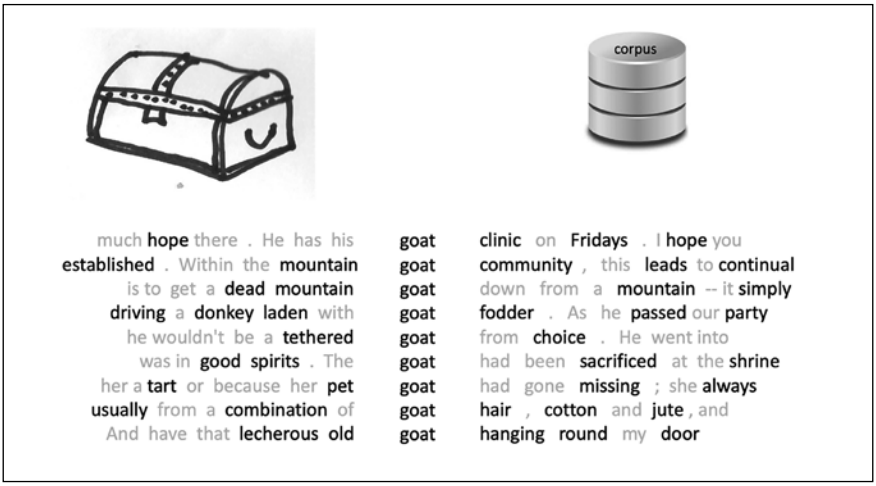


FIGURE 20

In the example I have shown, the vocabulary that we are using as the basis for our study is this set of farm words. In my native language German, there is a word for vocabulary that translates into treasure of words, "Wortschatz", which why you see the treasure chest here.

All elements should go into the treasure chest. Once we have the vocabulary, we can choose a corpus and retrieve concordances for all of our vocabulary items. In this case, we retrieve several concordance lines for the word *goat*. You see the keyword *goat* in the middle and a couple of words to the left and right. This data undergoes several processing steps. In the first processing step, grammatical words and other high frequency items are removed from the concordances. Those are the words that I have shown in grey on this slide. Pronouns like *he* or deictic elements like *there* or articles like *the*, all of those are deleted so that what remains are really just the contentful lexical items like *Friday*, *community* or *fodder* that form the context of *goat* to the left and right, and that presumably tell us something about the key word.

These are the words that the word *goat* collocates with.

All of these context items are collected in what we can call "a bag of words". Why is it called a bag of words? That is a technical term from corpus linguistics. It is called a bag of words, because in a bag all elements are mixed up and there is no linear structure to them anymore. In a sentence or in authentic language use, words follow each other in a linear way, but once you put them in the bag of words, you really only know how frequent they are, not how they are usually ordered.

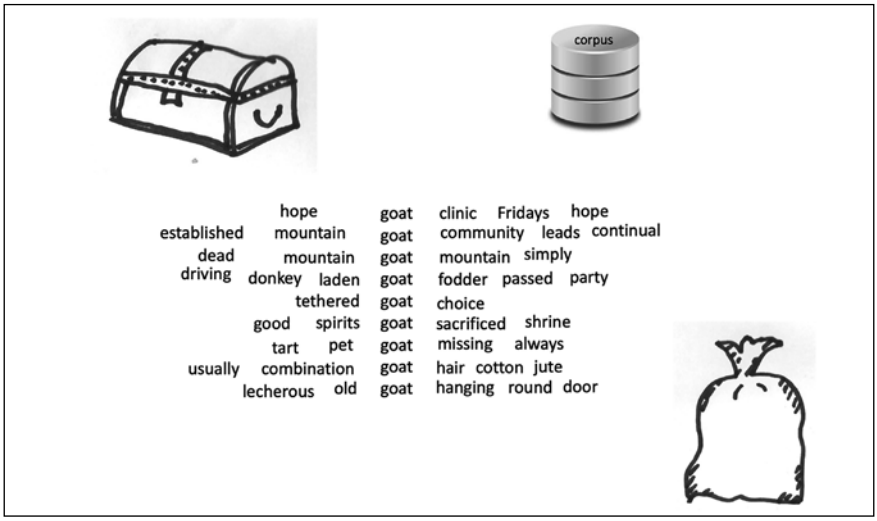


FIGURE 21

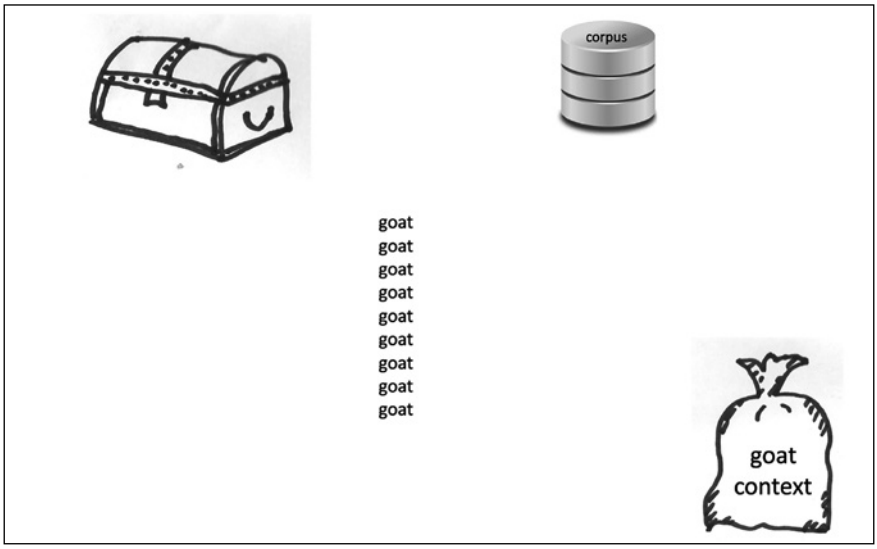
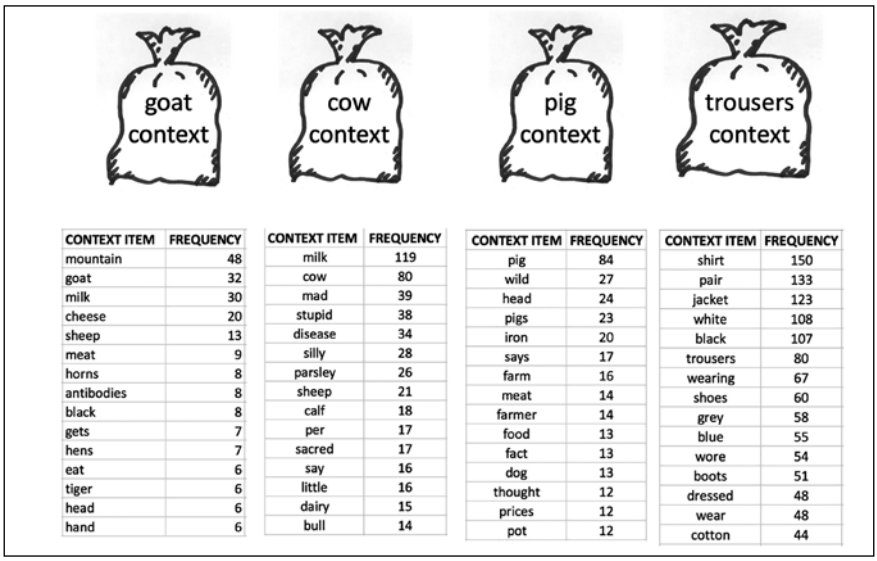


FIGURE 22

A bag of words like this can be easily transformed into a frequency word list. In the list on this slide, you see the most frequent elements that occur with the word *goat*. In the corpus that I used, the word *mountain* appears as the most frequent collocate. It is no coincidence that the word *goat* appears as a context item of itself. There is *milk*. There is *cheese*. The word *sheep* is a frequent



CONTEXT ITEM	FREQUENCY
mountain	48
goat	32
milk	30
cheese	20
sheep	13
meat	9
horns	8
antibodies	8
black	8
gets	7
hens	7
eat	6
tiger	6
head	6
hand	6

CONTEXT ITEM	FREQUENCY
milk	119
cow	80
mad	39
stupid	38
disease	34
silly	28
parsley	26
sheep	21
calf	18
per	17
sacred	17
say	16
little	16
dairy	15
bull	14

CONTEXT ITEM	FREQUENCY
pig	84
wild	27
head	24
pigs	23
iron	20
says	17
farm	16
meat	14
farmer	14
food	13
fact	13
dog	13
thought	12
prices	12
pot	12

CONTEXT ITEM	FREQUENCY
shirt	150
pair	133
jacket	123
white	108
black	107
trousers	80
wearing	67
shoes	60
grey	58
blue	55
wore	54
boots	51
dressed	48
wear	48
cotton	44

FIGURE 23

collocate. There are *meat*, *horns* and other items that we associate with goats. These frequency lists are the basic information the computer works with. They are representations, if you like, of the semantic profile of a word. We produce these frequency lists not only for one word that we are interested in, but rather for all the others that we have as well.

Here are three more frequency lists, one for *cow*, one for *pig*, and one for the word *trousers*. The reasoning goes that words with similar frequency lists should have some semantic relation between them. When we examine these frequency lists on the slide, we find, for example, that *goat* has the word *milk* as a frequent collocate. Also, *cow* has the word *milk* as a frequent collocate. *Goat* has *sheep* as a frequent collocate, and so does *cow*.

You get the basic point. Words with similar meanings should occur with similar collocates at similar frequencies. When we compare *goat*, *cow* and *pig*, we should see a relatively high degree of similarity, as opposed to contrasting *goat* with *trousers*. The most frequent collocates of *trousers* are *shirt*, *pair*, *jacket*, *white*, *black*, *trousers*, and *wearing*. No *meat*, no *horns* and no *milk*. It has a different semantic profile. Words with similar frequency lists have some semantic relation, which can be one of synonymy, antonomy, partonomy, and some other -onymies that would come into play there.

That is the first work step: We choose a vocabulary of key words and we retrieve frequencies of context items from a corpus. That is what I just described to you.

The second work step is an arrangement of these frequency lists in a table of frequencies that lists all the vocabulary items and how often they co-occur with all the context items. This slide shows the vocabulary items in the columns, the context items will appear in the rows. Let me show you how this works. We have our vocabulary items like *goat*, *cow*, *pig*, *trousers*, *jacket*, *boots* and others.

Then we have all context items from our bag of words. We have the table of vocabulary items and the context items.

In the cells of the table, we have the co-occurrence frequencies of each vocabulary item with each context item, that is, in the table, we have, for instance, the information that the word *goat* co-occurs with the word *cheese* 20 times. Some cells in the table are 0, because not all vocabulary items occur with all context items. The word *pig* in this sample does not occur with the word *grey*. *Goat* occurs once with *grey*. *Cow* also does not occur with *grey*. That tells you not only about what kind of words occur frequently with each other, but also where there is an absence of co-occurrence.

That information allows us to compare pairs of vocabulary items. We can use a mathematical measure to determine how similar or how different they are. We can carry out pairwise comparisons, subtract the values in the cells from each other and arrive at a measure of dissimilarity, a measure of distance of these items from each other. Earlier this morning I talked about how this can be done with complement-taking verbs that have different profiles with regard to the syntactic structures that they occur with. Here we perform the same analytical step, except that we are dealing with a table that is much larger. We

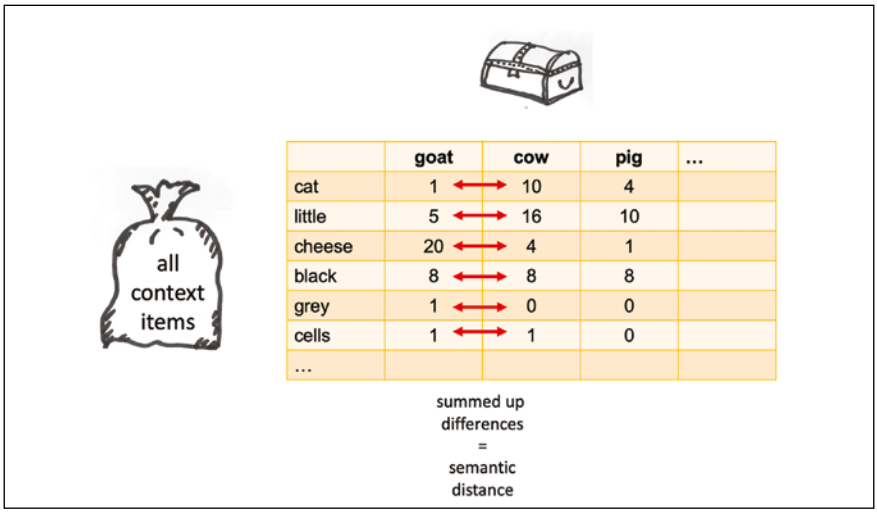


FIGURE 24

do not only compare five different linguistic units, but rather thousands of different words that are the context items of *goat* and *cow*.

We can take the absolute differences, we can sum them up and that would give us a measurement of how semantically distant or similar two items are. We would expect a larger difference for the pairing of *goat* and *trousers* because one is an animal and the other is a clothing item. We would expect a smaller distance measure for *goat* and *cow*, because both are animals.

So far, our table contains raw co-occurrence frequencies of vocabulary items and context items. These frequencies cannot be taken as such. They need to be transformed with a collocation measure. Let me explain why.

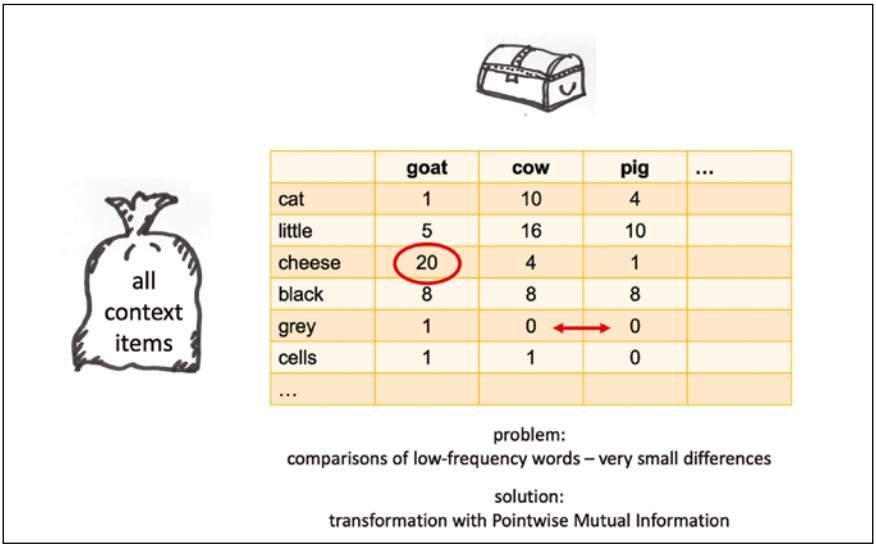


FIGURE 25

Comparisons between low frequency vocabulary items will involve many cells in this table that are 0. Let's say that our vocabulary contains two very rare verbs: *procrastinate* and *exempt*. Both of them have for many context items the value 0, which means that there are many cells where the two verbs are identical. This may yield the false impression that the two verbs are semantically very similar. However, *procrastinate* and *exempt* are not semantically similar. They're just similarly infrequent. That is why we need to perform a transformation of these raw frequencies in order to arrive at a more realistic assessment of how semantically similar or dissimilar two items are. This can be done with Pointwise Mutual Information, a measure of collocation which I would like to explain today in a little more detail.

Let me take a concrete example. In this table, we have 20 co-occurrence instances of *goat* and *cheese*. Is that more than what we would expect by

chance? Is that about what we would expect by chance? Or is this actually less than what we would expect, given the basic frequencies of the word *goat* and the word *cheese*? Pointwise Mutual Information can give us an indication of that.

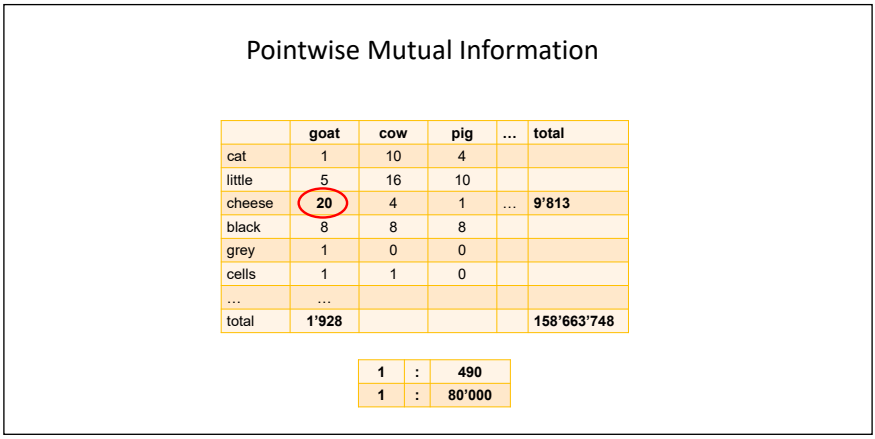


FIGURE 26

In order to compute a value of Pointwise Mutual Information, we need the marginal frequencies of this table here. We have all the instances of *goat*, which are about 2000. We have all the instances of *cheese*, which are about 9800. Then there is a very large number in the lower right corner of the table. That number indicates that we have about 160 million collocation pairs in the table. Point Mutual Information performs a comparison of ratios. Is 20 to 9800 about the same, or more or less than 1900 to 160 million? The ratio up 20 and 9800 is about 1:490. The ratio of 1900 to 160 million is 1:80,000. That means *goat cheese* is heavily over-represented in the data. This makes sense because we know that *goat cheese* is a lexical item that denotes a specific type of cheese.

The third work step is thus a transformation of the raw frequencies with Pointwise Mutual Information (PMI). This brings us to the fourth and final step, namely the visualization of the data. Once we have the table with its PMI values, we can transform that data into a visualization like the one I showed earlier with the farm words. That kind of data yields a display in which semantically close words are shown in very close proximity, and semantically distant words are further away from each other. Before we get back to asymmetric priming, there is one more issue that I need to explain. In order to answer our research question, we needed to turn to an extension of the general idea of distributional semantics.

In the graph you saw earlier, every item in the graph represented a word type, that is, there are several hundred concordance lines of the same word that

determine in the end where on the graph a certain word appears. If you want to compare lexical *going* and grammatical *be going to*, those two constructions actually have the same form *going*. That means that this type-based approach is only of limited use. However, there is an extension of the technique: token-based semantic vector spaces. That approach can reveal semantic differences between word tokens of the same type. Let's say that we have a number of key word tokens, each of which is represented by one concordance line. Through the use of token-based semantic vector spaces, these concordance lines can be compared against each other, and we can investigate differences in meaning of the same key word.

How can we do that? If we base the analysis only on the words that are contained in the concordance lines, we have very little material to go on, which would be problematic. We have to find a way to make the concordance lines as informative as possible. The solution is to use not only the collocates of the key word as such, but in fact collocates of collocates. Let me explain how this works.

What we need for that are basically two data resources. The first resource is a type-based semantic vector space of the kind that I showed you just a minute ago, the one with the *farm* words, except this one holds a much larger vocabulary. It contains about 20,000 key words and 20,000 context items. We thus have a large vocabulary that we have characterized in terms of their collocates. That is our first resource.

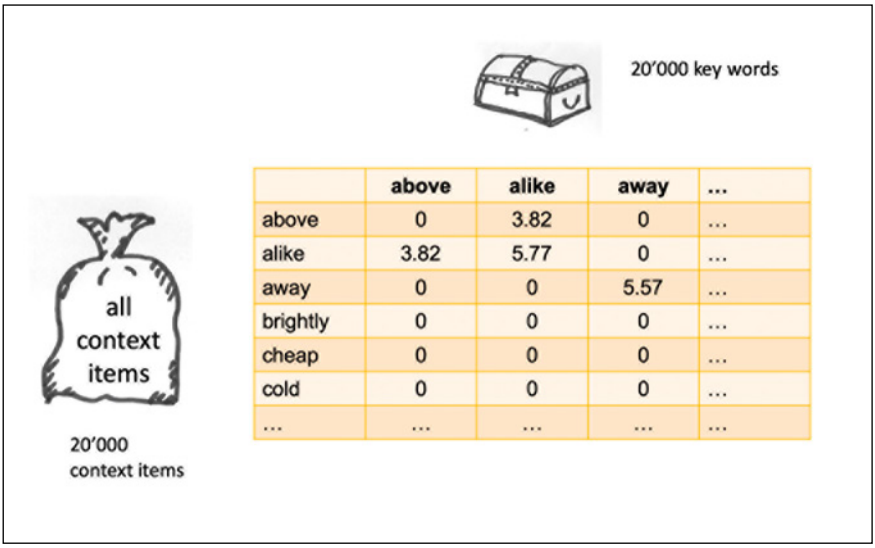


FIGURE 27

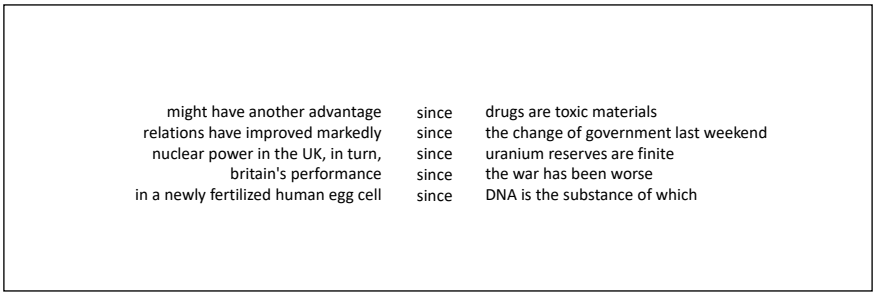


FIGURE 28

The second resource consists of concordance lines for a word that we would like to model. Here we have a short concordance of the word *since*. Let me just read the first example, "*might have another advantage since drugs are toxic materials*", and the second "*relations have improved markedly since the change of government last weekend*". What we are trying to do is to create a semantic vector space that gives us a measure of similarity between these concordance lines. How do we do that?

We proceed as we did before. We first exclude all the stop words: the pronouns, the articles, auxiliaries and prepositions. Then, in order to create a semantic representation of that very concordance line, we look up its context items in the type-based semantic vector space that we created earlier. For concordance line 1, we look up the context items *advantage*, *drugs*, *toxic* and *materials*. All of these words are represented in the type-based semantic vector

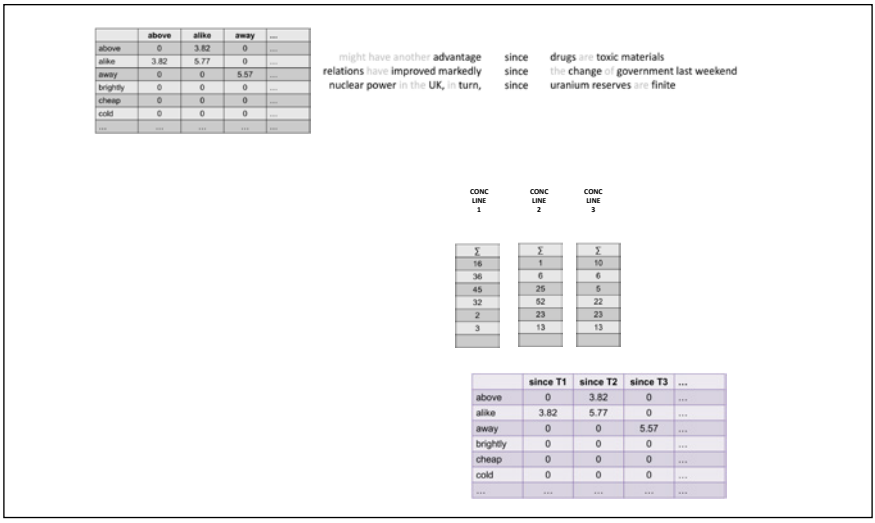


FIGURE 29

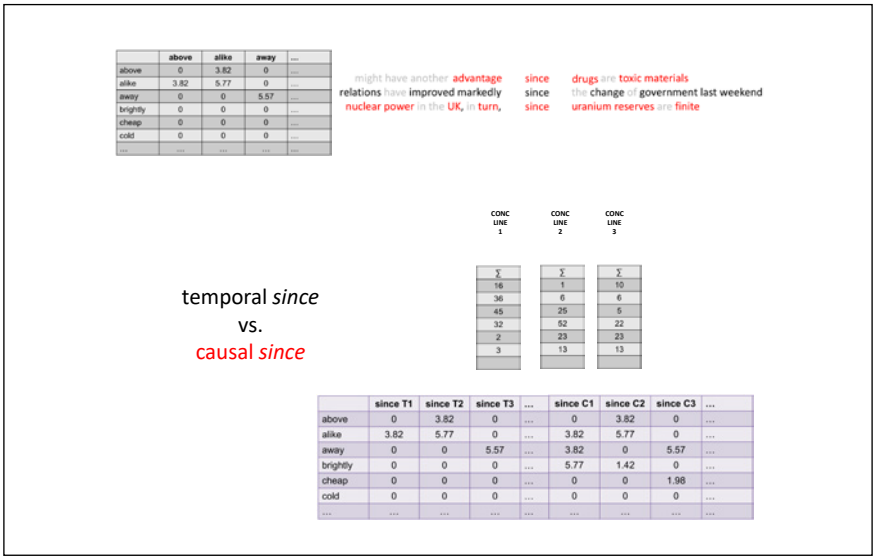


FIGURE 30

space. We just pull their context vectors out of the large vector space, and we create an average of their collocate vectors. That process yields a representation of our concordance line 1 in the analysis.

We then do the same for concordance line 2, for line 3 and for line 4, and so on and so forth, until we have a table that represents all lines of our concordances. With that kind of table, we can now create a visualization of a semantic space that is populated not by different word types, but rather by different concordance lines.

Let me tell you something about the three concordance lines you see here. They're all uses of *since*, but there is actually a semantic difference between the second example and the other two. The second *since*, "*since the change of government last weekend*" carries temporal meaning, and the first and the third express causal meaning. The semantic difference between temporal and causal meaning actually shows up in a visualization of a larger concordance of *since*.

In this graph here, you see little black dots and little red dots. The red dots are causal uses of *since* and the black dots are temporal uses of *since*. Each dot represents a concordance line. The positions of these dots in the graph have been calculated on the basis of their second order collocates, i.e. the context vectors of the words that are present in each concordance line. What you see in this graph is that there is some overlap. There is no perfect separation of temporal and causal *since*, but what you do see is that there is some structure in the graph. The red dots that represent causal *since* are more toward the right of the

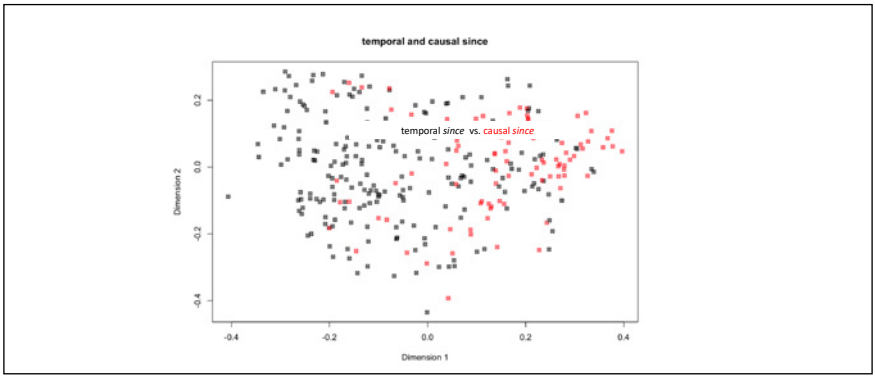


FIGURE 31

graph. On the left is a large area that is almost exclusively populated by black dots. In other words, with a token-based semantic vector space we can distinguish between two grammaticalized items, temporal *since* and casual *since*.

This is the point where I thank you for your patience, because this observation finally brings us back to asymmetric priming.

Data of this kind can be exploited to test the asymmetric priming hypothesis. Causal *since* is a case of secondary grammaticalization. Temporal *since* is the source that should, according to the asymmetric priming hypothesis, prime causal *since*, but not vice versa. According to the asymmetric priming hypothesis, if two instances of *since* follow one another, switches from temporal to causal should be more frequent, because time should make you think of causality but not the other way around. In every instance in which A turns into B, A should prime B, but not vice versa. The change from temporality to causality is exactly that kind of phenomenon.

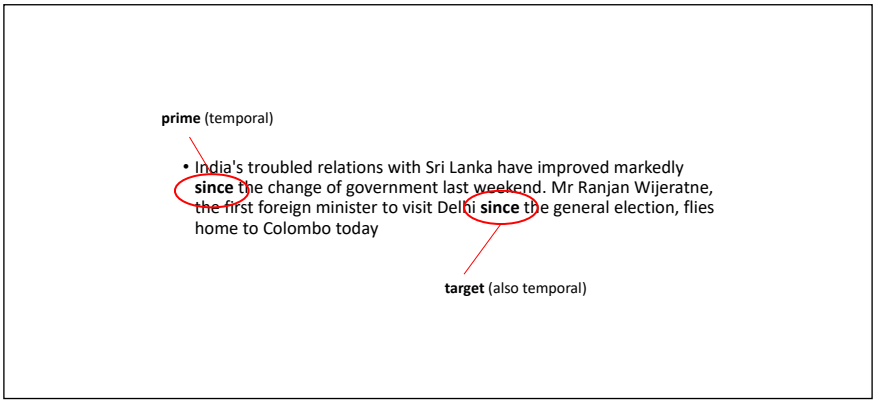


FIGURE 32

The red and black dots that you saw in the graph earlier came from examples where two *sinces* were following each other. We took corpus examples such as the one shown on this slide: "*India's troubled relations with Sri Lanka have improved markedly since the change of government last weekend*". That sentence contains the first *since*. The following sentence contains another token of *since*: "*Mr Ranjan Wijeratne, the first foreign minister to visit Delhi since the general election, flies home to Colombo today*". The example thus involves two instances of *since*, and in this particular example, both are temporal. The first *since* is what we call the prime. That example expresses time. The second *since* is the target, and it also carries temporal meaning. Both *sinces* in this cases express the same type of meaning.

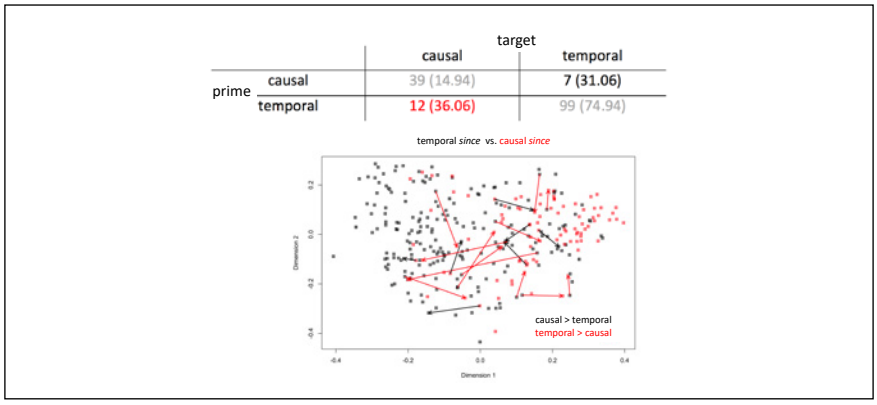


FIGURE 33

The question now is, when we have pairs like this, do we see switches from temporal to causal more often than we see switches in the other direction? This is something that you can easily count.

Let's first look at the table that you see right above the graph.

The numbers that you see are the observed frequencies. In brackets behind that, you see the expected frequencies. Let us examine the switches from temporal to causal. In our data, we observe 12 switches from temporal to causal, but we would have expected 36, given the overall distribution of causal and temporal *since*. The asymmetric priming hypothesis predicts that switches from causal to temporal should be underrepresented, which is confirmed by our data. But the most important information in this table is that most pairings are sequences from causal to causal and from temporal to temporal. We observe 99 instances of two temporal instances of *since*, but we would have expected only 74. We find 39 pairings that go from causal to causal, but we would only have expected 14.

A second piece of evidence follows from the graph with the token-based semantic vector spaces. In the graph, I have connected a number of pairs of data points with arrows. Red arrows indicate changes from temporal to causal, as predicted by the asymmetric priming hypothesis. Black arrows represent changes from causal to temporal. According to the asymmetric priming hypothesis, we would expect many more red arrows than black arrows. We would expect the arrows to pattern in a certain way, namely, we would expect them to start on the left side of the graph, which represents temporal meaning, to the right side of the graph, which represents causal meaning. That expectation is not met, as the arrows point towards different directions. This means that also our corpus-based analysis did not produce any evidence to support the asymmetric priming hypothesis.

You might be skeptical about the example of temporal and causal *since*, since it is just one example, and furthermore a case of secondary grammaticalization. In order to test that, we performed the same kind of analysis for other pairs of grammatical words and their lexical counterparts.

In this graph, we see a contrast of the lexical verb *used* and the grammaticalized habitual marker *used to*. The black data points represents lexical tokens, the red ones are habitual tokens. The analysis distinguishes nicely between them. The habitual data points show up at the left side of the graph and the lexical ones take up most of the graph. Again, when we look at the number of switches from lexical to habitual and from habitual to lexical, there is no significant difference between those. Neither the direction of the arrows nor the length of the arrows shows any difference, so there is no priming asymmetry towards habitual meaning.

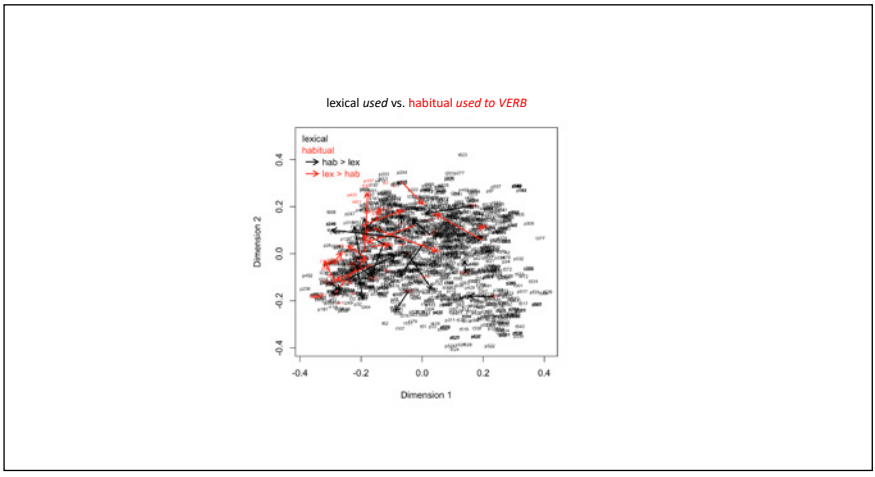


FIGURE 34

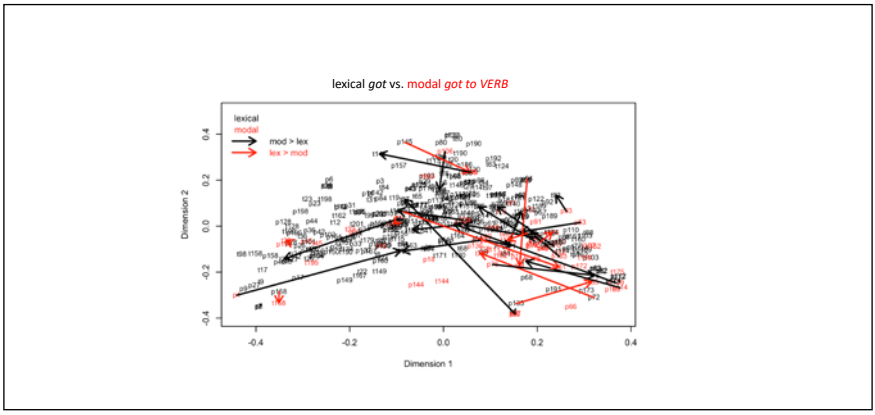


FIGURE 35

This slide shows another example. We performed the same analysis for lexical *got* versus modal *got to* and again the result is the same. There is no priming asymmetry towards modal *got to*.

In the last analysis, we compared deontic *may* as in "*You may now kiss the bride*" and epistemic *may* as in "*That may be a mistake*". The result is the same, no asymmetric priming effect.

What we do find in all cases though is that first, the semantic vector space allows us to discriminate rather neatly between the different meaning categories. Second, we find strong effects of within-category priming. Most pairings stay within their respective category, so that the sequences go from lexical to lexical and from grammatical to grammatical. This can be interpreted as a priming effect, but crucially not as asymmetric priming from lexical to grammatical.

With that I would like to come to my conclusions. The gist of the asymmetric priming hypothesis is that lexical *going* should prime grammatical *be going to*, whereas grammatical *be going to* should not prime lexical *going*. What actually happens is that lexical *going* strongly slows down grammatical *be going to*, but grammatical *be going to* does not slow down lexical *going*.

Why do we see this particular effect? The phenomenon of *horror aequi* suggests itself as an explanation. Processing the same form twice within the space of a few words is difficult, and speakers tend to avoid it. I have taken the liberty to construct a few sample sentences that we can subject to your intuition. The sentence "*The boys need new shoes that we need to buy*" involves two uses of *need*, the first is lexical and the second is grammatical. Here's another example, "*You need to make a list of things you need*". Again, two instances of *need*, but here the first is grammatical and the second is lexical. This strikes me as

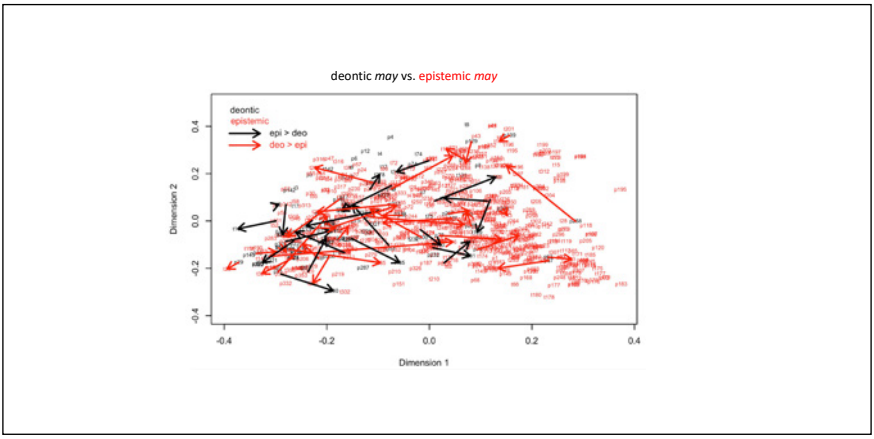


FIGURE 36

somehow better than the first one. If horror aequi has a role to play, then why should its effect be asymmetric? The best explanation that I can offer at this point is that semantic specificity may account for this.

When forms grammaticalize, their original lexical meaning fades. The more grammaticalized the form is, the less it will engender *horror aequi* effects. When I say "*My parents have had a dog*", which involves two instances of *have*, right next to one another, there is no *horror aequi*. Or consider "*I am going to go there*", in which two forms of *go* appear in close proximity, without any *horror aequi* effect. Compare that to "*He used to use a typewriter*", which is perhaps not terrible, but certainly not elegant, or "*It happened to happen on a Tuesday*", which sounds like someone deliberately played with language.

The first two are okay, because these are strongly grammaticalized and schematic constructions. The others are only weakly grammaticalized forms. Essentially, we believe that what we measured in the experiment was the degree of grammaticalization of our grammatical stimuli. For his dissertation, David Correia Saavedra has carried out a study of corpus-based measurements of grammaticalization. He has very interesting results that I hope you will find out more about soon.

Let me come to an end here. What we found in the corpus-based test of the asymmetric priming hypothesis is that lexical forms prime themselves, and grammatical forms prime themselves. We find self-priming of both lexical forms and grammatical forms, but no priming asymmetries towards the more grammatical variant.

In conclusion, both our experimental results and our corpus-based results detract from the asymmetric priming hypothesis. That is a little disappointing.

I still think the hypothesis is a fascinating idea. Very importantly, this negative result does not call into question the power of priming as a force on language use. It also does not call into question the basic tenet of usage-based approaches that cognitive processes operating in the here and now shape language structure and language change. It just means that asymmetric priming is not the explanation for unidirectionality in semantic change. With that, thanks again for your attention.

LECTURE 9

The Upward Strengthening Hypothesis

Welcome back to Ten Lectures on Diachronic Construction Grammar. This is the ninth lecture. In the last lecture, I discussed unidirectionality in language change and the idea that grammatical semantic change proceeds in one direction only, from more concrete, more specific meanings to more abstract meanings. The empirical studies that I discussed were designed to test whether this particular characteristic of language change could be explained through the cognitive process of priming. Since many processes of language change in grammaticalization are clearly asymmetrical, it is a very tempting idea that these processes of change are also triggered by a psychological mechanism that is asymmetrical, as in the case of priming relations that go in one direction only. I explained that the predictions of the asymmetric priming hypothesis ultimately could not be substantiated, but there is still reason to believe in the basic idea that historical processes of language change are to be explained in terms of cognitive processes that are at work in the here and now. When I talk to you, certain cognitive processes are active, and those are the same processes that are responsible for language change in the long run. This basic idea is also of interest for my lecture today. The title for this lecture is "The upward strengthening hypothesis". I have used this term in a paper of mine that was published in 2015 in the journal *Cognitive Linguistics*. I used it in order to combine several notions that are widely shared in Construction Grammar and cognitive linguistics in general, but that have not been put together in quite the same way that I thought would be insightful. Let me explain what I mean.

At the heart of the upward strengthening hypothesis is the notion that concrete usage events send activation through the network of construction that is your knowledge of language. As you hear me talk as I do right now, the words that you hear activate your mental representations of constructions in your knowledge of language.

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691269.

Let's take a simple example. When I say the word *hypothesis*, then the node in your network of constructions that corresponds to that word *hypothesis* is activated. When I say *the upward strengthening hypothesis,* there are four words that are activated, *the*, *upward*, *strengthening* and *hypothesis*, but at the same time, also a syntactic construction is activated, namely the "definite noun phrase" construction. *The upward strengthening hypothesis* instantiates a syntactic pattern and that pattern is a construction in your knowledge of language that is being activated at that point.

The upward strengthening hypothesis is thus based on the idea that all of your knowledge is a network of form-meaning pairs. It crucially relates to the concept of entrenchment. Mental representations of linguistic structures are strengthened if they are activated very often.

Over these past days, you've heard me pronounce certain words very often, as for instance *Diachronic Construction Grammar*. This particular phrase has been strengthened, it has been entrenched in your mind. The more often I pronounce a word or a construction, the more activation is sent to that particular node in your network of constructions, and the more entrenched that node becomes. The upward strengthening hypothesis relates to activation of your linguistic knowledge and the gradual entrenchment of that knowledge. That represents a broad consensus that we currently have in usage-based linguistics, but there is more to the upward strengthening hypothesis.

Another component of the hypothesis is what I presented as basic idea #3, the idea that constructions vary in terms of their degrees of complexity and schematicity. When you hear me say *the upward strengthening hypothesis*, you hear the concrete words, *the*, *upward*, *strengthening* and *hypothesis*, which are constructions that are not schematic and not very complex. Those words are activated in your mind, and they become a little more entrenched. But you also registered that these words combine to form a definite noun phrase, which is a constructional schema, a construction that is both complex and schematic. What happens to that schema when you hear a new instance of it? The idea, reasoning from the basic principles of usage-based linguistics, would be also that schema is activated and strengthened as you hear me pronounce that phrase.

Your knowledge of language is organized in different layers of complexity and schematicity that characterize the nodes of the constructional network. Schematic constructions are higher up in the network and specific constructions are further down.

The question that is inherent the upward strengthening hypothesis ultimately is this one: When I hear a construction, a simple one like the word *dog*, does the activation that runs through my knowledge of language only go to

that construction itself, or does it also strengthen more abstract constructions like the noun construction? Is there activation that starts with the *dog* and then flows up and strengthens a node that is higher up in the network?

Let's take another example. When I hear an expression such as *the hypothesis*, which is a bit more complex, as it consists of two words, does the activation go to that particular string only, or does it also go to the schematic construction that this string instantiates? That schema would be the definite noun phrase construction. We do not have to stop there. Does *the hypothesis* activate the definite noun phrase construction and in turn a general noun phrase construction? How far up does the activation go? That is the underlying question that I would like to address.

This question relates to one of the controversies that I mentioned on the first day, namely the controversy between the views of complete inheritance on the one hand and of redundant representations on the other. Just to refresh your memory, the relevant issue is whether speakers cognitively represent grammatical information just once, at the highest level of abstraction, or whether they represent it several times, across different layers of abstraction in the constructional network.

According to the view of complete inheritance, information is stored only once at the most abstract level, which is very economical and elegant, but which ultimately might not correspond to the psychological facts. The view of redundant representations, on the other hand, would hold that information is stored at several levels of abstraction, so that speakers remember forms that they technically they would not need to remember, such as the plural noun *cats*.

On the view of complete inheritance, the answer would be that cognitive strengthening affects all layers up to the most abstract construction, because hearing a phrase like *the hypothesis* would require you to look up certain pieces of information that are stored in the definite noun phrase construction. What does it mean to be definite? That is information that is stored in the definite noun phrase construction. When you hear *the hypothesis*, you would further need to access the general noun phrase construction in order to look up information such as that noun phrases can be subjects or objects in the clause. All of that information is only stored once, and as a hearer, you have to activate the respective level to find that information.

Let me say a few words about inheritance as it is normally understood in Construction Grammar. This concerns the relations between more abstract constructions and more concrete constructions.

Earlier, I have given you the example of an idiom such as *kick the habit*, which inherits information from transitive *kick*, which inherits information

from the transitive construction, which further inherits information from the general verb phrase construction. On the complete inheritance view, lower-level constructions do not redundantly represent information that is inherited. Information that can be looked up does not have to be stored. By contrast, on the view of redundant representations, low-level constructions have rich representations.

Crucially, these two views of inheritance have different implications for the upward strengthening hypothesis. Namely, when you hear *the hypothesis*, according to the idea of complete inheritance, you have to access the definite noun phrase in order to look up features like definiteness, and you have to access the general schematic noun phrase construction in order to look up features that are common to all noun phrases. In other words, with regard to the upward strengthening hypothesis, the view of complete inheritance would imply that every time you hear a construction, activation spreads through the entire network all the way up to look up the most general features of that construction. Specifically, when you hear *dog*, also your representation of the noun construction becomes a bit more entrenched. When you hear *the hypothesis*, the general noun phrase construction becomes a little bit more entrenched. By contrast, the view of redundant representations would allow you to stop at a certain level, which already represents all the relevant features redundantly, so you hear *dog*, but nothing happens to the noun construction. *Dog* is a very frequent word that you have stored, so no further activation spreads to the noun construction as such.

I need to come back for a minute to the idea of constructionalization, which I want to contrast with the idea that I call upward strengthening. Just to remind us, constructionalization is the creation of a new form-meaning pair in the network of constructions. Constructionalization requires the repeated exposure to concrete tokens of language use. You hear a certain type of expression again and again, which then leads to the formation of a new abstract node in the network of constructions.

What I would like to stress is that constructionalization and upward strengthening are not the same phenomenon. Grammatical constructionalization is the establishment of a new node in the network, and upward strengthening is concerned with its subsequent entrenchment through the experience of language use. Crucially, the experience of a linguistic unit may strengthen not only a representation of that unit itself. When you hear a word or an expression, not only that expression itself, not only that construction itself may be strengthened, but also the more abstract schema that licenses that expression in the first place. The claim that I want to make in the following is that upward strengthening is a process of grammaticalization.

Let me also stress that upward strengthening is not the same as entrenchment. Entrenchment, as it is typically understood, means that you hear a linguistic unit and the representation of that linguistic unit is strengthened in your mind. What I would like to discuss is how abstract patterns become entrenched.

Abstract patterns are not something that you hear. You hear strings of words, and the more you hear a certain string, the more routinized that string becomes, and the better you know it. But what about abstractions of these strings? Let's take an example. Take the expression *kind of funny*, which consists of three words that frequently occur together. Over time, the string may become entrenched in your mind, so that you recognize it as an idiomatic phrase of English. That string instantiates a more abstract schema, namely *kind of* ADJ. Speakers can describe entities and events as *kind of sad*, *kind of complicated*, *kind of expensive* and so on and so forth. How does that abstract pattern become entrenched? *Kind of* can also appear with other types of phrases. The final element does not have to be an adjective. In the utterance "Well, he's kind of a jerk", the final element is a noun phrase, not an adjective.

Let us take a concrete example. You hear *kind of funny* and that speech event sends activation to that very expression. That exact string becomes more strongly represented in your mind. Let's for the moment assume that this speech event also passes on a certain measure of activation to a more abstract schema like *kind of* ADJ. That would mean that this kind of phrase higher up in the network would become a little bit more entrenched. Going even one step further, if this the *kind of* ADJ schema is activated and passes on part of its activation to an even more abstract schema, that would lead to the entrenchment of an abstract constructional schema in your mind.

I have already mentioned the conflicting predictions of the complete inheritance view and the view of redundant representations. The view of complete inheritance invites an understanding of upward strengthening that I would like to call the "naive" strengthening hypothesis. What that would mean in practice is that every time you hear an expression, activation would spread upward, as far as it can possibly go. Let us take the expression *You drive me crazy* as an example. On the naive strengthening view, each component part of this would be strengthened, so *You drive me crazy* sends activation to each word that it contains: *you*, *drive*, *me* and *crazy*. As a construction, *You drive me crazy* would further be assumed to send upward activation to the *drive crazy* construction, which is more abstract. That construction does not only have *you* as a subject. It also allows for any kind of object, not just *me*. On the naive strengthening view, we do not redundantly store information at lower levels of representation. The idiomatic *drive crazy* construction inherits information from the

resultative construction, SUBJ VERB OBJ ADJP. That construction licenses not only "*You drive your parents crazy*", but also "*You hammer the metal flat*" and "*He painted the wall blue*". In order to process and understand *You drive me crazy*, we thus need to look up some information at the level of the *SUBJ drive OBJ crazy/nuts/insane* construction, which we can only process and understand if we look up some further information at the level of the SUBJ VERB OBJ ADJP construction. All of this triggers a surge of upward activation, and it does not stop there either. In order to process the SUBJ VERB OBJ ADJP construction, the hearer has to look up information that is represented at an even higher level of abstraction, namely, at the level of the Subject Predicate construction, which tells us that subjects can occur with verb phrases and that this yields a complete sentence of English. The complete inheritance view makes a very strong prediction here and I would argue that it is a questionable prediction. Do we really have to access the Subject Predicate construction, a very abstract schema, if we hear someone say, *You drive me crazy*?

So what is wrong with the naive strengthening view? I think there are some examples that illustrate why we should be skeptical. There are some linguistic units for which it is doubtful that they strengthen any constructions that are more abstract. An example for that would be linguistic units that simply do not have overarching categories, so that there is no abstract schema to send activation to. Let me give you an example from my native language, German. When a child sneezes, the parent may comment that by pronouncing the word *hatschi*. That is an onomatopoeic word that echoes the sound of a sneeze. In English, you might respond to a sneeze by saying *bless you*. The social routine is similar, but the words are not the same. I would say that *hatschi* is a word in German, but do not ask me what kind of word it is. Is it a noun? There are perhaps arguments for that analysis, but in any event *hatschi* is not a very prototypical noun. When someone says *hatschi*, I have grave doubts that hearers will activate of general noun construction in German.

Let's take a different example of that sort, *hallelujah*. What kind of word is *hallelujah*? You might say that it is an exclamation. There are others like *damn it*, which consists of a verb and a pronoun, but which is not really a sentence either. These are problem cases. These are linguistic units that are impossible to categorize. Whenever a word is very difficult to categorize, I find it highly unlikely that we look up some kind of information that is stored in a more abstract category to understand that very item.

De-categorialized units represent another type of example. When lexical items grammaticalize, they shed some of their original category features. The word *long* is an adjective in "*as long as a python*", and a part of a clause connector in *as long as you do not get caught*. When I say *as long as you do*

not get caught, do hearers activate and strengthen the English adjective? In *as long as*, the adjective *long* has been de-categorialized. It no longer instantiates its upward link with the category adjective, so it can't send any activation that way.

The last type of example I'd like to mention concerns highly frequent constructs, as for instance the phrase *Oh my God*. That phrase contains the noun *God*, but it is debatable whether *Oh my God* sends activation to the English noun construction. I do not think so, because *Oh my God* has become entrenched as a holistic unit. When a string of words becomes entrenched as a holistic unit, the parts become increasingly hard to access by themselves.

All of this, in my view, casts doubt on the complete inheritance view. I think it is much more plausible to assume that information is stored locally, redundantly in the case when there are still connections to the higher levels, but non-redundantly if there are no connections, if those connections are cut as in the case of de-categorialized or highly frequent units.

Let me ask the question the other way around. Which structures do trigger upward strengthening? There are three points that I would like to make here. Expressions that you hear trigger upward strengthening if three conditions are met. Linguistic units require you to activate higher levels in the constructional network if they contain a strong cue for an overarching generalization, if the construct is infrequent and if the construct is not very similar to already known instances of the overarching generalization.

I would like to illustrate this point with another non-linguistic example. Let's say that you observe an unusual cat. Cats can be seen every day and they all look fairly similar, with fur, a tail, and four paws. Suppose that you see one that has no fur, ears that look a bit bigger than usual, and eyes that are bright blue. This kind of stimulus would force you to reconsider a category that you have, an abstract schema of cats in this case. That is the theoretical foundation of what I will have to say about the upward strengthening hypothesis.

Since it is a hypothesis, I want to test it. I want to test the idea that stimuli of this kind actually have the effect of strengthening an existing category. I will do that on the basis of diachronic corpus data. There is a particular case study that I want to discuss, namely, English noun-participle compounds. I have mentioned that construction type a few times already. It is instantiated by forms such as *doctor-recommended* or *chocolate-covered*.

There are three questions that I would like to ask with regard to these noun-participle compounds. First of all, how have they changed chronically? It has been pointed out that this construction, noun-participle compounding, is related to the passive construction in English. I wanted to know whether the two constructions have changed in comparable ways. Are the developments

in the compounding construction related to changes that have been going on in the passive? Finally, I want to use the empirical findings that emerge here for theoretical considerations that concern the upward strengthening hypothesis. Specifically, I will examine whether upward strengthening can be profitably viewed as a process grammaticalization, and if not, what else this phenomenon might be.

Let me give you a quick overview of the rest of this lecture. First, I will discuss how we can imagine the relation between noun-participle compounding and the English passive. Based on that discussion, I will present findings from a diachronic corpus analysis that explores how the compounding construction has changed and whether these changes are paralleled by changes in the passive. Then I will raise the question of whether what we see here can be brought under the umbrella of grammaticalization.

What are English noun-participle compounds? Huddleston and Pullum (2002) in *The Cambridge Grammar of English* describe this construction and they offer examples such as *drug-related*, *home-made*, *safety-tested*, and *taxpayer-funded*. The construction is a very productive, very common pattern of English grammar. Huddleston and Pullum state that these compounds generally correspond to syntactic passives with a prepositional phrase. *Drug-related* can be paraphrased as *related to drugs*, *home-made* means *made at home*, *safety-tested* means *tested for safety* and *taxpayer-funded* means *funded by taxpayers*.

The noun in this construction can instantiate the agent of an action that would be represented by the *by*-phrase of a passive construction. However, there are other roles. In *taxpayer-funded*, the noun instantiates an agent, so something was funded by the taxpayer. The noun can also instantiate a location. The word *home-made* does not refer to something that was made by the home, it refers to something that was made in the home or at home. In *drug-related*, the noun expresses a cause, so something is caused by drugs, and in *safety-tested*, the noun expresses a purpose, so something was tested for the purpose of safety. This suggests that there are very few constraints on the roles that can be expressed in the noun of a noun-participle compound.

However, there is one rather fundamental constraint, which is identified by Bauer and colleagues (2013) in the *Oxford Reference Guide to English Morphology*. Bauer and colleagues state that "The first element cannot receive an object interpretation." In the compound *doctor-recommended*, the doctor would be the subject of an action. In *arsenic-exposed*, someone was exposed to arsenic. The noun would correspond to a prepositional object of the verb, not its subject. Then, there are examples that do not work, as for instance *lunch-eaten*. You cannot say "**The lunch-eaten, participants returned to the conference*",

even though that would be a meaningful concept. I mentioned this in the first lecture that constraints of this kind allow you to identify that something is a construction. Here we have an example of this. There is a constraint that we can call the no-object constraint, which would be one piece of evidence suggesting that what we are looking at here is really a construction, something that you have to learn. I hope you agree that the no-object constraint is something of a puzzle. Why is it that the direct object is the only grammatical relation that does not work in this construction? Some work has been done on this particular problem.

One explanation has been offered by Rochelle Lieber (1983), who argues that the no-object constraint in noun-participle compounding is the same phenomenon that we see in the passive, namely that an object is promoted to the role of the subject. Consider an active sentence such as "*The dog bit the man*". In the passive, *the man* appears in the subject position, "*The man was bitten by the dog*". Lieber (1983) argues that this is something that we also see in noun-participle compounding. She argues that the past participle "expels" or drives out the direct object from its ordinary position. This happens in both the passive and a noun-participle compounding. In the passive, the object of a transitive verb is no longer part of the verb phrase in which it originally appears. When I say "*The strawberries were picked by hand*", the *strawberries* appear outside the verb phrase. The verb *pick* has a direct object, *you pick the strawberries*, so the *strawberries* are part of the verb phrase in a canonical active sentence. In the passive, the object no longer appears in its original phrasal environment. Instead, it appears as the grammatical subject.

In noun-participle compounding, we see something very similar. The object of a transitive verb is no longer part of the compound phrase itself. I can say something like *the hand-picked strawberries*, where the *strawberries* are no longer part of the construction that contains the verb. This is parallel to the passive. In both cases, we have an object that is evicted, that is driven out of its place of origin.

Lieber (1983) presents a syntactic tree that is meant to illustrate this. She posits argument structure traits that are thought to move upwards in the syntactic tree and that cause the no-object constraint. The governing node imposes selection restrictions, which determine what can occur in this kind of phrase. What it states is that there should be no direct object in its scope. Cognitive linguists may view this syntactic representation with some skepticism, but actually the idea of inheritance makes use of the same idea. Inheritance means that abstract schematic construction impose restrictions on concrete constructs. Generative linguistics and construction grammar are really not that far apart from each other in that particular respect. Both the passive construction

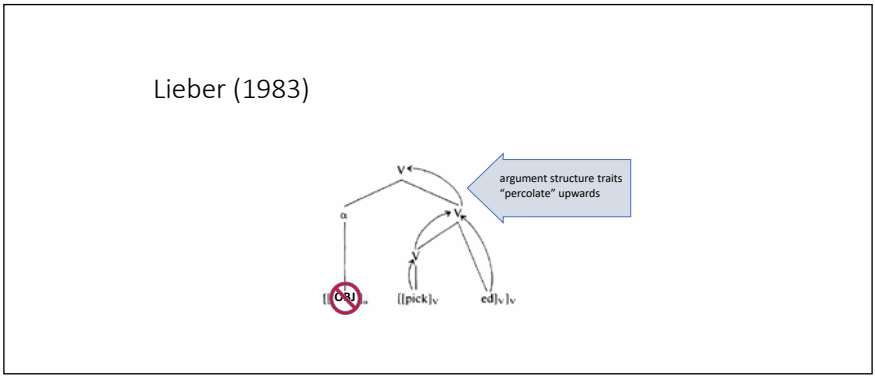


FIGURE 1

and noun-participle compounding inherit the argument structure of the past participle, which prohibits any direct object in its domain. But what are the implications of this?

What does it mean that both noun-participle compounding and the passive inherit the no-object constraint from the participle? It could mean that there is a higher-order schema, i.e. an overarching generalization. It could also mean that speakers treat them as two separate constructions, so that both constructions inherit characteristics from the participle, but are still represented individually.

In earlier lectures, I have talked about the controversy of higher-order schemas. Before we return to these two hypotheses, let us look at empirical data.

For this analysis, I used the COHA one more time. I searched the COHA for noun-participle compounds. I implemented several search criteria that gave me the examples that I was looking for.

The forms that I retrieved were *able-bodied*, *age-old*, *absent-minded*, *air-conditioned* and so on and so forth. Not all of them are target cases. For example, *old* is not a participle, so I weeded out those cases manually. I ended up with a dataset that consists of about 31,000 types, that is, different noun-participle compounds such as *god-abandoned*, *tax-abated*, *self-absorbed*, *man-abused* and so on. Across those 31,000 types, the dataset contains about 150,000 tokens. Many types are only represented in one or two decades. What can we learn from this dataset?

A first observation that I would like to share with you is that the construction quadruples in text frequency over the past two centuries. Noun-participle compounding is a success story in the recent history of English grammar.

This increase in token frequency is accompanied by an increase in the variety of types that are used. There are more and more different noun-participle

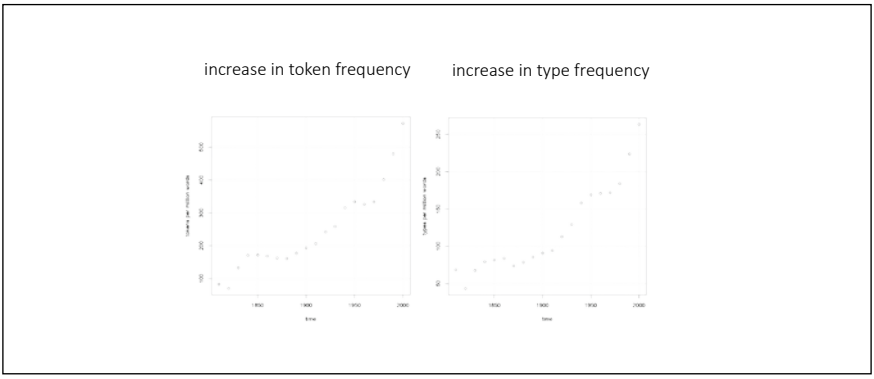


FIGURE 2

compounds that people are using. In other words, the construction is developing. In the rest of the time, I will be concerned with the question of what exactly this development is.

How does the construction change with regard to the participles that are found in it? I was interested in the following questions. How does the increase in type and token frequencies come about? Which are the participles that carry the increase?

The types that I retrieved from the COHA form families of different sizes around the same participle. A small family of this kind is arranged around the participle *wetted*, "made wet by something". The database contains compounds such as *dew-wetted*, *gall-wetted*, *snow-wetted* and *tear-wetted*, "made wet by tears". You understand the concept, but I very much doubt that you've heard this word very often. The participle *yellowed* occurs in the compounds *age-yellowed*, *fear-yellowed*, *opium-yellowed*, or *smoke-yellowed*. This is a slightly larger family. A large family is illustrated by the participle *coated*. Compounds with that participle include *aluminum-coated*, *bearskin-coated*, *beech-coated*, *blood-coated*, *candy-coated*, *caramel-coated*, *carbon-coated*, *cement-coated*, and many others.

I would like to show you how the noun-participle compounding construction changed with regard to its participle families. This graph shows the participle families that were most frequent in the early 19th century. Each bubble represents a participle family, and the bubbles are positioned with respect to their normalized type frequency on the *y*-axis, and with regard to their combined token frequency on the *x*-axis. If a participle appears high up, that means that it occurs with many different nouns. If a participle appears far to the right, it means that there are many tokens that involve that participle, regardless of how many types there are. Bubble size also corresponds to token frequency, so

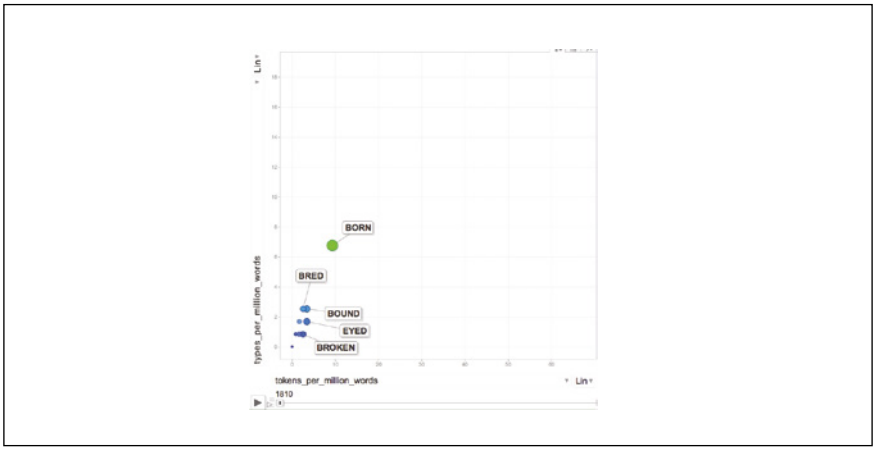


FIGURE 3

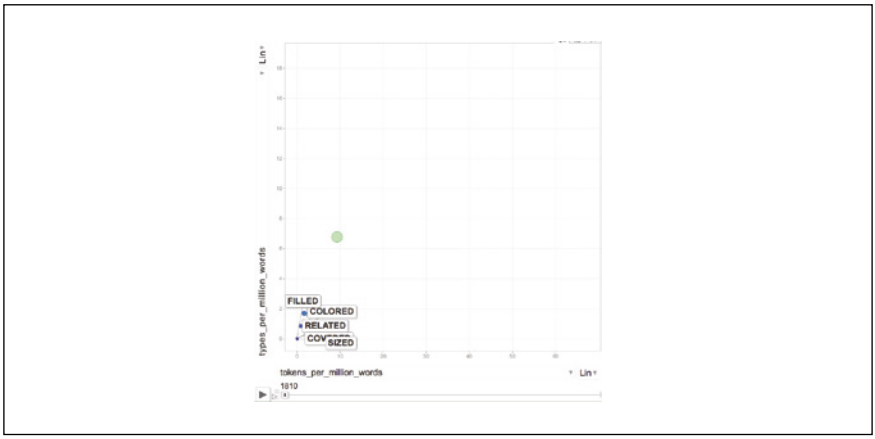


FIGURE 4

large bubbles account for many instances in the corpus. Among the participles with the highest type and token frequencies, we find *born* as in *Alaska-born*, *bred* as in *home-bred*, *bound* as in *Britain-bound* and then *eyed* and *broken*, as in *eagle-eyed* or *heart-broken*. We'll talk about those forms in a minute. For now, what I would like you to do is watch how these participle families have developed over the years. We start in the year 1810 and finish in the first decade of the 2000s.

In the first decades, there are small and unsystematic changes, up to the beginning of the 20th century. Then there are several participles that become more frequent. In the second half of the 20th century, there is one participle

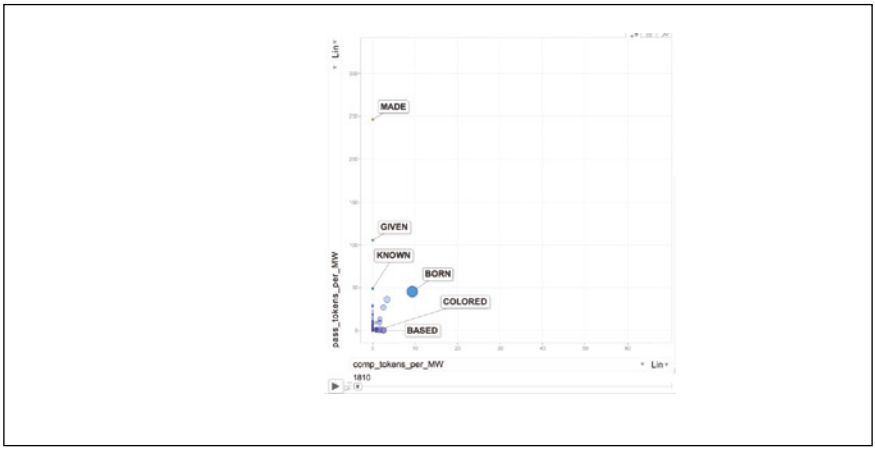


FIGURE 5

family that breaks out of the field and makes its way up to the upper right, where it has many different types and is represented by very many tokens.

Let's look at the same development again. This time, I have marked up the participles that turn out to be very successful in the late 20th century. Again, we see some back and forth in the first decades. Then *colored* and *shaped* emerge and stay in the vanguard of this field for quite a while. Then in the 1950s and 1960s, we have a few late-comers that join the field. By the 1970s, it is the participle *based* that escapes from the rest of the field and emerges as the most frequent participle in terms of types and tokens. You see that it draws the entire field with it to a certain extent. If we compare the 1810s to the 2000s, we see a clear difference.

Let me summarize the development. The participles change. In the 1810s, frequent participle include *born*, *bred*, *bound*, and *eyed* as in *Harvard-bred* or *context-bound*. In the 1900s, *colored*, *shaped* and *covered* dominate. In the 2000s, frequent participles are *based, related* and *sized*, as in the compounds *Houston-based*, *work-related*, and *toddler-sized*.

As I said earlier, one of my leading questions was whether or not noun-participle compounding would show diachronic developments that are similar to those of the passive. If for instance, passive uses of the verb *relate* increased diachronically, then the increase of forms such as *work-related* would be just an epiphenomenon of the history of the word *relate*, not a consequence of anything happening to the compounding construction.

How does the compounding construction compare to the passive? I went back to the COHA and retrieved examples of the passive, specifically passages where a form of *to be* is followed by a past participle. This covers only a subset of all passives, but a substantial subset. I retrieved three million examples that

fall into about 5000 different participle types. I identified participle types that occurred both in noun-participle compounding and in the passive, and visualized that data in order to see whether the overlapping types show similar frequency developments.

This chart shows the participle types arranged according to their token frequency in the passive, which is shown on the *y*-axis, and token frequency in the compounding construction, shown on the *x*-axis. Color represents passive token frequency, and size represents compound token frequency. Before we watch what happens here, let me discuss what I expected to see. If the passive and the compounding construction behave in similar ways, then the participles should move one step on the *x*-axis and one corresponding step on the *y*-axis. We should see diagonal developments in some way or other, either both increase or both decrease. If on the other hand, both constructions develop in relative autonomy, we should see movements that are either just horizontal or vertical.

The result is the latter. There are clearly no diagonal developments. Most of the changes that you see are either on the horizontal plane or on the vertical plane. The horizontal movement of *based* is clear to see. We further see a vertical fall of *made*, which indicates a frequency decrease in the passive.

The data indicate that we are observing independent developments. Changes in noun-participle compounds are independent of changes that happen in the *be*-passive, and the participle types that stand out most in the history of noun-participle compounding do not correspond to passive sentences. Sentences like "*The company is based in Houston*", do not have an active counterpart. "**Houston bases the company*" is not a grammatical sentence of English. The same goes for "*The problem is related to drug abuse*" or "*The car is sized just right*". This speaks for hypothesis B that I presented earlier. It suggests that there is no higher order schema for noun-participle compounding and the passive, but instead two independent generalizations.

With that, I would like to come to the theoretical part of my argument. Is upward strengthening grammaticalization? In earlier lectures, I used Hopper and Traugott's (2003) definition of grammaticalization: "*The change whereby lexical terms and constructions come in certain linguistic contexts to serve grammatical functions, and, once grammaticalized, continue to develop new grammatical functions.*"

Having adopted this definition, let me elaborate on grammatical functions and what they are. I have discussed procedural meaning before, which expresses who did what to whom, when something happens and how the elements of clause are related. I view noun-particle compounding as an argument-structure construction, very much like the ditransitive construction that Adele Goldberg has worked on. Goldberg (1995: 39) has formulated

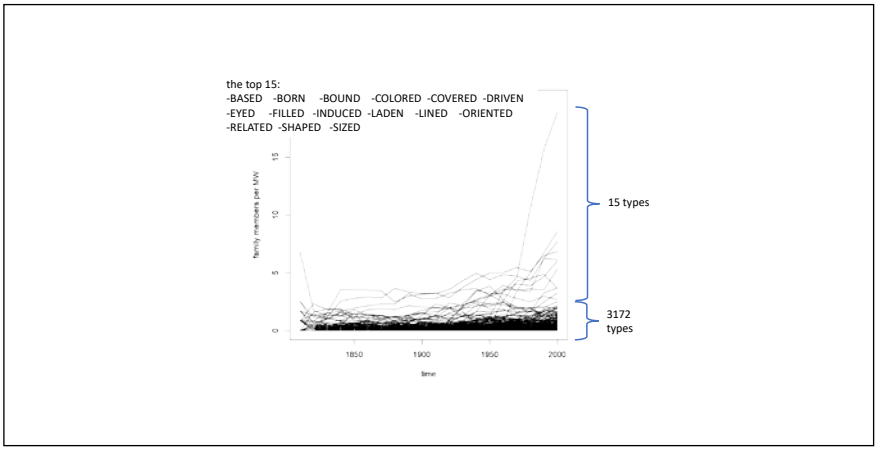


FIGURE 6

the "scene-encoding hypothesis", the idea that basic morphosyntactic patterns correspond to basic recurrent patterns of human experience. I think the meaning of noun-participle compounding fits reasonably well into that idea.

I have shown that noun-participle compounding has undergone dramatic frequency increases. We have a pattern with a grammatical function that is apparently on the move. Does that mean that we are looking at a process of grammaticalization? I am actually not convinced that this is the case, and I would argue against it in the following way.

This graph visualizes the development of the participle families in noun-participle compounding in a static way. It is based on the same data that you have seen before. Each line is a participle family, and you can see for each decade how prolific that family is. The sharply increasing line at the end represents the participle *-based*. What is interesting is that the terrific frequency increase of the construction is carried by a very small set of participle families, 15 families to be exact, *-based*, *-born*, *-bound*, *-colored*, *-covered* and a few others. The other 3000 types are the many infrequent participle families that did not contribute substantially to the frequency increase. This leads me to formulate the following points.

The increase in type and token frequencies is not carried by a broad base of different lexical items, only by a small "elite". The developments do not show systematic relaxation of constraints, so the no-object constraint is still there. Agents, instruments, causes, etc. are found in the noun slot at all times. When we consider morphosyntactic form, there is no formal change in the pattern as such. There is host-class expansion to some extent, but not much either. If what we see is not grammaticalization, it is reasonable to ask what else it might be.

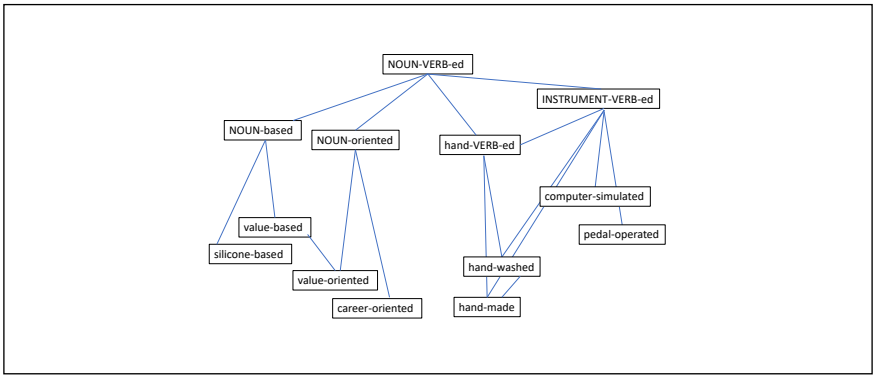


FIGURE 7

I view this as an instance of constructional change. Noun-participle compounding is a construction, a generalization that speakers make about form-meaning pairs in language use. Over time, this construction changes. Certain patterns become relatively more prominent and "emancipate" themselves. We can say that a pattern like *NOUN-oriented* emerges, and we have words such as *relationship-oriented* or *career-oriented* that illustrate that. Others sub-patterns decrease in relative prominence and become less productive. There is a participle in English with the form *stricken*. It is found in compounds such as *panic-stricken*, affected by panic, and *woe-stricken*, affected by emotional pain. That participle is no longer productive, so that node in the network has faded away. I like to think of this as change in a network of constructional patterns, where a lot of strengthening affects lower levels of the network, but crucially, there is no strengthening of the top node, so that the general noun-participle construction does not receive any upward strengthening.

This slide shows the kind of network that we can imagine. At the top of the network, we have the very abstract generalization of noun + participle. Then there are more concrete instances of this, like *Noun-based* or *Noun-oriented*, or *hand*-participle like *hand-washed*, *hand-made*, or *Instrument*-participle like *computer-simulated* or *pedal-operated*. There are generalizations at all levels of abstractions, and some are strengthened, some are weakened. Overall, this highest node does not become any stronger over the course of the time period that I have investigated.

One last point that I want to address is how we can model that kind of change in a constructional network. Diachronic corpora show us which instances of a construction are used during a given period of time. Changes in the instances that come in lead to strengthening of subconstructions and potentially the overarching construction. Let me come back to the example of the strange-looking cat. Each usage event potentially can cause some upward

strengthening, and the amount of upward strengthening depends on cue validity, frequency, and salience.

Cue validity relates to the following question: Is the usage event easily categorized as a combination of noun and past participle? If categorization is easy, then there would be more upward strengthening. Something like *chocolate-covered* where *-covered* is clearly recognizable as a participle, this sends a lot of activation. We can contrast that with the participle *-stricken*, which is not easily recognized as a participle. As a consequence, there would be less activation and less upward strengthening. The form *-eyed* is not even a true participle, so there would be even less activation.

Frequency captures how often hearers encounter linguistic usage events. I have discussed patterns like "*Oh my god*" or "*Damn it*", or "*Thank you*". Frequent forms emancipate themselves from overarching schemas, which leads to less upward strengthening. Conversely we can argue that the lower the frequency, the more upward strengthening should occur. Take an example such as *garlic-infused*. The concept is clear, but the word itself is infrequent. That sends a lot of upward strengthening. *Aluminum-coated* is more frequent, so it would trigger less upward strengthening. The compound *oil-based* is even more frequent and would be hypothesized to trigger an even smaller amount of upward strengthening.

Salience relates to the similarity between a usage event and other usage events of the same category. A compound such as *oven-roasted* would be similar to the form *oven-baked*. It is not very unusual in that regard. The reasoning would be that more unusual items send more activation to the category. The compound *war-acquired* is not very similar to previously seen compounds, so it sends more activation to the overarching construction. The compound *bacon-wrapped* is not so unusual and would therefore not trigger as much upward activation.

I computed a measure of upward strengthening over time on the basis of the COHA data. In each decade, constructs appear in language use. These constructs create the maximal amount of upward strengthening if they are new, if they contain a recognizable participle, and if they are relatively low in frequency, and if they are semantically different from the rest.

How did I define semantically different? This relates to concepts from distributional semantics that I discussed yesterday.

I created a semantic vector space for the participles in the database, which left me with a semantic space like the ones that you saw yesterday. The idea would be that participles from a sparsely populated area would result in more upward strengthening than participles that we find in the center, which is more densely populated. In this graph, sparse areas are shown in yellow. The darker areas of orange and red are more densely populated.

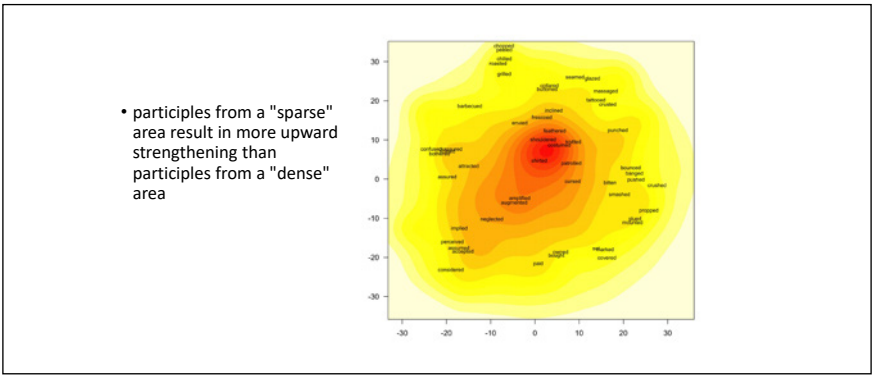


FIGURE 8

Constructions can also be forgotten. Generalizations that are not strengthened undergo decay if no upward strengthening occurs. My modeling thus included a decay function. When a construction is not regularly updated, then it is eventually forgotten.

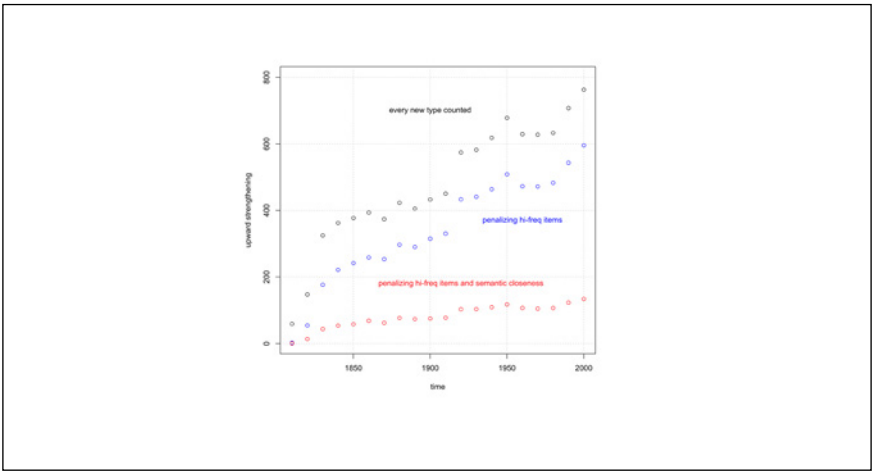


FIGURE 9

On the basis of these measurements, I created the graph you see on this slide. The graph is based on the following information. First, I determined for each period of time how many new instances are attested. I also determined how saliently they illustrate the category and how similar they are to already existing types. Based on these factors, I calculated a measure of upward strengthening. If every new type is counted, without adjusting for frequency or salience, then the result is a strongly increasing curve that would suggest a lot of upward strengthening. If high-frequency forms are penalized, so that they send less

activation to the higher node, the curve is already a little bit flatter. When high-frequency items and semantic closeness are penalized, the resulting curve is just about flat. That would imply that the highest node in the network does not receive any amount of upward strengthening.

This would explain why I do not see this as a case of grammaticalization despite increases in type and token frequencies. If we do not attribute upward strengthening to types that are not clear members of the participle category, if we do not attribute upward strengthening to types that are highly frequent like *road-tested*, and if we do not attribute up strengthening to types that are semantically similar to already existing types, then we can motivate why speakers do not strengthen their overarching generalization over all noun-participle compounds.

Asked the other way around, what new types would lead to upward strengthening, and thus to grammaticalization?

According to what we know about grammaticalization and about known cases of grammaticalization, upward strengthening would occur when speakers continually produce new types which clearly instantiate the construction, and which are semantically dissimilar to already existing new types. That is, if speakers always try to stretch the limits of what a construction can do, and if they produce instances that are a little different from what is already known, that would result in upward strengthening and in grammaticalization. Once a construction is grammaticalized to a high degree, then upward strengthening should naturally cease to increase.

I am summing up here. I have discussed grammaticalization as a specific type of change in the network of constructions. I have invited you to think about that kind of change in terms of gradient strengthening of nodes and connections in that network. How do nodes in the upper layers of the network come into being and how are they subsequently strengthened? That is, I think, an important idea to think about. The upward strengthening hypothesis is a part of a more general enterprise that tries to understand grammaticalization in constructional terms. Like other issues that I have been discussing, as for example higher-order schemas, I think that this is an area that requires further research. With that, I would like to come to an end. Thank you for your attention once more.

LECTURE 10

Constructional Change and Distributional Semantics

This brings us to the tenth and last lecture of these Ten Lectures on Diachronic Construction Grammar. I wish I could say that this lecture ties it all together and explains all the remaining questions. That is not really what I am going to do here. What I will try to do is to tie up some loose ends with regard to constructional change and the distributional methods that I have been talking about in earlier lectures. I feel that when you hear about distributional semantics and semantic vector spaces for the first few times, it can be very demanding. In this lecture, I want to take things a little more slowly and present some of the issues that I have already talked about in a little more detail. I also want to give you two more examples of analyses that you can do with this method.

Without further ado, let me talk about the motion charts that you have seen in earlier lectures. I was introduced to motion charts by Hans Rosling. A few years ago, I watched a video of a talk that he gave, which had the somewhat sensationalist title *Hans Rosling Shows the Best Statistics that You've Ever Seen*.

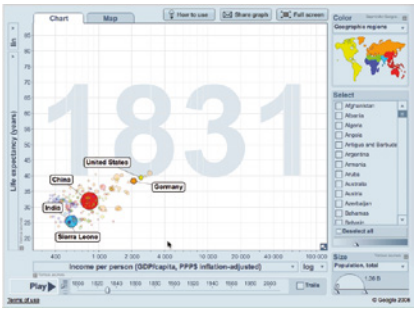


FIGURE 1

All original audio-recordings and other supplementary material, such as any hand-outs and powerpoint presentations for the lecture series, have been made available online and are referenced via unique DOI numbers on the website www.figshare.com. They may be accessed via this QR code and the following dynamic link: https://doi.org/10.6084/m9.figshare.13691296.

The graph in the video showed a two-dimensional scatter plot. The points on the graphs are countries, so you see China, and United States, and Germany, and India. Bubble size represents population size. There are two axes. The graph shows income on the *x*-axis. The further to the right, the richer is a country. The *y*-axis shows life expectancy. The higher up a country, the longer people live. The two variables are positively correlated, that is, more income means higher life expectancy. What is fascinating is that you can see how these countries develop over time. In the 1860s, life expectancy in places like India and Sierra Leone is really low. As we move into the 20th century, especially the second half of the twentieth century, life expectancy rises all across the world. Life expectancy in India today, in the early 2000s, is about as high as it was in Germany in 1955. Hans Rosling used these statistics as an argument to correct the misperception that there is a clear dichotomy between Europe and so-called developing countries. That is clearly not what the world is like.

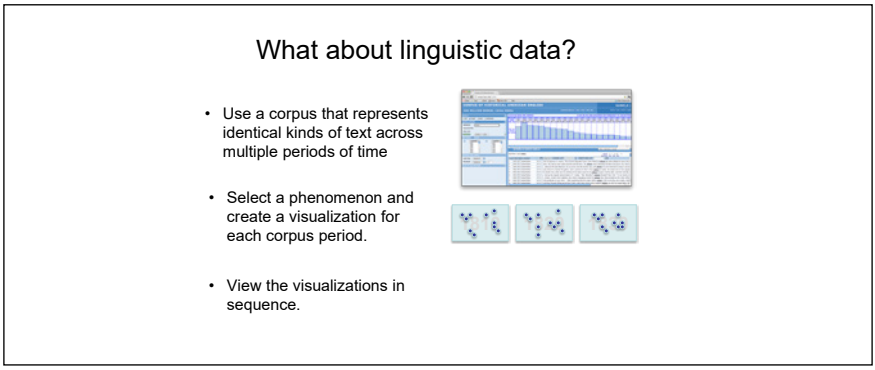


FIGURE 2

Watching this fascinated me. I began to think about ways of using this technique with language data, The ingredients were available. Diachronic corpora are readily available, and we can represent that data visually. Our bubbles wouldn't be countries. They would be constructions, or maybe speakers or dialects or other kinds of linguistic entities.

To further facilitate matters, the software that Hans Rosling uses for his presentations has been made available as a package for R. It really just takes a few clicks to make your own charts. If you're interested in that, on my homepage, you can find a folder with sample files and instruction videos that tell you exactly how to do this.

One of the first analyses that I tried out is the graph that you see on this slide. The graph represents the development of negative contraction in English. English verbs can be negated in two different ways, with *do not* or with *don't.*

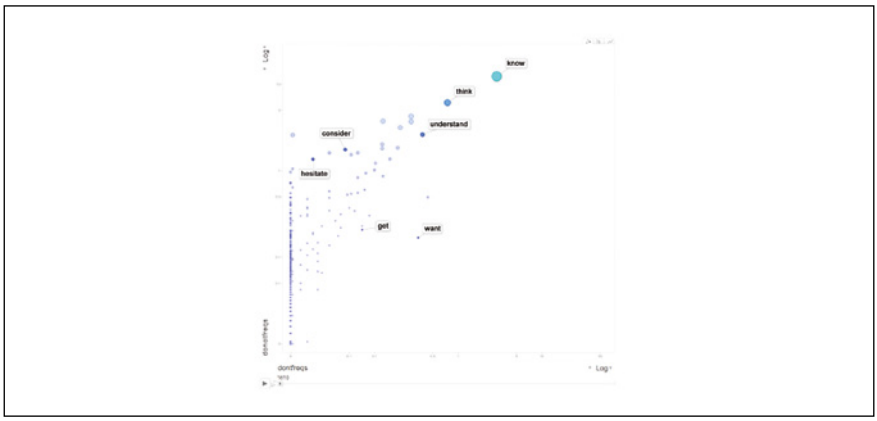


FIGURE 3

The two variants are represented on the two axes. On the *y*-axis, you see the frequencies of verbs as they are used with *do not*, with the full uncontracted form. On the *x* axis, you see how these words are used with the contracted form. In the still graph, you see data from the early 19th century. Many verbs are exclusively used with *do not*. But in fact, the more frequent a verb is, the more likely it is that both variants are used. In the early 19th century, the verb *know* occurs with *do not* and with *don't*. Frequently negated verbs include *I don't know*, *I don't think*, *I don't understand*. On the upper left side of the graph, we have verbs that prefer the non-contracted form, i.e. *do not consider* and *do not hesitate*. On the right side of the graph, further towards the *x*-axis, we see verbs that are more informal and therefore prefer the contracted variant. Diachronically, there is a trend towards the contracted form, as English writing over time has become more and more informal.

As we move through time, the verbs drift as a whole group more and more towards the contracted variant. Towards the end of the 20th century, the whole field drops a little bit, and the verbs that take exclusively the uncontracted variant become fewer and fewer. We can consider this a stylistic development. It has become more and more acceptable in English writing to use the contracted form.

While this is interesting, it is not the kind of grammatical change that I have been talking about in the previous lectures. It is not on a par with constructionalization or constructional change. Nonetheless, I thought I would start with this kind of example to show you how I came into contact with that kind of method.

As you know, one of the central ideas that shaped my thinking about language the most in these past years is that constructional meaning is reflected

in associations between syntactic patterns and lexical elements. The fact that *give* is the most typical verb for the ditransitive construction, and the *way*-construction is attracted to verb such as *elbow* or *force* or *squeeze*, this to me really underscores the point that syntactic patterns have meaning and are not meaningless templates. It is this idea that I will also elaborate on in the two case studies that I have brought for this final lecture.

We will thus come back to constructions and their collocates and what we can learn from diachronic shifts in those collocates. I have presented several cases of this kind in my previous lectures, discussing how these changes relate to grammaticalization. We can see progressive semantic change towards more abstract, grammatical meanings in shifting collocational profiles. In this lecture, I would like to address other questions that also lend themselves to a treatment with this kind of method.

The first case study is a pattern that is anachronistic and marginal in the English language. It is a construction that many of my students and in fact many second language learners of English find strange when they see it for the first time. What I am talking about is the *many a NOUN* construction.

The *many a NOUN* construction is exemplified by utterances such as *Many a day will pass until this construction is properly understood* or *I've thought that many a time myself*. In these two examples, *many a day* and *many a time* are expressions that relate to time. Both *day* and *time* are words that occur frequently in this frame. We also find *many a month*, *many a year*, *many a century* and so on and so forth. The construction is also found with words that denote human beings, as in *many a father*. For an anachronistic construction, it would be typical behavior to retreat into a narrow semantic niche, so that speakers can only use certain types of words in that particular syntactic frame. Not so with *many a NOUN*. In this kind of construction, we find nouns that do not easily fit into any kind of semantic category. Let me give you one example that is taken from a travel report: "*During my time in Australia I enjoyed many a sausage roll for brekkie*". *Brekkie*, in case you do not know, is an Australian term for breakfast, and *a sausage roll* is a pastry that has sausage baked into it.

We find all kinds of nouns in this construction. This I found puzzling, because with *many a NOUN*, we can observe in diachronic corpus data that its usage frequencies steadily fall until the construction is barely existent in the 2000s. It is very marginal in present-day usage.

This example contrasts with the V-*ment* construction that I have discussed in an earlier lecture That construction is instantiated by many types, but it is no longer productive. Here, we observe the opposite. The construction is infrequent, but still speakers create new types. I was intrigued by this contrast. Specifically, I wanted to investigate changes in the semantic spectrum of *many*

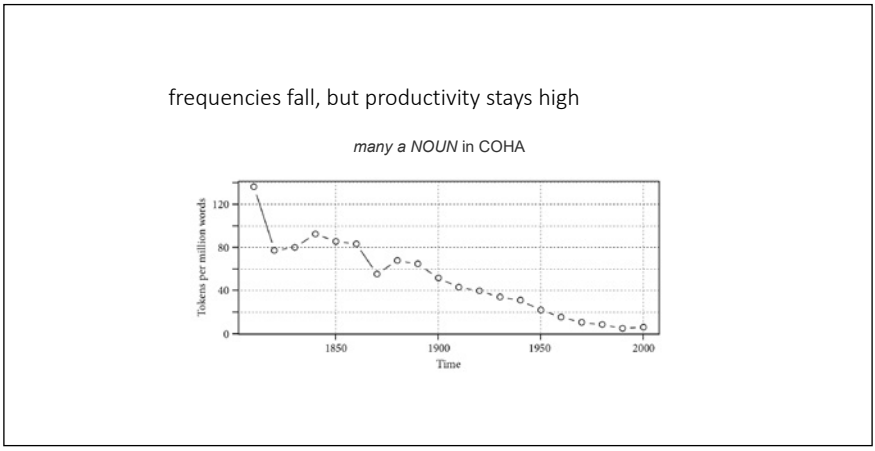


FIGURE 4

a NOUN. Can these changes explain why speakers can still create new types like *many a sausage roll*?

In order to investigate these issues, I retrieved data from the Corpus of Historical American English (COHA). I searched for sequences of words that started with the quantifier *many*, then continued with the indefinite determiner *a* or *an*, and continued with a noun. I allowed for intervening adjective, so as to allow examples such as *many a frustrated voter*. That search procedure yielded about 3000 different types that are spread out across 15,000 tokens. The tabular overview on this slide shows the token frequencies for each decade and for each type. For this analysis, I chose to focus only on a subset of the data, namely the 230 most frequent noun types. The 230 most frequent types actually account for more than sixty percent of the data. As in many other constructions, the most frequent types account for a large part of the data, and there is a long tail of infrequent forms at the other end of the spectrum.

For those 230 most frequent types, I constructed a semantic vector space, following the analytical steps I outlined yesterday: You select the vocabulary items, you retrieve the context items, you determine the co-occurrence frequencies, and you compute a collocation measure such as Pointwise Mutual Information. Then you analyze the vocabulary items in terms of their relative similarities, and you try to visualize that. The idea would be that words that occur with the same collocates are judged to be similar, and that these similar words then would be represented in relatively close proximity on a semantic map. Just to take a quick example, given a word such as *church*, what are the words that co-occur with *church* in a window of 4 words to the left and 4 words to the right?

BNC collocation frequencies (four-left, four-right window)

terms

collocates	church	heart	eye	sigh
abbey	30	0	1	0
above	13	8	26	0
activities	21	2	0	0
always	28	28	23	5
half	8	0	10	0
long	28	11	7	48
family	47	10	4	1
gave	9	13	12	109
christ	141	8	1	0
walk	15	2	1	0

FIGURE 5

Here are some collocation frequencies from the British National Corpus that illustrate the principle. *Church* occurs next to words such as *abbey*, *Christ*, and *family*. This kind of frequency profile would be very different from other words such as *heart* or *eye* or *sigh*.

The raw data that is necessary for a vector space is the numerical frequency data that you see in these columns. This data is obtained if you collect concordances for each of the vocabulary items, and determine the frequencies of lexical elements that occur in a chosen context window. Each column is what you could call a vector of frequencies. Each word has its own collocational profile. Some words are quite different from another, others are more similar. In this table, the vocabulary items *heart* and *eye* are both body parts, so we would expect that there is at least some similarity between them. This indeed can be visualized in a kind of map such as the one that you've seen yesterday.

from collocate frequencies to a map

	church	heart	eye	sigh
abbey	30	0	1	0
above	13	8	26	0
activities	21	2	0	0
always	28	28	23	5
half	8	0	10	0
long	28	11	7	48
family	47	10	4	1
gave	9	13	12	109
christ	141	8	1	0
walk	15	2	1	0

```
         church    heart      eye
heart 144.50606
eye   153.54153  23.60085
sigh  186.34645 106.44717 110.46266
```

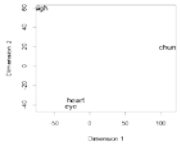


FIGURE 6

On this slide, we see the frequencies, we see distance matrix that we derive from those frequencies, and we see a visualization of that data in a semantic map that confirms that *heart* and *eye* are indeed very similar to one another, but different from *church* and from *sigh* respectively. With this reminder in place, I want to get to the actual data to show you what happened with *many a NOUN*.

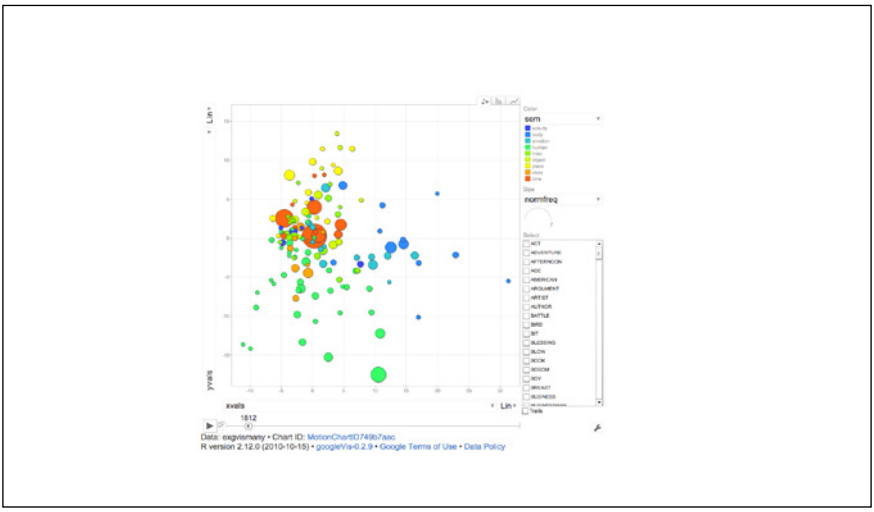


FIGURE 7

This graph represents a semantic vector space of the 230 most frequent nouns in the *many a NOUN* construction. What you see is a two-dimensional surface with a cloud of dots that are spread out over the surface. I have colored the dots according to several semantic categories. The relative positions of the bubbles reflect similarity in collocational behavior, and the sizes of the bubbles are based on the text frequencies of the nouns in the *many a NOUN* construction in the Corpus of Historical American English. Over time, some of these elements can become more or less frequent, or they might even disappear. The bubbles can grow or shrink, and more bubbles can show up. Before I show you how everything changes, let me give you a more guided tour of the semantic landscape that is represented in this graph.

In the center of the graph, you see a cluster of bubbles that I colored in red. They all belong to the category of time nouns. The biggest bubbles are *time*, *year*, *hour*, *night*, and *morning*. There is structure in the semantic space, so *summer* and *winter* are really close to one another, and there will be *spring* later on. These time words are positioned at the center of this semantic landscape. The *x* and *y* coordinates of this graph come together exactly where the noun *time* is in this graph.

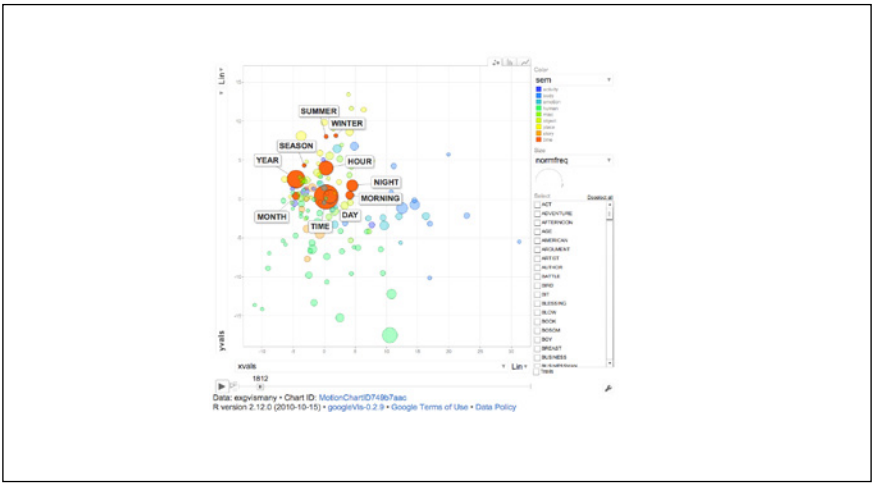


FIGURE 8

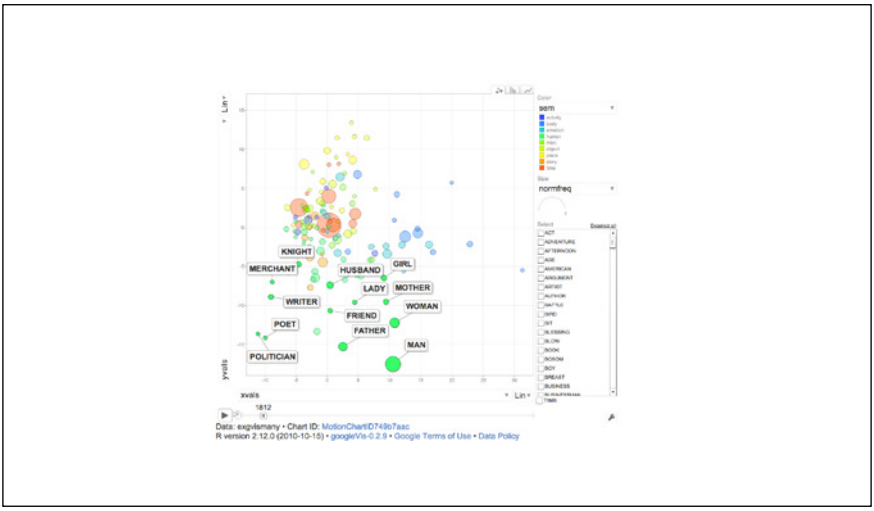


FIGURE 9

In the lower half of the graph, there is a large contiguous space that is filled with nouns that denote *human beings*. The spectrum goes from very general nouns, such as *man*, *woman*, *mother* and *girl* on the right-hand side, to family relations in the middle, like *husband*, *father* and *friend*, to very specific occupations and professions on the left-hand side, like *writer*, *politician* and *poet*, you also see *merchant and knight*. So we have time nouns in the center, and in the lower half of the graph, there are nouns that refer to human beings in different roles.

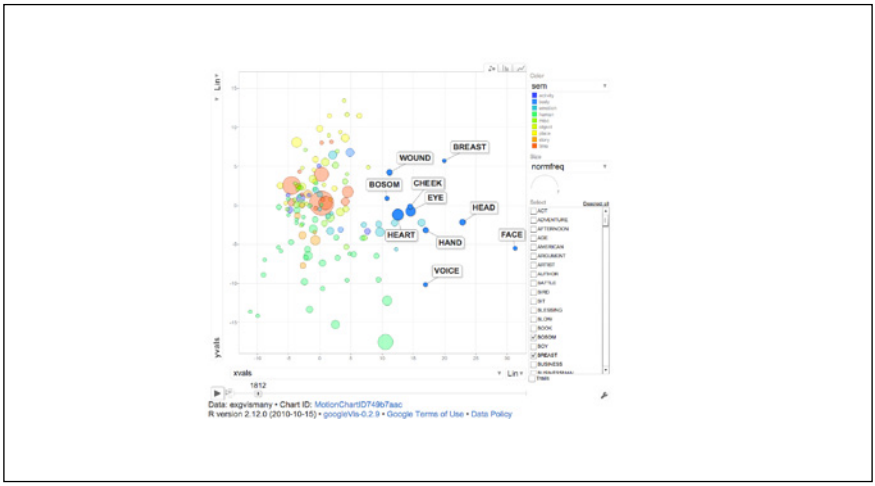


FIGURE 10

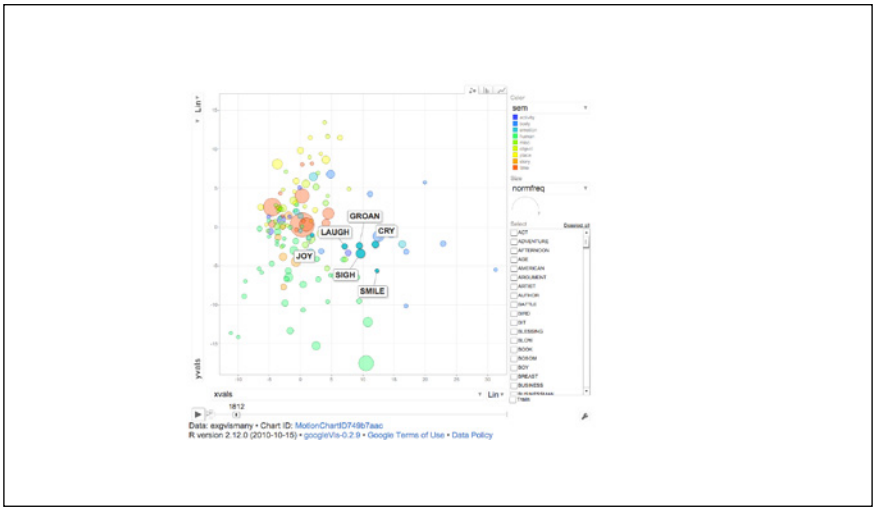


FIGURE 11

Moving on to the right-hand side, the graph shows body part nouns like *head*, *hand*, *heart*, *eye* and *face*. I decided to include *voice* in this group, even though it is not technically a body part. In the *many a NOUN* construction, these typically stand metonymically for the entire person. For example, the expression "*Many an eye shall weep*" means that many people will cry.

In between body parts and human beings, we have emotion-related nouns, exemplified by *sign*, *smile*, *cry* or *joy*. They form a small contiguous cluster in this area here. There are further categories that I annotated, but I do not want

to go through all of them. You see that the separation between them is not perfect, which is something that we would expect.

Let us examine how this landscape changed over time. I already told you that the *many a NOUN* construction becomes less frequent as we go along. However, that could mean different things with regard to this picture. It could mean that all of the bubbles become smaller, or it could mean that perhaps some of the bubbles disappear. It would also be possible that one area of the graph empties and another area stays relatively well filled. It turns out that what happens is a combination of everything. As we are going from the first half of the 19th century into the second half, more types are coming into the construction. But then, as we enter the second half of the 20th century, the space thins out fairly radically, so nearly all the bubbles become a lot smaller, and many of them disappear. In other words, the semantic space of the *many a NOUN* construction becomes more sparsely populated.

I would like to show you each category individually, starting with the development of *time* nouns. We started with a fair number of them. During the second half of the 20th century, all of them become a lot less frequent, but most of them are still in use. This is an aspect of the construction that remains intact. However, the construction as such is used less often with time nouns than it used to be the case.

This development contrasts with that of body parts, which in this construction stand metonymically for entire human beings. In the 19th century, we have lots of them, including *foot*, *breast*, *bosom*, *cheek*, *eye*, *hand*, *head*, *face* and so on and so forth. They first become less frequent, and then as we move into the 20th century, they gradually fade away one by one, until only the word *soul* remains. The body part category radically thins out and then disappears.

What happened to the human beings? They form a large set in the 19th century. Their development is similar to that of the time nouns. They become overall less frequent as we go along, but many of them stay in usage even throughout the second half of the 20th century. If you take highly frequent nouns such as *man* and *woman*, it is probably not so surprising that we still have them in the construction, but there are also anachronistic lexical elements such as *fellow* and *maiden*. That testifies to the fact that this construction continues to have the function of referring to human beings.

What can we learn from these observations? Can we come back to the question of *the sausage roll* and answer why speakers can still use the construction in this way? The answer I would like to propose is that the *many a NOUN* construction does not recede into a narrow semantic niche like many idiomatic constructions do. We have very frequent and very general nouns that are part of the most typical nouns that are used in this construction, i.e. *time* and

man. These words are highly diffuse in their collocational behavior and in the syntactic contexts that you find them. That makes it very difficult for speakers to experience the *many a NOUN* construction as semantically restricted. If we find a construction that is used with general nouns such as *time* and *man*, which have very different meanings, then we are likely to conclude that this construction is actually semantically unrestricted, and that it can be used with any kind of lexical element.

I would like to move on to the second case study that Florent Perek and I have conducted together. In this study, the issues of distributional semantics and grammaticalization are addressed together. Distributional change allows us to examine how forms grammaticalize. The construction that I will be looking at here is the English verb *get* in one of its grammatical functions.

English *get* is something of a linguistic Swiss Army knife. It has many functions. *Get* has the lexical meaning of 'receiving' as in *Look what I got for my birthday!* Then there is the *get*-passive that I have been mentioning a couple of times in these lectures, as in *Nobody move, nobody get hurt.* There is a causative construction with *get*, *Can I get you to deliver a message?* There is what we could call inchoative *get*. In expressions such as *It gets worse and worse*, *get* functions as a copula, it allows us to express a predication. Lastly, there are many idiomatic uses of *get*. *I get up at seven* means that I wake up at seven o'clock. *I do not get it* means that I do not understand it.

I want to focus on another function of *get*, namely the use of *get* that expresses permission. This slide shows three examples: *In the movies the prisoners always get to make one phone call*. This means that in movies, the prisoners are allowed to make one phone call, usually to their lawyer. *This is a big day for the guards. They get to remind us who's boss*. This means that the guards have the possibility to remind us who's boss. *I want to be a Marine. They get to wear swords*. This means that Marines have the permission and the privilege of wearing swords. This permissive meaning takes *get* into the grammatical domain of modality. There are other modal verbs that express permission, like *may* or *can*. *Get* has entered that paradigm. The examples that I just read to you differ subtly in the meanings that they express. They can of course express a permitted action. In *They get to remind us who's boss*, it is however not permission in the strict sense. It is not really that someone allowed the guards to remind someone else who's the boss, but rather, we understand that they have the opportunity to remind us. Similarly, in *They get to wear swords*, an authority has given them the privilege to wear swords. There is permission in that sense, but the Marines do not really have a choice. They are given a sword as a sign of their status.

In the next minutes, I would like to address the following questions: How did permissive *get* emerge? How did we end up with this multi-functional verb adopting yet another function and where does this come from? I also want to examine what has been said about permissive *get* in earlier work. Lastly, I want to use corpus data and distributional semantics to figure out how this construction developed and what we can assume as a source. When these polysemous constructions develop into a new grammatical construction, the problem is that there are several source meanings to choose from. Which one is actually the source construction for permissive *get*? As we will see, there are different proposals with regard to that.

To give you an overview of the next minutes, I will first focus on two conflicting accounts of permissive *get*. The grammaticalization of permissive *get* has been described in terms of two possible grammaticalization pathways. One of these is called the causative-to-permissive pathway, and the other one is the acquisitive-to-permissive pathway. The first pathway is proposed by Gronemeyer (1999). The second is put forward by Van der Auwera and colleagues (2009) in a typological, cross-linguistic study of modals that derive from verbs of acquisition. I will criticize both of these proposals and suggest a third one.

On the basis of data from the COHA and the use of distributional evidence we will look at developments in the semantic spaces of inchoative *get* and permissive *get*. Then I will end with some conclusions, and that will take us to the end of this lecture series.

As I said, there is currently no consensus on how permissive *get* emerged.

One possible scenario is the causative-to-permissive pathway that is proposed by Gronemeyer (1999: 1). She works out a complete semantic map of the history of *get* and states the following:

> Using diachronic data, I show that possession leads to movement as well as to stative uses (possession and obligation), movement develops into the causative and inchoative, from which the passive develops, and the infinitival causative gives rise to permission and to ingressive aspect.

What Gronemeyer describes is a developmental pathway that starts with possession. That would be uses such as *I have got a new book*. That meaning gives rise to obligation and movement. Obligation, that is *I have got to make a call*, and movement is expressed by examples such as *who never gets home*. Movement further splits up into causative and inchoative meaning. Inchoative get is illustrated by *You've got to get mad*, in which *get* is a copula that is followed by a predication. Movement, according to Gronemeyer, also gives rise to causative meaning, as in *John got the students to work on the problem*. This

causative meaning, Gronemeyer claims, gives rise to permissive meaning, as in *They get to use Linda's car*.

She presents a syntactic argument for this claim. The crucial context, according to Gronemeyer (1999), is found in causative examples such as *I got him to be a chaplain*, which is an authentic example. *I got him to be a chaplain* expresses causation, but also implies the fact that he was permitted to be a chaplain. I got him to experience a kind of privilege. I got him to be a chaplain. The caused event and the permitted event refer to the same idea. Gronemeyer suggests that speakers may have treated *get* in *I got him to be a chaplain* as an anticausative verb. Anticausative verbs participate in the causative alternation. For example, verbs like *melt* can be used transitively, *The sun melted the ice*, and intransitively, *The ice melted*. Gronemeyer argues that *get* acquired the syntactic properties of that class. *I got him to be a chaplain* represents the transitive use that goes along with causative *get*. If *get* is used not transitively but intransitively, we have sentences such as *He got to be a chaplain*, which do not have a causer argument. That sentence only has the implied permissive meaning, which can then conventionalize. Gronemeyer's analysis is a classic syntactic account that takes a phenomenon that is well-documented, the causative alternation, and then uses that phenomenon as an explanation for something else. That sounds plausible enough, but not everybody is convinced by the causative-to-permissive pathway of *get*.

Van der Auwera and colleagues (2009: 284) explicitly contradict this proposal: "Gronemeyer (1999: 30–32, 35) actually claims that what she calls 'permissive' *get* derives from 'causative' *get*", as illustrated in *John got me to clean his car*. They continue, "This is not very plausible though".

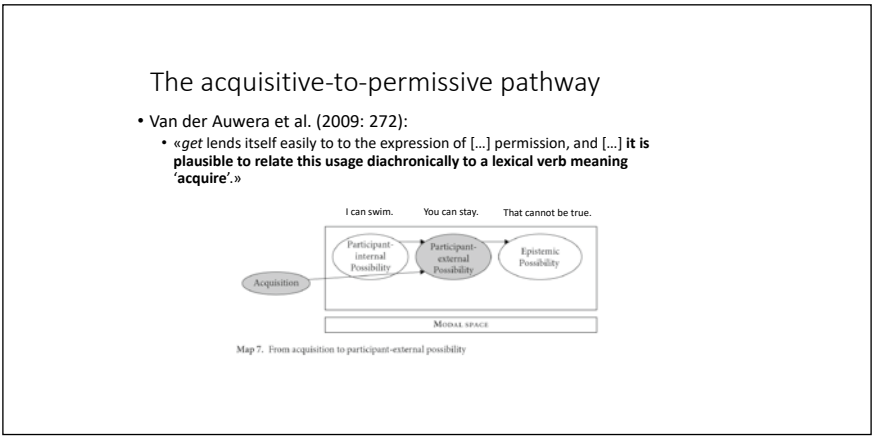


FIGURE 12

In their view, permissive *get* is one of many examples in the world's languages that instantiates the pathway from acquisitive meaning to permissive meaning. They state that "*get* lends itself easily to the expression of permission, and it is plausible to relate this usage diachronically to a lexical verb meaning 'acquire'." (Van der Auwera *et al.* 2009: 272)

The graph that you see on the slide is a typological semantic map that represents different modal meanings and their diachronic relations. On the left side, we have ability or participant-internal possibility, which is expressed by the English auxiliary *can* in *I can swim*. Participant-internal possibility can give rise to permission or participant-external possibility. Uses of English *can* such as *You can stay* express permission. If we have a linguistic element that can express participant-internal possibility, that meaning might be extended to participate-external possibility. Possible extensions are shown as arrows in the semantic map. One such extension connects ability and permission. Permission, in turn, can give rise to epistemic possibility, as in *That cannot be true*, which is signified by the arrow from participant-external possibility to epistemic possibility. The logic of a semantic map is that forms can express meanings in a contiguous space. This means that there is no way for a language to have a verb that conveys the meanings of ability and epistemic possibility, but not the meaning of permission. Semantic maps thus make predictions.

The box in the graph represents modal meanings. You see that outside that box of modal space, there is a different meaning, namely, acquisition, a lexical meaning that is conveyed by the lexical use of *get*. There is an arrow that goes from acquisition directly to permission. That arrow represents the observation that verbs of acquisition tend to give rise to markers of permission across many languages. Those elements can then move on to develop further meanings, such as epistemic possibility.

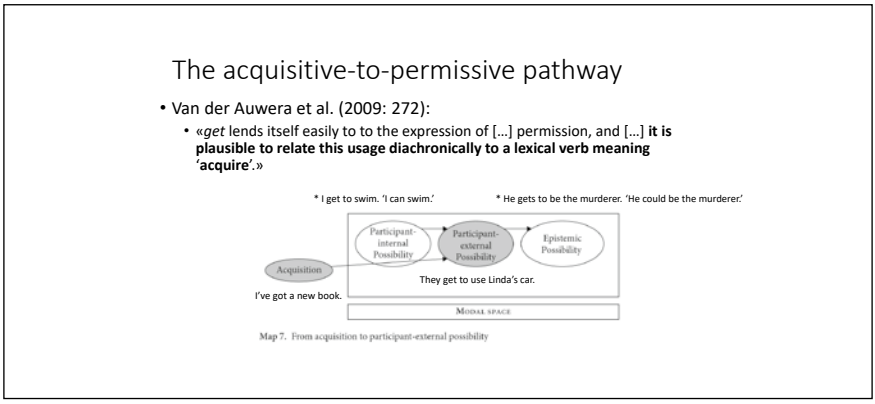


FIGURE 13

Acquisitive meaning typically does not give rise to participant-internal possibility. It is a direct source for participant-internal possibility. English actually illustrates that. Lexical *get* would be illustrated by *I have got a new book*, and permissive *get* is illustrated by *They get to use Linda's car*. Crucially, English *get* is not a marker of ability. In English, we cannot say **I get to swim* with the meaning of *I can swim* or *I am able to swim*.

Likewise, English *get* has not acquired the meaning of epistemic possibility, so it cannot be used to express logical possibilities. We cannot say **He gets to be the murderer* and convey the idea that *He could be the murderer*. English *get* occupies just the two areas that are shown in grey, acquisition and participant-external possibility. Ultimately, Van der Auwera and colleagues (2009) analyze a greater sample of languages and find that the picture is somewhat more complex. In some languages, acquisition actually does give rise to the meaning of ability. They also make a discovery that goes against strict unidirectionality in grammaticalization. They find cases where permissive meaning gives rise to the meaning of ability, so that meaning goes back and forth inside the general area of modality. Yet the broad tendency is that acquisitive meaning develops into permissive meaning. This is very suggestive of a scenario in which English *get* acquired permissive meaning because it conveyed a sense of acquiring something. That leaves us with two conflicting accounts, one based on a syntactic argument, the other based on a typological argument. Both offer valid points. But as I said earlier, I have doubts about both of them.

I would like to work out an alternative hypothesis that brings another semantic facet of *get* into focus, namely, its inchoative meaning. This alternative hypothesis is the inchoative-to-permissive pathway. Its starting point is the meaning of *get* that denotes a change of state, an onset of a new activity or a new state of affairs. Let me give you some examples, such as *It gets worse and worse*, or *I got into the habit*, or *You're getting to be a big girl now*. All of those mean that some state of affairs is about to change or is currently changing. These examples are morphosyntactically quite diverse. We have complements of *get* that are adjectival, as in *It gets worse and worse*. We have prepositional phrases, as in *I got into the habit*. There is a verbal complementation structure in *You're getting to be a big girl now*. The crucial context in which permissive *get* can conventionalize as a meaning would be an inchoative change of state that is simultaneously a privilege or fortunate turn of events. Examples from the data that illustrate this are examples such as *I guess we won't get to see Colonel Morrison after all* or *Some day she'd get to be an editor herself*, or *Oh thank you and you'll get to meet our new minister then.* Examples such as these may have served as bridging contexts between inchoative and permissive meanings of

get. The verbalized message is that there is a change of state, but there is an implicature that this change of state was granted by an authority, that is, some person or entity that gave permission for this change of state. At this point, one argument in support of our hypothesis is that the other two accounts have to stipulate their semantic shifts. They do not offer bridging contexts that would show how a hearer could actually understand the verb *get* in a way that would allow the semantic shift to happen naturally, through the conventionalization of an implicature.

Besides this qualitative argument, I would like to use the rest of my time here to develop a second argument that is based on quantitative evidence, which brings me back to the COHA. The Corpus of Historical American English is a large corpus covering the past two centuries. Here, I only use data from the 1860s onward, because the corpus has a more even representation of genres after the 1860s. From the COHA, I extracted uses of *get* followed by an infinitive, which resulted in some 30,000 examples, which then were manually annotated for five different semantic categories, namely, permission, obligation, causation, possession, and inchoative meaning. All examples were annotated for the lexical verb in the infinitive, which for the sentences on this slide would be *make*, *leave*, *confess* and *be*.

On this slide you see the frequency developments of the five semantic types. What can be seen is that permissive *get* is clearly on the rise. It starts slowly and then increases in frequency, despite the fact that it emerged more than 100 years before the COHA data, so permissive *get* is actually fairly old. We find the earliest examples in the English of Shakespeare's times.

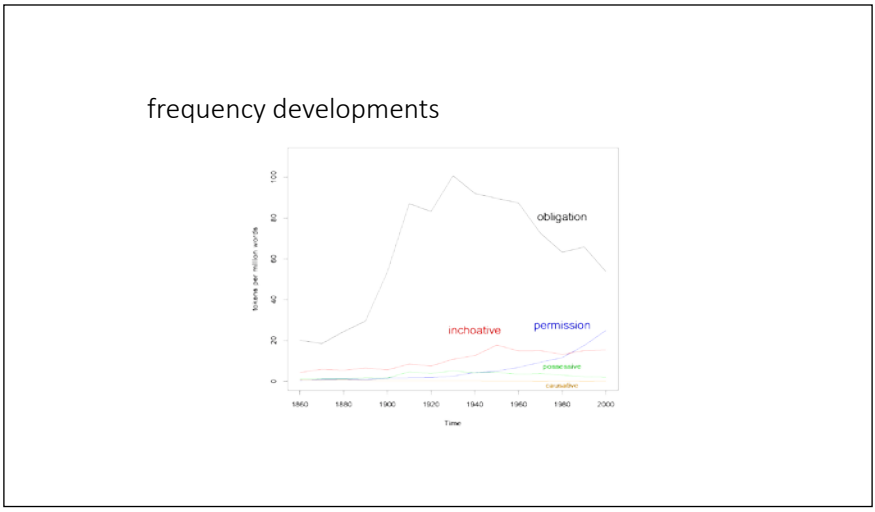


FIGURE 14

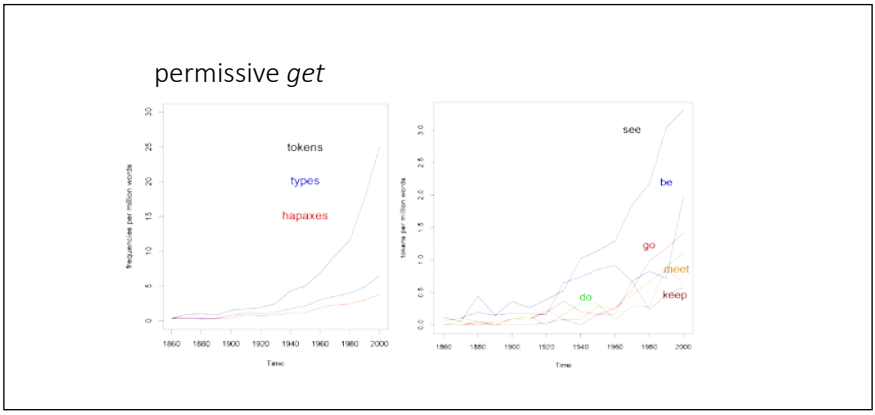


FIGURE 15

These graph on the left shows diachronic increases across token frequencies, type frequencies and hapax legomena of permissive *get*. Among the most frequent verbs we find *see*, *be* and *go*, which harmonize with the meaning of a privilege. If *you get to see the Eiffel tower*, that means that you are lucky enough to enjoy that privilege, and not that someone allowed you to behold the Eiffel tower. What we see in these collocational preferences can be interpreted as a persistence effect of the inchoative meaning that I would like to argue is the source for permissive meaning.

To strengthen this point further, Florent Perek and I turned to distributional semantic methods. Our hypothesis is that permissive *get* emerges through secondary grammaticalization from inchoative *get*. We derived two predictions from that hypothesis. First, there should be what Hopper (1991) has called lexical persistence. Grammaticalized constructions retain traces of their lexical history. Second, we predict that permissive *get* undergoes host-class expansion. Grammaticalized constructions gradually expand their range of lexical fillers (Himmelmann 2004), and that should be observable in the data that we have.

Based on this, we formulated two empirical questions. First, do inchoative and permissive *get* collocate with similar sets of verbs? You remember that I asked a similar question in my analysis of concessive parentheticals. The same analytical tool is applied here. Second, to what extent does permissive *get* emancipate itself from inchoative *get*? As in the previous studies of distributional semantics that I have talked about, we created visual representations of the semantic areas that are occupied by the two constructions. The similarities in collocations reflect similarities between word meanings.

This slide recapitulates our methodological steps. They correspond to the work steps that I have described in more detail in earlier lectures.

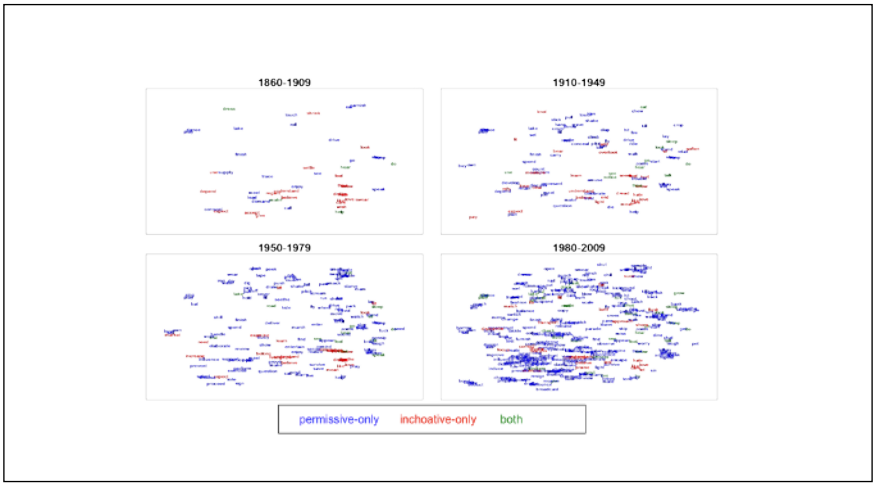


FIGURE 16

This slide offers a first summary of our results. You see one plot for each of the four historical periods that we investigated. The plots show the verbs that are attested with permissive *get*, with inchoative *get* and with both of the constructions. Permissive-only verbs are shown in blue, inchoative-only verbs are shown in red, and overlapping verbs are shown in green.

What we expected to see was a gradual decline of green and increasing diversification into red and blue. That kind of development is not apparent in the data. Instead, we see that the two constructions occupy overlapping areas of semantic space. We further see that permissive *get* expands semantically over time both inward and outward. Inward, it fills areas of semantic space that were previously not filled. Outward, it expands into areas that were previously not covered at all.

The fact that the two constructions occupy the same semantic areas can be interpreted as a lexical persistence effect. The fact that permissive *get* expands into different areas over time can be seen as host-class expansion. What I just described to you is of course merely a qualitative interpretation of these data. Can these effects actually be measured? Can we quantify how much lexical persistence or how much host-class expansion there is in both permissive *get* and inchoative *get*?

In order to answer these questions, we partitioned the semantic space into areas by clustering the different verb types that we found in inchoative and permissive *get*. We arrived at a solution of 12 verb clusters that we used to divide the semantic space into 12 semantic areas. Let me illustrate the verb clusters that we found. Cluster 1 consists of verbs that refer to speech and sound, such

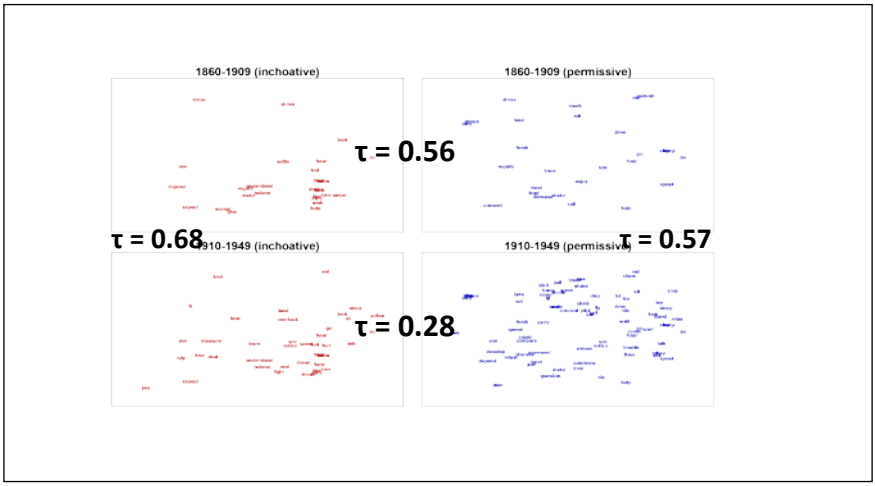


FIGURE 17

as *say*, *tell*, *ask*, *hear*, *speak*, *answer*, and *laugh*. Cluster 3 includes verbs of emotion and cognition. Examples are *love*, *enjoy*, *hate*, and *hurt*. Cluster 5 is about ingestion, as in *eat*, *drink*, and *swallow*. Cluster 9 denotes manipulation and force, as in *turn*, *open*, *shake*, or *pull*.

What we can investigate on the basis of these clusters is whether the same areas are populated in the same way by the two constructions. We determine how densely each semantic cluster was populated in each construction and in each period. We ran correlation statistics to obtain similarity measures between the distribution of two constructions at different points in time and also, crucially, between the same construction at different points in time.

On this slide, we see the graphs of the semantic distribution again. This time, however, there are separate graphs for inchoative *get* on the left-hand side and permissive *get* on the right-hand side. We can now use the partitioning into different verb clusters in order to quantify similarities between the two constructions. Between-construction comparisons would be between left and right, within-construction comparisons would be on the vertical axis, so that you compare inchoative *get* of the first period to inchoative *get* of the second period. We can correlate the population density of the semantic space in both of these ways. If we run a correlation statistic that tells us how similar or dissimilar two distributions are, we obtain values such 0.56, which tells us that in the first period, inchoative and permissive are relatively similar. A correlation of 1 would mean that they are identical. A correlation of 0.56 means that they are different, but more similar to each other than in the second period, where we only have a correlation of 0.28. From the first to the second period, these

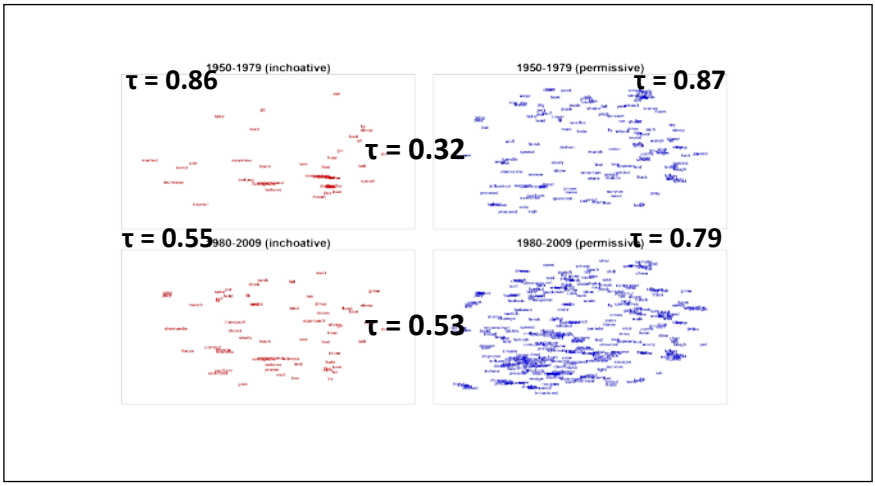


FIGURE 18

two constructions diversify. They become more different in terms of their collocational behavior.

When we compare the constructions to themselves over time, what result do we get? The correlation values are actually higher. Between the first period and the second period, inchoative *get* changes a little bit, which results in a correlation of 0.68. Permissive *get* changes a bit more, so here the correlation is 0.57.

Let's move on to the next two periods. In the third period, both inchoative and permissive *get* are similar to themselves. The correlation values between the third period and the preceding second period are 0.86 and 0.87 respectively. They continue to be very different from each other, so they are drifting further apart. A bit surprisingly, the fourth period shows that the two constructions are re-approaching one another. When we look at the comparisons within the constructions, there are still fairly high values, but the similarity is not as strong as between the constructions themselves earlier in time. We observe values of 0.55 for the inchoative and 0.79 for the permissive.

What can we conclude from this? There is a decrease in similarity between inchoative and permissive *get* from period 1 to period 3. There is overall less change in inchoative *get* than in permissive *get*, and inchoative *get* regains more types in the last period and becomes more similar to permissive *get*.

How did permissive *get* emerge? With these data, we can make a plausible case that inchoative *get* is the source for the grammaticalization of permissive *get*. There are bridging contexts that provide a motivation for that source. The

distributional evidence portrays a quantifiable trajectory of grammaticalization in terms of persistence and host-class expansion. Methodologically, the combination of clustering and correlating the distributions is a new method to assess the semantic spread of constructions, both between constructions and in the same construction over time.

I am coming to the end of this lecture series. I have talked about a number of theoretical and methodological issues, starting with notions such as constructionalization, constructional change, how constructional change is different from grammaticalization and from language change, what the underlying notions are, what the relevant controversies are, and what methods we can use to come to new insights with regard to constructional change. If you were hoping for final answers, I am afraid I will disappoint you. However, I hope to have given you an appetite for this kind of topic. We've covered a lot of ground, including auxiliaries and their verbal complements, complement-taking verbs and their syntactic subcategorization frames, morphological constructions like the *V-ment* construction and compounding constructions, and grammaticalization paths such as the one that led to permissive *get*. Diachronic Construction Grammar interfaces with many different areas of research. It can engage with theoretical debates and arguments, and it can bring a new perspective to these areas. One notion that I think is crucial and that I would like to end with here is the notion of links.

Construction Grammar started as a research project that placed form-meaning pairs at the center of the study of language. That was very much a necessary step at the time. Following that step, much attention focused on the characteristics of form-meaning parings. What kind of constraints do they have? What kind of structural relations do they entertain with each other? How do they change? There is an increasing need to rephrase the questions that we have about constructions and constructional change in terms of connections between them. I have talked about what I called the fat node problem, the fact that if we transcribe information directly into the nodes of constructions, we are sidestepping crucial questions and we are opening ourselves up to serious criticism from other related fields. Construction Grammar has the ability to re-conceptualize many of its core notions in terms of relations between constructions, for instance, in terms of associations between constructions and the lexical items that occur within them.

Some examples of link-based thinking are already part of constructional research. On the very first day, I mentioned Adele Goldberg's work on statistical preemption that shows how links in the construct-i-con allows speakers to

learn constructional constraints. Of course, a lot of work remains to be done. I am very excited about the future work that undoubtedly is to come in this area.

Finally, this overview of Diachronic Construction Grammar has of course vastly overemphasized my own work. I apologize for that, but I thought I would rather speak about topics that I am very familiar with. There is a lot of work in Diachronic Construction Grammar that differs in substantial ways from what I have been presenting here. I strongly encourage you to engage with that important work. With that, I would like to express my gratitude one final time. It was a great honor to spend this week with you all. Thank you so much for your hospitality. It has been a wonderful time, and I look forward to seeing you again soon.

References

Anshen, Frank, and Mark Aronoff. 1999. Using dictionaries to study the mental lexicon. *Brain and Language* 68(1–2): 16–26.

Barðdal, Jóhanna. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. Amsterdam: John Benjamins.

Bauer, Laurie. 2001. *Morphological Productivity*. Cambridge: Cambridge University Press.

Bauer, Laurie, Rochelle Lieber and Ingo Plag. 2013. *The Oxford Reference Guide to English Morphology*. Oxford: Oxford University Press.

Bloomfield, Leonard. 1933. *Language*. New York: Henry Holt.

Booij, Geert. 2013. Morphology in Construction Grammar. In Thomas Hoffmann and Graeme Trousdale (eds), *The Oxford Handbook of Construction Grammar*. New York: Oxford University Press, pp. 255–273.

Börjars, Kersti, Nigel Vincent, & George Walkden. 2015. On constructing a theory of grammatical change. *Transactions of the Philological Society*. 113(3): 363–382.

Brinton, Laurel J. 2008. *The Comment Clause in English: Syntactic Origins and Pragmatic Development*. Cambridge: Cambridge University Press.

Bybee, Joan L. 2003. Cognitive processes in grammaticalization. In Michael Tomasello (ed.), *The New Psychology of Language, Volume II.* Mahwah: Erlbaum, pp. 145–167.

Bybee, Joan L. 2003. Mechanisms of change in grammaticization: The role of frequency. In Brian D. Joseph and Richard D. Janda (eds.), *Handbook of Historical Linguistics*. Oxford: Blackwell Publishers, pp. 602–623.

Bybee, Joan L. 2010. *Language, Usage, and Cognition*. Cambridge: Cambridge University Press.

Bybee, Joan L., William Pagliuca and Revere Perkins. 1991. Back to the future. In Elizabeth Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization*, *Vol. II*. Amsterdam: John Benjamins, pp. 17–58.

Bybee, Joan L., Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect and Mood in the Languages of the World*. Chicago: University of Chicago Press.

Croft, William A. 2000. *Explaining Language Change: An Evolutionary Approach*. Longman Linguistics Library. London: Pearson Education Limited.

Croft, William A. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Croft, William A., and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Dąbrowska, Ewa. 2008. Questions with long-distance dependencies: A usage-based perspective. *Cognitive Linguistics* 19(3): 391–425.

Dąbrowska, Ewa. 2010. Naive v. expert intuitions: An empirical study of acceptability judgments. *Linguistic Review* 27: 1–23.

Dąbrowska, Ewa. 2017. *Ten Lectures on Grammar in the Mind*. Leiden: Brill.

Dalton-Puffer, Christiane. 1996. *The French Influence on Middle English Morphology: A Corpus-based Study of Derivation*. Berlin: Mouton de Gruyter.

Davies, Mark. 2010. *The Corpus of Historical American English (COHA): 400+Million Words, 1810–2009*. Available online at http://corpus.byu.edu/coha.

De Smet, Hendrik, Frauke D'hoedt, Lauren Fonteyn and Kristel Van Goethem. 2018. The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics* 29(2): 197–234.

Diessel, Holger. 2019. *The Grammar Network: How Linguistic Structure is Shaped by Language Use*. Cambridge: Cambridge University Press.

Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of let alone. *Language* 64(3): 501–538.

Fillmore, Charles J., Russell R. Lee-Goldman, and Russell Rhodes. 2012. The FrameNet Constructicon. In Hans C. Boas and Ivan A. Sag (eds), *Sign-Based Construction Grammar*. Stanford: CSLI, pp. 283–299.

Firth, John Rupert. (1957). A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis*. Blackwell, Oxford, pp. 1–32.

Fischer, Olga. 2007. *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: Oxford University Press.

Gadde, Fredrik. 1910. *On the History and Use of the Suffixes -ery (-ry), -age and -ment in English*. Lund: Berlingska Boktryckeriet.

Givón, Talmy. 1971. On the verbal origin of the Bantu verb suffixes. *Studies in African Linguistics* 2(2): 145–163.

Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Goldberg, Adele E. 2002. Surface generalization: An alternatice to alternations. *Cognitive Linguistics* 13(4): 327–356.

Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7(5): 219–224.

Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.

Goldberg, Adele E. 2019. *Explain Me This: Creativity, Competition and the Partial Productivity of Constructions.* Princeton and Oxford: Princeton University Press.

Gries, Stefan Th. 2009. *Statistics for Linguistics with R: A Practical Introduction*. Berlin, New York: Mouton de Gruyter.

Gries, Stefan Th., Beate Hampe, and Doris Schönefeld. 2005. Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16(4): 635–676.

Gries, Stefan Th., and Martin Hilpert. 2008. The identification of stages in diachronic data: Variability-based neighbor clustering. *Corpora* 3(1): 59–81.

Gries, S. Th. and Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics*, 9(1): 97–129.

Gronemeyer, Claire. 1999. On deriving complex polysemy: The grammaticalization of *get*. *English Language and Linguistics* 3(1): 1–39.

Hilpert, Martin. 2006. Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory* 2/2, 243–57.

Hilpert, Martin. 2008. *Germanic Future Constructions: A Usage-based Approach to Language Change.* Amsterdam & Philadelphia: John Benjamins.

Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.

Hilpert, Martin. 2014/2019. *Construction Grammar and Its Application to English*. Edinburgh: Edinburgh University Press.

Hilpert, Martin. 2015. From hand-carved to computer-based: Noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics* 26(1): 1–36.

Hilpert, Martin, and Florent Perek. 2015. Meaning change in a petri dish: Constructions, semantic vector spaces, and motion charts. *Linguistics Vanguard* 1(1): 339–350.

Hilpert, Martin, and David Correia Saavedra. 2016. The unidirectionality of semantic changes in grammaticalization: An experimental approach to the asymmetric priming hypothesis. *English Language and Linguistics* 22(3): 357–380.

Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticization: Opposite or orthogonal. In Walter Bisang, Nikolaus P. Himmelmann, and Björn Wiemer (eds), *What Makes Grammaticalization: A Look from Its Components and Its Fringes*. Berlin: Mouton de Gruyter, pp. 21–42.

Hopper, Paul J. 1991. On some principles of grammaticalization. In Elizabeth C. Traugott and Bernd Heine (eds.), *Approaches to Grammaticalization* (*vol. 1*). Amsterdam: John Benjamins, pp. 17–35.

Hopper, Paul J., and Elizabeth C. Traugott. 1993/2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hudson, Richard. 2008. Word grammar and construction grammar. In Graeme Trousdale and Nicholas Gisborne (eds.), *Constructional Approaches to English Grammar*. Berlin and New York: Mouton de Gruyter, pp. 257–302.

Hudson, Richard. 2007. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press.

Hudson, Richard. 2015. Review of Rolf Kreyer, The nature of rules, regularities and units in language: A network model of the language system and of language use. Mouton De Gruyter, 2014. *Journal of Linguistics* 51/3, 692–696.

Israel, Michael. 1996. The *way* constructions grow. In Adele E. Goldberg (ed.), *Conceptual Structure, Discourse and Language*. Stanford: CSLI Publications, pp. 217–230.

Jäger, Gerhard, and Anette Rosenbach. 2008. Priming and unidirectional language change. *Theoretical Linguistics* 34(2): 85–113.

Kay, Paul, and Fillmore, Charles J. 1999. "Grammatical Constructions and Linguistic Generalizations: The What's X Doing Y? Construction." *Language* 75: 1–34.

Langacker, Ronald W. 2000. A dynamic usage-based model. In Michael Barlow and Suzanne Kemmer (eds.), *Usage-based Modes of Language*. Stanford: CSLI Publications, pp. 1–63.

Langacker, Ronald W. 2003a. Constructional integration, grammaticalization, and serial verb constructions. *Language and* Linguistics 4(2): 251–278.

Langacker, Ronald W. 2003b. Construction in Cognitive Grammar. *English Linguistics* 20: 41–83.

Langacker, Ronald W. 2003c. Extreme subjectification: English tense and modals. In H. Cuyckens, T. Berg, R. Dirven and K. U. Panther (eds.), *Motivation in Language: Studies in Honor of Günter Radden*. Amsterdam & Philadelphia: Benjamins, pp. 3–26.

Langacker, Ronald W. 2005. Construction grammars: Cognitive, radical, and less so. In de Mendoza Ibáñez, F. J. Ruiz, Cervel, P., and Sandra, M. (eds.), *Cognitive Linguistics: Internal dynamics and interdisciplinary interaction*. Berlin: Mouton de Gruyter, pp. 101–159.

Langacker, Ronald W. 2013. *Essentials of Cognitive Grammar*. Oxford and New York: Oxford University Press.

Lehmann, Christian. 2015. *Thoughts on Grammaticalization* (*3nd edition*). Berlin: Language Science Press.

Lieber, Rochelle. 1983. Argument linking and compounds in English. *Linguistic Inquiry* 14(2): 251–285.

Lightfoot, David. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.

Lightfoot, David. 1999. *The Development of Language: Acquisition, Changes and Evolution*. Malden, MA: Blackwell Publishers.

Michaelis, Laura A. 2004. Type shifting in Construction Grammar: An integrated approach to aspectual coercion. *Cognitive Linguistics* 15(1): 1–67.

Millar, Neil. (2009). Modal verbs in TIME: frequency changes 1923–2006. *International Journal of Corpus Linguistics* 14(2): 191–220.

Noël, Dirk. 2007. Diachronic construction grammar and grammaticalization theory. *Functions of Language* 14: 177–202.

Perek, Florent. 2012. Alternation-based generalizations are stored in the mental grammar: Evidence from a sorting task experiment. *Cognitive Linguistics* 23(3): 601–635.

Perek, Florent and Martin Hilpert. 2014. Constructional tolerance: Are argument structure constructions equally powerful across languages? *Construction and Frames* 6(2): 266–304.

Perek, Florent. 2016. Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics* 54(1): 149–188.

Perek, Florent and Adele E. Goldberg. 2017. Linguistic generalization on the basis of function and constraints on the basis of statistical preemption. *Cognition* 168: 276–293.

Perek, Florent and Martin Hilpert. 2017. A distributional semantic approach to the periodization of change in the productivity of constructions. *International Journal of Corpus Linguistics* 22(4): 490–520.

Pijpops, Dirk and Freek Van de Velde. 2016. Constructinal contamination: How does it work and how do we measure it? *Folia Linguistica* 50(2): 543–581.

Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. Berlin: Mouton de Gruyter.

Roberts, Ian, and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization.* Cambridge: Cambridge University Press.

Salkie, Raphael, Pierre Busuttil and Johan van der Auwera. 2009. *Modality in English: Theory and Description*. Berlin and New York: Mouton de Gruyter.

Stefanowitsch, Anatol. 2006. Distinctive collexeme analysis and diachrony: A comment. *Corpus Linguistics and Lingustic Theory* 2(2): 257–262.

Stefanowitsch, Anatol. 2011. Constructional preemption by contextual mismatch: A corpus-linguistic investigation. *Cognitive Linguistics* 22(1): 107–129.

Stefanowitsch, Anatol and Stefan Th. Gries. 2003. Collostructions: Investigating the interaction of words and constructions. *International Journal of Corpus Linguistics* 8(2): 209–243.

Szmrecsanyi, Benedikt. 2006. *Morphosyntactic Persistence in Spoken English: A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics, and Discourse Analysis*. Berlin: Mouton de Gruyter.

Szmrecsanyi, Benedikt. 2013. *Grammatical Variation in British English Dialects: A Study in Corpus-Based Dialectometry*. Cambridge: Cambridge University Press.

Tomasello, Michael, Malinda Carpenter and R. Peter Hobson. 2005. The Emergence of Social Cognition in Three Young Chimpanzees. *Monographs of the Society for Research in Child Development* 70(1): 1–152.

Torrent, Tiago (2015), 'On the relation between inheritance and change: The Constructional Convergence and the Construction Network Reconfiguration Hypotheses', in

Jóhanna Barðdal, Elena Smirnova, Spike Gildea, and Lotte Sommerer (eds), *Diachronic Construction Grammar*, Amsterdam: John Benjamins, pp. 175–214.

Traugott, Elizabeth C., and Graeme Trousdale. 2013. *Constructionalization and Constructional Changes*. Oxford: Oxford University Press.

Wolk, Christoph, Joan Bresnan, Anette Rosenbach, and Benedikt Szmrecsanyi. 2013. Dative and genitive variability in Late Modern English: Exploring cross-constructional variation and change. *Diachronica* 30(3): 382–419.

About the Series Editor

FUYIN (Thomas) LI (born 1963, Ph.D. 2002) received his Ph.D. in English Linguistics and Applied Linguistics from the Chinese University of Hong Kong. He is professor of linguistics at Beihang University, where he has organized the *China International Forum on Cognitive Linguistics* (cifcl.buaa.edu.cn) since 2004. As the founding editor of the journal *Cognitive Semantics* (brill.com/cose), the founding editor of *International Journal of Cognitive Linguistics*, editor of the series *Distinguished Lectures in Cognitive Linguistics* (brill.com/dlcl; originally *Eminent Linguists' Lecture Series*), editor of *Compendium of Cognitive Linguistics Research*, and organizer of ICLC-11, he plays a significant role in the international expansion of Cognitive Linguistics.

His main research interests involve Talmyan cognitive semantics, the overlapping systems model, event grammar, causality, etc. with a focus on synchronic and diachronic perspectives on Chinese data, and a strong commitment to usage-based models and corpus method.

His representative publications include the following: *Metaphor, Image, and Image Schemas in Second Language Pedagogy* (2009), *Semantics: A Course Book* (1999), *An Introduction to Cognitive Linguistics* (in Chinese, 2008), *Semantics: An Introduction* (in Chinese, 2007), *Toward a Cognitive Semantics, Volume I: Concept Structuring Systems* (Chinese version, 2017), *Toward a Cognitive Semantics, Volume II: Typology and Process in Concept Structuring* (Chinese version, 2019), both volumes were originally published in English, written by Leonard Talmy (MIT, 2000).

His personal homepage: http://shi.buaa.edu.cn/thomasli

E-mail: thomasli@buaa.edu.cn; thomaslfy@gmail.com

Websites for Cognitive Linguistics and CIFCL Speakers

All the websites were checked for validity on 14 February 2020

Part 1 Websites for Cognitive Linguistics

1. http://www.cogling.org/
   Website for the International Cognitive Linguistics Association, ICLA

2. http://www.cognitivelinguistics.org/en/journal
   Website for the journal edited by ICLA, *Cognitive Linguistics*

3. http://cifcl.buaa.edu.cn/
   Website for China International Forum on Cognitive Linguistics (CIFCL)

4. http://www.brill.com/cose
   Website for the journal *Cognitive Semantics* (*ISSN 2352-6408/ E-ISSN 2352-6416*), edited by CIFCL

5. http://www.degruyter.com/view/serial/16078?rskey=fw6Q2O&result=1&q=CLR
   Website for the Cognitive Linguistics Research [CLR]

6. http://www.degruyter.com/view/serial/20568?rskey=dddL3r&result=1&q=ACL
   Website for Application of Cognitive Linguistics [ACL]

7. http://www.benjamins.com/#catalog/books/clscc/main
   Website for book series in Cognitive Linguistics by Benjamins

8. http://www.brill.com/dlcl
   Website for Distinguished Lectures in Cognitive Linguistics (DLCL)

9. http://refworks.reference-global.com/
   Website for online resources for Cognitive Linguistics Bibliography

10. http://benjamins.com/online/met/
    Website for Bibliography of metaphor and Metonymy

11. http://linguistics.berkeley.edu/research/cognitive/
Website for the Cognitive Program at UC Berkeley

12. https://framenet.icsi.berkeley.edu/fndrupal/
Website for Framenet

13. http://www.mpi.nl/
Website for the Max Planck Institute for Psycholinguistics

Part II Websites for CIFCL Speakers and Their Research

14. CIFCL Organizer
Thomas Li, thomasli@buaa.edu.cn; thomaslfy@gmail.com
Personal Homepage: http://shi.buaa.edu.cn/thomasli
http://shi.buaa.edu.cn/lifuyin/en/index.htm

15. CIFCL 19, 2019
Francisco José Ruiz de Mendoza Ibáñez
francisco.ruizdemendoza@unirioja.es

16. CIFCL 19, 2019
Martin Hilpert
martin.hilpert@unine.ch

17. CIFCL 18, 2018
Arie Verhagen, A.Verhagen@hum.leidenuniv.nl
Website: http://www.arieverhagen.nl/

18. CIFCL 18, 2018 (CIFCL 12, 2013)
Stefan Th. Gries, stgries@linguistics.ucsb.edu
http://www.stgries.info

19. CIFCL 17, 2017
Jeffrey M. Zacks, jzacks@wustl.edu
Lab: dcl.wustl.edu
Personal homepage: https://dcl.wustl.edu/affiliates/jeff-zacks/

20. CIFCL 16, 2016
Cliff Goddard, c.goddard@griffith.edu.au
https://www.griffith.edu.au/humanities-languages/school-humanities-languages-social-science/research/natural-semantic-metalanguage-homepage

21. CIFCL 15, 2016
Nikolas Gisborne, n.gisborne@ed.ac.uk

22. CIFCL 14, 2014
Phillip Wolff, pwolff@emory.edu

23. CIFCL 13, 2013 (CIFCL 3, 2006)
Ronald W. Langacker, rlangacker@ucsd.edu
http://idiom.ucsd.edu/~rwl/

24. CIFCL 12, 2013
Stefan Th. Gries, stgries@linguistics.ucsb.edu
https://www.linguistics.ucsb.edu/people/stefan-th-gries

25. CIFCL 12, 2013
Alan Cienki, a.cienki@vu.nl
https://research.vu.nl/en/persons/alan-cienki

26. CIFCL 11, 2012
Sherman Wilcox, wilcox@unm.edu
http://www.unm.edu/~wilcox

27. CIFCL 10, 2012
Jürgen Bohnemeyer, jb77@buffalo.edu
Personal webpage: http://www.acsu.buffalo.edu/~jb77/
The CAL blog: https://causalityacrosslanguages.wordpress.com/
The blog of the UB Semantic Typology Lab: https://ubstlab.wordpress.com/

28. CIFCL 09, 2011
Laura A Janda, laura.janda@uit.no
http://ansatte.uit.no/laura.janda/
https://uit.no/om/enhet/ansatte/person?p_document_id=41561&p_dimension_id=210121

29. CIFCL 09, 2011
Ewa Dabrowska, ewa.dabrowska@northumbria.ac.uk

30. CIFCL 08, 2010
William Croft, wcroft@unm.edu
http://www.unm.edu/~wcroft

31. CIFCL 08, 2010
Zoltán Kövecses, kovecses.zoltan@btk.elte.hu

32. CIFCL 08, 2010
(Melissa Bowerman: 1942–2014)

33. CIFCL 07, 2009
Dirk Geeraerts, dirk.geeraerts@arts.kuleuven.be
http://wwwling.arts.kuleuven.be/qlvl/dirkg.htm

34. CIFCL 07, 2009
Mark Turner, mark.turner@case.edu

35. CIFCL 06, 2008
Chris Sinha, chris.sinha @ling.lu.se

36. CIFCL 05, 2008
Gilles Fauconnier, faucon@cogsci.ucsd.edu

37. CIFCL 04, 2007
Leonard Talmy, talmy@buffalo.edu
https://www.acsu.buffalo.edu/~talmy/talmy.html

38. CIFCL 03, 2006 (CIFCL 13, 2013)
Ronald W. Langacker, rlangacker@ucsd.edu
http://idiom.ucsd.edu/~rwl/

39. CIFCL 02, 2005
John Taylor, john.taylor65@xtra.co.nz
https://independent.academia.edu/JohnRTaylor

40. CIFCL 01, 2004
George Lakoff, lakoff@berkeley.edu
http://georgelakoff.com/