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](https://doi.org/10.6084/m9.figshare.13691296)

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.

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*. 
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*
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 3,000 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?
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.

---

**Figure 5**

BNC collocation frequencies (four-left, four-right window)

<table>
<thead>
<tr>
<th>terms</th>
<th>church</th>
<th>heart</th>
<th>eye</th>
<th>sigh</th>
</tr>
</thead>
<tbody>
<tr>
<td>abbey</td>
<td>30</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>above</td>
<td>13</td>
<td>8</td>
<td>26</td>
<td>0</td>
</tr>
<tr>
<td>always</td>
<td>28</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>half</td>
<td>8</td>
<td>0</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>long</td>
<td>28</td>
<td>11</td>
<td>7</td>
<td>48</td>
</tr>
<tr>
<td>family</td>
<td>47</td>
<td>10</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>gave</td>
<td>9</td>
<td>13</td>
<td>11</td>
<td>109</td>
</tr>
<tr>
<td>christ</td>
<td>141</td>
<td>8</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>walk</td>
<td>15</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

**Figure 6**

from collocate frequencies to a map
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](image.png)

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.
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.
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
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”.

---

**The acquisitive-to-permissive pathway**

- Van der Auwera et al. (2009: 272):
  - “*get* lends itself easily to the expression of [...] permission, and [...] it is plausible to relate this usage diachronically to a lexical verb meaning ‘acquire’.”

![Figure 12](image-url)
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.

---

**The acquisitive-to-permissive pathway**

- Van der Auwera *et al.* (2009: 272):
  - “*get* lends itself easily to the expression of [...] permission, and [...] it is plausible to relate this usage diachronically to a lexical verb meaning ‘acquire’.”

---

**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.
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.
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 as
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 as 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
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.
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.