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
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 Petréd 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.*
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.
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...
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 6**

Michael? Penzley here. Just calling to confirm that you’ve got the new Bundy ads ready for us today.

OK, before we begin the interview itself, I’d like to confirm that you have read and signed the informed consent form.
The starting point of the development is a form-meaning pair that can be expressed as \( \text{FORM}_1, \text{MEANING}_1 \). If that pair develops a new form, that constitutes constructional change, as we have seen with \textit{confirm} and its host-class expansion. As a result of that process, the original \( \text{FORM}_1, \text{MEANING}_1 \) pair is now accompanied by a second one, in which \( \text{FORM}_2 \) is associated with \( \text{MEANING}_1 \). Everyone would agree that this is constructional change, but then the development continues, resulting in a second meaning that is attached to \( \text{MEANING}_2 \). That would be \textit{confirm} with \textit{that}-clauses with the new meaning. Applying Traugott and Trousdale’s definition, we could say that the pair of \( \text{FORM}_2, \text{MEANING}_2 \) qualifies as a case of constructionalization. A first form-meaning pair, \( \text{FORM}_1, \text{MEANING}_1 \), gives rise to a second one, namely \( \text{FORM}_2, \text{MEANING}_2 \). There is however a complication.
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\textsubscript{0} on this slide. In that case, the original starting point, FORM\textsubscript{1}-MEANING\textsubscript{1}, is reached through one meaning change, which represents one step of constructional change. Once the new form develops from that, we have FORM\textsubscript{2}-MEANING\textsubscript{1}, counting back from FORM\textsubscript{1}-MEANING\textsubscript{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\textsubscript{1} and MEANING\textsubscript{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.
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 “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 pairings. 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,
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.

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
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.
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 grow 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.

**FIGURE 13**

**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**

- You shall know a word by the company it keeps!
- Modals that are semantically similar are expected to occur with similar collocates (at similar frequencies)
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.
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
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
Distinctive collexemes of *may*

<table>
<thead>
<tr>
<th></th>
<th>1860s</th>
<th>OBS</th>
<th>EXP</th>
<th>COLL STR</th>
</tr>
</thead>
<tbody>
<tr>
<td>be</td>
<td>8099</td>
<td>7517.13</td>
<td>46.43</td>
<td>130.24</td>
</tr>
<tr>
<td>say</td>
<td>277</td>
<td>206.94</td>
<td>15.95</td>
<td>70.9</td>
</tr>
<tr>
<td>do</td>
<td>190</td>
<td>136.74</td>
<td>14.31</td>
<td>31.36</td>
</tr>
<tr>
<td>add</td>
<td>79</td>
<td>50.67</td>
<td>12.30</td>
<td>11.21</td>
</tr>
<tr>
<td>judge</td>
<td>51</td>
<td>31.13</td>
<td>10.94</td>
<td>22.39</td>
</tr>
<tr>
<td>hope</td>
<td>46</td>
<td>28.08</td>
<td>9.87</td>
<td>11.69</td>
</tr>
<tr>
<td>form</td>
<td>36</td>
<td>21.98</td>
<td>7.72</td>
<td>9.56</td>
</tr>
<tr>
<td>trust</td>
<td>34</td>
<td>20.76</td>
<td>7.29</td>
<td>9.09</td>
</tr>
<tr>
<td>meet</td>
<td>43</td>
<td>37.47</td>
<td>7.02</td>
<td>8.77</td>
</tr>
<tr>
<td>suppose</td>
<td>31</td>
<td>18.92</td>
<td>6.65</td>
<td>7.38</td>
</tr>
</tbody>
</table>

Comparing *may*$_{1860}$ and *may*$_{2000}$

- Permissive examples in the 1860s
  - *If I may say so,...*
  - *You may do that if you like.*
  - *Mrs. Chapman, I may add here, was an old friend.*
  - *It’s neoclassist, if I may judge by the character of its frescos.*

- Epistemic examples in the 2000s
  - *I may have told you ...*
  - *If the hives are itchy, antihistamines may help.*
  - *The police may want to speak with you.*
  - *If fillets are large, you may need to cook them in two batches.*
  - *This is not so radical a step as it may sound.*

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.
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.
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.
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,
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.