

GENERATIVE LINGUISTICS: AN INTRODUCTION

Анотація. Дана стаття є коротким загальним оглядом генеративної лінгвістики як розділу когнітивної науки та основних питань, які вивчаються генеративними лінгвістами. В роботі пояснюються принципові відмінності між когнітивною та генеративною лінгвістикою та наводяться приклади того, як генеративні лінгвісти підходять до вивчення мови, зокрема обговорюються правила утворення іменної групи в англійській мові.

Аннотация. Данная статья дает общий обзор генеративной лингвистики как части когнитивной науки, а также главных вопросов, которые изучаются генеративными лингвистами. Работа объясняет принципиальные различия между когнитивной и генеративной лингвистикой и приводит примеры того, как генеративные лингвисты подходят к вопросу изучения языка, в частности обсуждаются правила строения именной группы в английском языке.

Abstract. This paper provides a general overview of the field of generative linguistics, intended primarily for philologists as the main target audience. The paper discusses the questions that generative linguists occupy themselves with in their pursuit of understanding of what language is and what principles underlie our remarkable ability as humans to acquire (any) language as children and what kind of mental capacity humans must have developed in order to be able to acquire such a complex system of knowledge as grammar. The paper also touches briefly on differences between cognitive linguistics and generative linguistics, arguing that the latter should be viewed as a branch of cognitive science that relies first and foremost on the scientific method in the study of language. To provide a concrete example of the rather abstract concepts discussed in text, the paper provides a succinct discussion of the rules underlying the grammar of noun phrases in English (borrowed from Abels (2015)), briefly demonstrating the difference between what we believe to be part of our innate knowledge of language (the so-called principles) and what should be viewed as learnable on the basis of evidence provided by empirical data (the parameters that are subject to cross-linguistic variation). Finally, the paper concludes by providing a brief description of the various categories that linguists (especially syntacticians) can be classified into based on their specific interests in language (due to Chris Collins) that should give the reader an appreciation of just how varied the work of a generative linguist can be.

1. What is language and how do we study it?

When we think about the things that make us who we are, that make us human, we think about such things as the fact that we are in charge of our own lives, that we are able to foresee consequences of our own actions and predict (to some extent) the actions of others, depending on the situation we find ourselves in. We also think about our ability to smile and appreciate jokes, form friendships and pass our knowledge and experience to our friends, colleagues and especially, to our children and future generations more generally. And of course we think about our ability to speak (or sign, in case of the deaf), our ability to understand at least the language(s) we call our native one(s), and our ability to accumulate knowledge and share ideas with others through language. What we rarely think about unless it is our job to think about language is just how remarkable language truly is and that in fact, most of the other accomplishments we humans can achieve largely depend on ability to acquire and use language, which Pinker (1995) calls our “preeminent trait”.

What is it that makes language so special? What do we really know when we know a language? Is it a long list of lexical items, the words (the vocabulary) and a set of rules? How are those lexical items and rules stored, and in what form are they represented in our brain, given we are not explicitly taught our native language? (if you think about it, you’ll agree we simply “pick it up” from our environment as kids). How is it possible that children learn their language with the same success and going through the exact same language development milestones, whether they come from households where their attempts to speak get plenty of reinforcement and encouragement from their parents (think middle-class parents everywhere in Western societies) or whether their parents or other caretakers barely talk to them beyond what’s minimally necessary (there are in fact societies where it is believed children are not interesting enough to talk to until they reach a certain age, - by which, of course, they are fully competent speakers of their native language (Pinker 1995, p.40). The latter question becomes especially puzzling if you consider how vastly different the languages of the world are, how much variation there is, which has led linguists to question the idea that languages are different at a fundamental, underlying level of organization, resulting in the (largely agreed upon) hypothesis that the differences are relatively superficial. When you think about the additional fact that the language that children hear around them does not even provide the perfect input - after all, both adults and other children who are fluent, **competent** speakers of a language occasionally make **performance** mistakes, use not fully grammatical or not fully formed sentences, employ ellipsis and rely heavily on context and pragmatics, you begin to fully appreciate the monumental task children acquiring their native language are faced with (this is what Noam Chomsky, the father of generative linguistics, referred to as “poverty of the stimulus” (Chomsky 1980).

It is exploring and providing answers to these fascinating questions that is at the core of the discipline known as **generative linguistics**. Just 60 years old (with

Chomsky's *Syntactic Structures*, published in 1957, laying the foundations for the discipline), this branch of **cognitive science** (which also includes psychology, computer science, philosophy and neurobiology) and the scientists who consider themselves generative linguists have made tremendous advances in providing answers to or non-trivial insights into these questions.

2. Generative Linguistics: Language as Science

For those who are not generativists, calling a discipline that occupies itself with language “science” may sound unusual, or even overly ambitious. However, generative linguistics is, indeed, deeply rooted in **the scientific method** and in the belief that one can study language employing the same rigorous scientific reasoning as is employed when studying physics, chemistry, astronomy or geometry.

So how does one study language as science? After all, our knowledge of a language as native speakers is not something that can be observed directly in the same way as celestial bodies can (in simple cases) be observed via telescopes or as human cells can be observed under the microscope. If human language and the rules that comprise the grammar are indeed properties of the mind, as generative linguists believe, we cannot observe and study language in the same way, by direct observation, and so studying language as property of the mind is like studying dark matter (only easier), - by having to examine available evidence and calculating and positing what must be true of our human language capacity in order for the language to manifest itself in humans the way it does. But neither does science more generally rely exclusively on that which is visible and always directly observable. Some of the most fascinating discoveries in science have been made relying on the kind of data and theories about the data that cannot be directly observed. Thus, just because a complex set of rules we call grammar, which we take to reside in the human mind, cannot be observed directly by scanning the brain, for instance, does not mean that we cannot study the grammar in a scientifically rigorous way. In fact, the body of linguistic knowledge accumulated by linguists since Chomsky's *Syntactic Structures* provides convincing evidence that the language can and should indeed be a subject of rigorous scientific study.

To study language by relying on the scientific method means that what we come to believe about what language is and how it functions comes from all and only the kinds of assumptions for which empirical data, empirical support can ultimately be found. This, of course, does not preclude one from forming complex theories and postulating objects, which are *not* directly observable, but the existence of which can be inferred from the linguistic data that *is* available. Occasionally, the existence of such abstract objects is fully confirmed empirically by coming across new linguistic data years after the principle has been postulated and relied on as a useful theoretical construct. This situation is very common in other sciences, such as astronomy, for instance, with the existence of objects and entities such as Neptune, dark matter, Higgs bosons, etc., being first predicted and

postulated, but not observed until later technological advances made this possible. In generative linguistics such unobserved/inferred objects were empty categories: PRO subjects, the traces that block “to” contractions in English, etc., (I am grateful to Richard Larson (p.c.) for a clarifying discussion of this point).

Saying that linguists rely on the scientific method in their exploration of language thus means linguists do not cherry-pick the data that suit their ideas, but instead form ideas based on what empirical evidence leads them to believe, formulating hypotheses about a particular aspect of language and verifying the predictions the hypotheses make. A good theory, about language or anything else that is studied using the scientific method, will be one, where the theory makes predictions that can be tested against more data and is thus falsifiable. At this point, all of this may and probably does sound very abstract, but we will see in a moment that this is indeed how linguists study language scientifically. A big part of what draws scientists to studying language and to becoming linguists is that every topic, from what structure noun phrases have to Genitive of Negation, to word order and especially its variability cross-linguistically, poses a puzzle, and every good linguist (indeed, any good scientist) loves solving puzzles.

Of course, as with every discipline and subject matter there are different ways to view the object under investigation and to approach the questions that arise in the process of studying it. To give one such salient example, any linguist will tell you that just as languages of the world have an astounding number of properties that they share, the number of differences between them (and which, unlike the similarities, are much more readily amenable to direct observation), is truly impressive. One of the questions to be asked in this context is, what do we attribute the vast number of differences between languages to and whether we take those differences to be important for the brains and minds of those who speak what appear to be dramatically different languages on the surface? One school of thought, taking its origin with the famous *Sapir-Whorf Linguistic Relativity Hypothesis* (see, for instance, Whorf 1956) takes the observable differences between languages having to do with how languages “carve up” the world they describe to directly affect native speakers, leading to cultural differences and differences in perception of objective reality. Thus, there are languages that have many more words to describe a certain phenomenon than others (many of you will think of the quite infamous example of the Eskimo having around 50 words to describe snow in its various forms (Boas 1911, pp.25-26) or certain other languages that lack lexical distinctions where yet other languages make them (think of the English word “blue” where Ukrainian has both “блакитний” and “синій”). Those who subscribe to the tenets of the *Linguistic Relativity Hypothesis* in one form or the other believe that such differences ultimately affect the way speakers of such distinct languages perceive reality (that is, that they have different **world views** depending on what their language allows or does not allow them to express). The Linguistic Relativity Hypothesis has found strong supporters among **cognitive linguists**, with George

Lakoff, focusing on the metaphoric uses of language in particular, rekindling interest in the idea. The idea has been revived again to some extent more recently, most prominently in the work of an American cognitive scientist Lera Boroditsky, whose clever psycholinguistic experiments on languages ranging from Russian to English to Mandarin have provided support for *the weak version* of the Linguistic Relativity Hypothesis.

This view, exciting as it is, departs drastically from the view of language taken by generative linguists, who believe that on some underlying level, all human languages are regulated by the same set of **principles** that allows any human being to acquire any human language as long as the humans in question find themselves in a community that speaks the language natively before the end of a **critical period** during which kids easily *acquire* any language natively (which stands in sharp opposition to the notorious difficulties adults encounter when attempting to consciously *learn* a foreign language). It is these **universal principles** underlying all human languages, as well as the **parameters** languages (minimally) differ on that are responsible for the observable variation between languages, along with the arguably innate mechanism for acquiring them, known together as **Universal Grammar** (henceforth UG; see Chomsky 1981).

The main goal of generative linguists then, irrespective of a particular set of personal research interests is to study language scientifically in an attempt to figure out which properties of any given language stem from easily observable in the input (and thus easily learnable) language-specific variation, and which properties reflect general properties of language, which tend to be too abstract and too complex to reasonably be expected to be figured out by children in the process of language acquisition, and which are thus believed to be part of our **innate language endowment, the Language Faculty** (Chomsky 1986). In the words of Steven Franks, "...the distinctly human capacity to "project" a grammar from primary language data can be explained in terms of some kind of language-learning cognitive system" (Franks 1995, p.viii). This way of studying the language, then, - figuring out what is language-specific and variable, and what is universal, is believed to ultimately help us provide deep insights into the workings of this language-learning cognitive system and the human mind more generally - the ultimate puzzle all cognitive scientists, linguists included, are puzzled and fascinated by. And so to learn about the mind, linguists study grammar, which can be understood as "an internalized system of rules and representations by which we compute and manipulate linguistic structures". We do this in the belief that "[b]y studying the properties of grammars, we may learn about the structure of our language faculty. Thus, we regard grammatical analysis as a theory of mind, which makes "specific empirical claims about abstract mental representation" (Franks 1995, p.viii).

3. Principles and Parameters

Thus, when generative linguists talk of Universal Grammar – the holy grail of linguistics, we talk of a set of **principles**, which are believed to be innate and universal, holding across languages, thus enabling the children to acquire any language they are exposed to effortlessly and in an amazingly short period of time (by the time children go to school they are already fluent speakers of their native language(s)). As already mentioned, what allows for the vast variability between the languages is believed to be the **parameters**, or the minute details of the implementation of these **universal principles** (on one of the views). Such view allows for an elegant explanation of how children manage to master the formidable task of figuring out the complex set of rules in their native language so easily: they come “prewired” with the knowledge of inborn, universal principles and all they have to do is learn the lexicon (the vocabulary of their language) and set the parameters to the value allowed by their particular native language (with the evidence for the parameter setting being readily observable in the input, that is, the language they are exposed to in their linguistic environment). It is important for the reader to understand that the Universal Grammar idea briefly sketched above is a theoretical construct, a framework, and as such may turn out to be imprecise in how we currently understand it. Yet it is currently our best guess as to how humans come to learn such a complex system as natural language in the absence of direct instruction and with usually impoverished information available in the language data children are exposed to.

One of the underlying tenets of generative grammar is thus that all sentences of human languages are generated through an application of a set of internalized (subconscious) rules that reside inside the human mind. It is what allows children to generate original sentences, rather than simply repeat the sentences they heard from adults over and over again. And if you have been around small children at all, you know how original their language use is, and if you pay enough attention, you will notice that they actively (if completely subconsciously) refine their understanding of the rules their language is governed by (throughout their language acquisition process they make grammatical mistakes that could have never been the result of simple memorization and repetition, but reveal the complex workings of their mental language faculty). And of course we know that the process of forming grammatical sentences is as natural and automatic to any native speaker as walking or discerning colors (see Barros 2014 for a more detailed discussion). These rules are thus drastically different from the conscious rules drilled into our heads at school, - they are not about where to put commas in sentences or not stranding prepositions in questions (if your language is English) or about avoiding split infinitives. We know these rules are subconscious, since any native speaker of a language has no problem distinguishing a grammatical sentence of their language, (exemplified in (1)), from an ungrammatical one (2), or from a grammatical

sentence that seems somehow “off” by virtue of not making much sense semantically but being nevertheless perfectly well-formed (3).

- (1) Mary gave her favorite book to John as a Christmas present.
- (2) *To John as a Christmas present gave her favorite book Mary.
- (3) Colorless green ideas sleep furiously.

As a standard notation practice, ungrammatical sentences are marked with a star/asterisk symbol at the beginning of the sentence; sentences that are borderline-grammatical or questionable in this respect are marked, quite appropriately, with one or more question marks at the beginning and sentences that are infelicitous for any reason are marked with a pound sign (#). Thus we could in principle mark the sentence in (3) with the latter to indicate that it is somehow “off”, even if completely grammatical in terms of its syntax. The sentence in (2), by the way, may well be a grammatical sentence of some other language (one that unlike English has a *free*, rather than *fixed word order* and employs a reordering mechanism known as *Scrambling*), but at the first glance most likely it will strike you as a sentence Master Yoda from *Star Wars* could have produced (not by virtue of its meaning, but by virtue of a “funny” word order that it employs). The sentence in (3) is probably one of the most famous sentences of its type, due to Noam Chomsky, who used it in his groundbreaking (1957) *Syntactic Structures* to show how we can form (and understand as such) fully grammatical sentences of our native language that are nevertheless nonsensical from the semantic point of view.

The goal of any generative linguist then (and syntactician in particular) is to examine the grammatical sentences of the language under investigation and to come up with a model or a set of rules that a native speaker of this language must have internalized as a child in order to produce all and only the grammatical sentences (with the ungrammatical sentences providing an important source of information and insights as well). Thus, to use a fancier expression, we might say that the task of a linguist is to “reverse engineer” human language, to recreate the grammar, - the complete set of subconscious rules that any native speaker possesses, so that we may gain insights into how our cognitive system operates (see Koenen and Zeijlstra (2017) for more discussion of this point). We should concede, of course, that this is an incredibly ambitious endeavor that has not been fully achieved yet, or else Google-translating (complex and occasionally less so) sentences from one language to another would not be such a painful and often embarrassing task. Yet, whatever successes are achieved in this area are in large part a credit to the work done by linguists and other cognitive scientists and if you are one of millions of the young and the ambitious dreaming to work at Google or Facebook at some point in the future, learning generative linguistics is as great an idea for a first step in this direct as learning how to code.)

To appreciate the rules of the language that linguists deal with and need to figure out, let us consider an example of rules governing the structure of noun phrases in English (the discussion below is borrowed from Abels (2015)). The recurrent phrase *those three green jellybeans* in the example sentences in (4) are referred to as a *noun phrase*, and simple transformations of these examples show that a noun phrase behaves as a single syntactic unit, which linguists refer to as a **constituent**.

- (4) a. Those three green jellybeans are tasty.
b. Michelle is looking at them.
c. She will give them to her mother.

We can see that *three green jellybeans* is a constituent since the pronoun *them* in (4b) and (4c) replaces the whole noun phrase without changing the overall meaning of the sentences (with *them* still meaning *three green jellybeans*). The data in (5), on the other hand, show that the words that comprise the noun phrase come in a specific, fixed word order, so that *jellybeans* always has to come last within the noun phrase:

- (5) a. (i) Those jellybeans are tasty.
(ii) *Jellybeans those are tasty.
b. (i) Three jellybeans are tasty.
(ii) *Jellybeans three are tasty.
c. (i) Green jellybeans are tasty.
(ii) *Jellybeans green are tasty.

Furthermore, the order of the words comprising the noun phrase when all of them are present is also fixed, thus the relative order of the words within the phrase (the demonstrative *those*, the numeral *three*, and the adjective *green*) with respect to each other and with respect to the final noun also matters, and when all of them occur within the noun phrase at the same time (as in (6d)), the order is completely fixed:

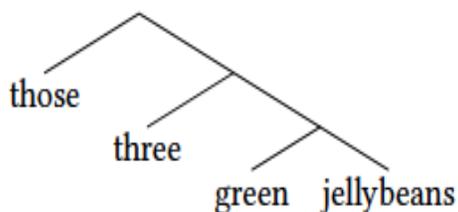
- (6) a. (i) Michelle is looking at those three jellybeans.
(ii) *Michelle is looking at three those jellybeans.
b. (i) Michelle is looking at those green jellybeans.
(ii) *Michelle is looking at green those jellybeans.
c. (i) Michelle is looking at three green jellybeans.
(ii) *Michelle is looking at green three jellybeans.
d. Michelle is looking at those three green jellybeans.

Finally, in addition to showing that the noun phrase *those three jellybeans* forms a syntactic unit, called a constituent, with further transformations (in this case, by employing **ellipsis**, – an operation where part of the original constituent phrase is deleted yet is nevertheless implicitly understood as being present in the sentence as it is previously mentioned in the immediately preceding and salient part of the sentence), we can show that this constituent has internal organization within the constituent, referred to as *internal syntactic structure* of the constituent. This is shown in (7) through (9), the point being that the missing/elided part of the sentence is still always understood as corresponding to the noun phrase mentioned overtly in the first part of the sentence, commonly referred to as **antecedent of the ellipsis** (with the antecedent marked by underlined bolded text). The examples (4) through (9) are borrowed from Abels (2015), with only a few small alternations made to them but with the point about the examples in the original discussion preserved. The notation used in the examples (7) through (9), such as underlining and small caps (used to mark stress) also reflects the original notation in Abels (2015).

- (7) I'll eat TWO **green jellybeans** and you can eat THREE.
- (8) I'll eat THESE **green jellybeans** and you can eat THOSE.
- (9) I'll eat THESE **three green jellybeans** and you can eat THOSE.

What the above examples in (7)-(9) demonstrate is that *green jellybeans*, *three green jellybeans* and *those three green jellybeans* are all constituents. This internal constituent structure can be represented in linguistics either by using **bracketed structures**, as in (10), or tree structures, as in (11):

- (10) [those [three [green jellybeans]]]
- (11)



Comparing the two types of notation, we can see that they represent the internal structure of the noun phrase in exactly the same way, thus showing that *green jellybeans* can serve as an independent constituent, and so can *three green jellybeans*, and so can *those three green jellybeans*, reflecting the relative word order of the demonstrative, the numeral and the adjective with respect to the noun phrase *jellybeans*.

As further shown in Abels (2015), by considering comparable data from

Spanish where the word order in the noun phrase is slightly different but the hierarchical structure is shown to be the same, we may (tentatively) conclude that such directly observable properties of English and Spanish as their word order is a manifestation of a **parameter** (something that can be inferred on the basis of observable data and thus can be acquired or learned), whereas the hierarchical structure (shown to be the same for noun phrases in both languages but which is not easily observable) is a **principle**, or part of what must be universally given. Clearly, one of the tasks that linguists face is describing the rules of various languages, so that the principles and parameters that we posit rest on generalizations derived from the empirical data of a large number of languages, with (possibly apparent) counterexamples taken seriously and evaluated with respect to the theory *du jour* in order to determine its accuracy.

4. What Kind of a Linguist are You?

For someone not familiar with generative linguistics the sheer number of ideas packed into this necessarily short introductory paper may seem a bit overwhelming at first. To convince the reader that linguistics is fun and linguists do in fact love what they do and wear the “linguist” label proudly on their sleeve, I will conclude this paper by providing a description of various types of linguists (or more precisely, the types of linguists doing mainly syntactic work), proposed by an NYU syntactician Chris Collins. What follows below is his description (reproduced here with his permission), which is both informative for a novice and can double as a fun “linguistic personality test”, - so read up to learn which types ring true to you (note that you may find, as most linguists do, that more than one or even two types are needed to describe what each individual linguist does or is interested in doing, with the author of this paper identifying as a 5-2-3, with some feeble attempts at 7 in the past and serious ambitions to qualify for a 6 at some point).

(1) The Skimmer

The skimmer looks at lots of languages (dozens) to address some theoretical point or to formulate a universal principle. They enjoy reading grammars and reading obscure papers on less studied languages. They are not so interested in giving an in-depth description of a particular language, although they might indulge in this from time to time. What is more important is to try to understand the nature of UG from the point of view of massive cross-linguistic comparison. The skimmer should not be confused with a typologist, who does not believe in generative syntax, and is often hostile towards it.

(2) The Theoretician

The theoretician starts from some clear set of assumptions and tries to give an original analysis in formal terms of some relatively well known phenomena (e.g., *that*-trace effects). Usually the theoretician does not push forward empirical

understanding, but his/her work often sets the stage for others who will.

(3) The Language Specialist

The language specialist can often be seen huddling at conferences with like-minded people all working on the same language or the same small group of languages. They have highly specialized and detailed knowledge of their language. They often use terms that are not understood outside their specialization. Their theoretical contributions attempt to make sense of the intricate and mysterious properties of their language.

(4) The English Syntactician

The English syntactician has a voluminous memory, knowing every possible counter-example to theoretical generalizations in English. They can also cite by date and title all the articles that have appeared in the early volumes of *Linguistic Inquiry*. Their golden grail is to discover some unknown and theoretically interesting fact about English (these are called "cool facts").

(5) The Syntax/Semantic Interface Specialist

This type of syntactician knows quite a bit of semantics. They have read all the classical texts and have mastered a formal semantics framework. To the lay person, they might be confused for a semanticist. But their work turns toward the syntactic, showing how syntactic and semantic principles interact or how to account for semantic phenomena using a sophisticated syntactic framework.

(6) The King/Queen of Construction

The King/Queen have built their career on a single construction (or closely related set of constructions). They wrote their thesis on that topic, and most of their current work is also on that topic. They serve as a kind of clearinghouse or point person for the field, writing periodic handbook articles, and giving invited talks on their construction.

(7) The Experimentalist

This is a new breed, which I may not be in a good position to characterize well. Their work approaches traditional syntactic question from an experimental perspective, meaning lab work and/or Mechanical Turk survey work. The experimentalist does not push forward theoretical or empirical understanding, but helps us to understand the status of various linguistic generalizations and principles that already exist.

5. Conclusion

In this very brief paper I have attempted to explain to a lay person what generative linguistics is and what generativists do, what questions occupy them and why this

is in fact a serious science to be reckoned with. I also hope to have convinced the reader that human language is far more fascinating and way more complex than is generally appreciated and most importantly, I hope to have convinced you that generative linguistics can indeed be studied the way “hard sciences” are and that this enterprise holds great promise for the whole field of linguistics (whatever linguistic stripes you may choose to wear) and to our understanding of the workings of human mind in general.

References

- Abels, Klaus. 2015. Syntax. In *Exploring Language and Linguistics*, Braber, Natalie, Cummings, Louise and Morrish, Liz (eds.), Cambridge University Press.
- Barros, Matthew. An Introduction to the Study of Language. Preliminaries. Ms, Rutgers University.
- Boas, Franz. 1911. Handbook of American Indian languages, 83 pages.
- Boeckx, Cedric. 2006. *Linguistic Minimalism: Origins, Concepts, Methods and Aims*. Oxford University Press, 246 pages.
- Boroditsky, Lera. 2001. Does language shape thought? Mandarin and English speakers' conceptions of time. *Psychological Science*, 13(2), 185–188.
- Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.
- Chomsky, Noam. 1980. *Rules and representations*. Oxford: Basil Blackwell.
- Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris Publications. Reprint. 7th Edition. Berlin and New York: Mouton de Gruyter.
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin, and Use*.
- Chomsky, Noam. 1986. *Barriers. Linguistic Inquiry Monograph Thirteen*. Cambridge, MA and London: The MIT Press.
- Franks, Steven. 1995. *Parameters of Slavic Morphosyntax*. Oxford University Press: New York.
- Koenenman, Olaf and Hedde Zeijlstra. 2017. *Introducing Syntax*. Cambridge University Press: Cambridge.
- Lakoff, George. 1987. *Women, fire, and dangerous things*. University of Chicago Press.
- Pinker, Steven. 1994. *The Language Instinct: How the Mind Creates Language*, Perennial.
- Sapir, Edward. 1921. *Language: An Introduction to the Study of Speech*, Harcourt, Brace.
- Whorf, Benjamin. 1956. Carroll, John B., ed., *Language, Thought, and Reality: Selected Writings of Benjamin Lee Whorf*, MIT Press.