1 What Is a Theory of English Syntax About?

1.1 Linguistic and Syntactic Competence

We language users believe that we 'know' a language, but the question is what we know when we know a language, like English or Korean. It may mean that we know how to create natural English sentences like (1a) but not unnatural sentences like (1b):[1]

(1) a. We can't pay for health care benefits like this, but you can.
b. *We can't keep paying for health care benefits like this, but you can keep.[2]

In the same way, speakers who know English may accept (2a) and (2c), but not (2b):[3]

(2) a. She swam.
b. *She swam the passengers.
c. She swam the passengers to three nearby boats.

This implies that knowing a language means that (English) speakers have linguistic knowledge sufficient to distinguish between 'acceptable' and 'unacceptable' sentences. However, when speakers are asked to articulate what kind of knowledge allows them to make these distinctions, it is not easy for them to describe it. This knowledge of language, often called **linguistic competence**, is the ability to speak a language. Knowing one's native language requires neither skill nor talent, but it is nonetheless an accomplishment worthy of investigation.

Linguistic competence involves several different levels of language structure. It includes phonetic and phonological competence: knowledge of the sounds

[1] The example in (1a) is from the corpus COCA (Corpus of Contemporary of American English), a collection of 560 million words of text from five different genres including spoken, fiction, magazine, newspaper, and academic texts. Throughout this book, we will use many corpus examples (extracted mainly from COCA) to portray English as it is actually spoken. We will, however, suppress their exact sources in the interest of readability.

[2] The notation * indicates that the particular example is ungrammatical or unacceptable. The notion of grammaticality (grammatical or ungrammatical) is closely related to that of acceptability (acceptable or unacceptable). Grammaticality has to do with whether a given sentence conforms to the rules and constraints of the relevant grammar, while acceptability has to do with whether a native English speaker would judge the sentence to be an instance of native English. Unless a distinction is required, we use these notions interchangeably.

[3] These examples are based on those used by Goldberg (1995). See Chapter 4.5 for discussion of such sentences.

of the language and their pronunciation variants. Linguistic competence also includes morphological competence. English native speakers, for example, can decide which words are (or could be) English and which are not. They (implicitly) know the rules for forming words, enabling them to make the past tense of an unfamiliar verb like *winter* or a new verb like *google*, as illustrated by the following corpus examples:

(3) a. Swallows wintered beneath the lakes.
b. She googled his name and discovered ninety-four hits.

Semantic competence includes the ability to determine the meaning of a particular sentence from the words of the sentence and their manner of combination. Native speakers distinguish the meanings of the following two sentences, which contain the same words but in different word orders:

(4) a. The dog chased the cat up a tree.
b. The cat chased the dog up a tree.

English speakers also interpret sentences flexibly, according to interactional context, enabling them to give appropriate responses to each. Consider the following utterances:

(5) a. Can you give me an aisle seat? (said at an airport check-in counter)
b. Can you pass the maple syrup, please? (said at a dining table)

The speaker's intent in uttering such sentences is not just to inquire about the hearer's ability but also to request an aisle seat and the syrup, respectively. The person to whom such a question is directed can infer that it is actually a directive.

The pivotal competence that we are concerned with in this book is syntactic competence: the ability to combine words into phrases that conform to the phrasal patterns of the language. Children learn these patterns without explicit training. How exactly they do so is a matter of controversy. Some linguists claim that certain aspects of grammar must be **innate**, because children do not receive enough data during early development to determine what the patterns are. Others argue that syntactic competence is in fact something that a child acquires through learning, and that the proponents of innate grammar have overlooked children's outstanding capacity to imitate adult routines and to infer patterns from rich but noisy input. We do not attempt to resolve this controversy here, because our focus is on what constitutes the adult's knowledge of language, and not the means by which it is achieved.[4]

Although children do not receive explicit instruction in their first language, they somehow gain the ability to produce all and only the grammatical sentences of their language and to distinguish grammatical sentences from ungrammatical

[4] We refer the interested reader to the rich literature on grammar learnability, which includes works by Goldberg (2006), Tomasello (2009), Newport (2016), and Chater and Christiansen (2018).

ones, as in (1). This kind of competence arises because language is **rule-governed**. One piece of evidence that it is a rule-governed system can be observed in word-order restrictions. If a sentence is an arrangement of words and we have five words, such as *player*, *ball*, *a*, *the*, and *kicked*, how many possible combinations can we have from these five words? Mathematically, the number of possible combinations of five words is 5! (factorial), equalling 120 instances. But among these 120 possible combinations, there is only a limited number of grammatical English sentences (including those that are semantically odd, as in (6c) and (6d)):[5]

(6) a. The player kicked a ball.
b. A player kicked the ball.
c. The ball kicked a player.
d. A ball kicked the player.
e. The ball, a player kicked.
f. . . .

Most of the combinations, a few of which are given in (7), are unacceptable to native speakers of English:

(7) a. *Kicked the player the ball.
b. *Player the ball kicked the.
c. *The player a kicked ball.

It is clear that there are certain rules in English for combining words. These rules constrain which words can be combined and how they can be ordered, sometimes in groups, with respect to each other.

Such combinatory rules also enable speakers to construct (or construe) complex sentences like (8a).[6] Whatever the combinatory rules are, they should give a different status to (8b), an example which is judged ungrammatical by native speakers even though the intended meaning is relatively clear.

(8) a. My parents decided to stay in the house they built.
b. *My parents decided to stay in the house they built it.

The fact that we require such combinatory knowledge also provides an argument for the assumption that we use a **finite** set of resources (expressions and rules) to produce and interpret grammatical sentences, and that we do not just rely on the meanings of the words involved. Consider the examples in (9):[7]

(9) a. I *(am) fond of that garden.
b. He *(is) angry at the not guilty verdict.

[5] Examples like (6e) are called 'topicalization' sentences: The topic expression (*the ball*), already mentioned or understood in a given context, is placed in a sentence initial position. See Lambrecht (1994), Gregory and Michaelis (2001), and references therein.
[6] In Chapter 2, we will begin to see these combinatory rules.
[7] The star * in front of the parenthesis symbols means that the expression within the parentheses cannot be omitted.

The omission of the copula verbs *am* and *is* would not prevent us from understanding the intended meaning, but the presence of these words is a structural requirement here.

In addition to being rule-based, syntactic competence powers the **creativity** (expressivity) that defines language ability. Speakers can produce and understand an infinite number of new grammatical sentences that the speaker has never spoken or heard before. For example, native speakers of English may have never heard, seen, or talked about the subject matter of sentences like (10) before, but they would have no difficulties producing or understanding such sentences:

(10) Forget intelligence or wisdom. A muscular physique might just be a more important attribute when it comes to judging a person's leadership potential, according to a new study.[8]

The expressivity intrinsic to grammatical competence is unbounded: A language user can produce and understand an infinite number of grammatical sentences. For example, given the simple sentence (11a), we can make a more complex one like (11b) by adding an adjective like *isolated*, which modifies the noun *nation*. To this sentence, we can add another adjective, *corrupt*, as in (11c). We could continue adding adjectives, theoretically enabling us to produce an infinite number of sentences:

(11) a. The nation faced sanctions.
b. The isolated nation faced sanctions.
c. The isolated, corrupt nation faced sanctions.
d. The isolated, corrupt, belligerent nation faced sanctions.
e. . . .

One might argue that since the number of English adjectives is limited, there should be a limit to this process. However, there are numerous examples in which we could keep such a process going, as shown by the following (Sag et al., 2003: 22):

(12) a. Some sentences can go on.
b. Some sentences can go on and on.
c. Some sentences can go on and on and on.
d. Some sentences can go on and on and on and on.
e. . . .

To (12a), we add the string *and on*, producing a longer one, (12b). To the resulting sentence, we once again add *and on* and make (12c). We could in principle go on adding without stopping: This is enough to prove that language has infinite creative potential (see Chomsky, 1957, 1965).

[8] Excerpt from the newspaper *Science Newsline*.

1.2 Generative Grammars

As discussed in the previous section, language is an infinite resource: There is no upper limit to the number of distinct grammatical arrangements of words that a native speaker can produce. Because any language user, at any time, can produce a grammatically licit string of words that she or he has never encountered before, it cannot be the case that language users have somehow managed to memorize every string of words that we might see or hear (Pullum and Scholz, 2002). Thus, a grammar cannot be an exhaustive list of possible word strings. It must instead be a model of our capacity to create (and understand) sentences of the language – a capacity referred to in the literature as competence. Nearly all syntactic theorists working today assume the following as a working hypothesis:

(13) All native speakers have **grammatical competence** that enables them to produce and understand an infinite number of grammatical sentences.

As reflected in (13), grammatical competence is a kind of discrete infinity: a limited repertoire of rules that allows us to make an infinite number of acceptable sentences. This grammatical competence is modeled by a **generative grammar**, which we then can define as follows (for English, in this instance):

(14) An English generative grammar is one that can generate an infinite set of well-formed English sentences from a finite set of rules or principles that do not generate any of the non-well-formed sentences.

The job of the syntacticians is thus to discover and formulate these rules or principles, which is also our goal here.

1.3 How We Discover Descriptive Rules

When talking about language rules or principles, we must recognize that there are two types of rules: prescriptive and descriptive rules. The key difference between the two is that **descriptive** rules capture naturally occurring language patterns, while **prescriptive** rules recommend certain usage practices. Prescriptive rules, as illustrated in (15), tell us how language ought to be used, rather than describing the language as it is:

(15) a. Do not end a sentence with a preposition.
b. Avoid split infinitives.
c. Use *who* rather than *that* to introduce a relative clause that describes a human.

However, the very existence of a prescriptive rule is good evidence that the targeted usage practice is commonplace, as suggested by the following attested 'violations':

(16) a. Who does she work with?
b. Young people need to try to boldly go where no one has gone before.
c. And she's the person that puts together the master list of songs.

Descriptive rules characterize whatever forms speakers actually use. One might have occasion to posit both prescriptive and descriptive rules, but the rule-governed grammar we are exploring in this book consists exclusively of descriptive rules.

The ensuing question is then: how can we discover the descriptive rules of English syntax – those that can generate all of the grammatical sentences but none of the ungrammatical ones? As noted earlier, these rules are part of our knowledge about language but are not consciously accessible; speakers cannot articulate their content if asked to do so. Hence we can discover the rules indirectly: We infer these latent rules from the observed data of a language. These data can come from speakers' judgments – known as intuitions – or from collected data of produced written or spoken language – often called corpora. Linguists use patterns in data to make inferences about an underlying phenomenon, and this is why we take linguistics to be an empirical discipline. The basic steps involved in doing such data-based linguistic research can be summarized as follows:

- Step I: Collect and observe data.
- Step II: Make a hypothesis to cover the first set of data.
- Step III: Check the hypothesis using more data.
- Step IV: Revise the hypothesis if necessary.

Let us now use these basic strategies to discover one of the grammar rules of English: the rule that distinguishes count and mass (noncount) nouns.[9]

Step I: Observing Data. To discover a grammar rule, the first thing we need to do is examine grammatical and ungrammatical variants of the expression in question. For example, let us look at the usage of the word *evidence*:

(17) Data Set 1: *evidence*
a. *The professor found some strong evidences of water on Mars.
b. *The professor was hoping for a strong evidence.
c. *The evidence that Jones found was more helpful than the one that Smith found.

What can you tell from these examples? We can make the following observations:

(18) Observation 1:
a. *evidence* cannot be used in the plural.
b. *evidence* cannot be used with the indefinite article *a(n)*.
c. *evidence* cannot be referred to by the pronoun *one*.

[9] The discussion and data in this section are adapted from Baker (1995).

In any form of scientific research, one example is insufficient to enable us to draw a conclusion. However, we can easily find more words that behave like *evidence*:

(19) Data Set 2: *equipment*
- a. *We had hoped to get three new equipments every month, but we only had enough money to get an equipment every two weeks.
- b. *This is a large truck which has an equipment to automatically bottle the wine.
- c. *The equipment we bought last year was more expensive than the one we bought this year.

We thus extend Observation 1 a little bit further:

(20) Observation 2:
- a. *evidence/equipment* cannot be used in the plural.
- b. *evidence/equipment* cannot be used with the indefinite article *a(n)*.
- c. *evidence/equipment* cannot be referred to by the pronoun *one*.

It is usually necessary to find contrastive examples to understand the range of a given observation. For instance, words like *clue* and *tool* act differently:

(21) Data Set 3: *clue*
- a. They hold vital clues to deciphering the history of the solar system.
- b. That would give us a good clue that something funny is going on.
- c. The clue that John got was more helpful than the one that Smith got.

(22) Data Set 4: *tool*
- a. The word clouds are good tools for engaging in critical thinking.
- b. Trade can be a powerful tool for global growth.
- c. The tool that Jones got was more helpful than the one that Smith got.

Unlike *equipment* and *evidence*, the nouns *clue* and *tool* can be used in the test linguistic contexts we set up. We thus can add Observation 3, different from Observation 2:

(23) Observation 3:
- a. *clue/tool* can be used in the plural.
- b. *clue/tool* can be used with the indefinite article *a(n)*.
- c. *clue/tool* can be referred to by the pronoun *one*.

Step II: Forming a Hypothesis. From the data and observations we have made so far, can we make any hypothesis about the English grammar rule in question? One hypothesis that we can make is the following:

(24) First Hypothesis:
English has at least two groups of nouns, Group I (count nouns) and Group II (mass nouns), diagnosed by tests of plurality, the indefinite article, and the pronoun *one*.

Step III: Checking the Hypothesis. Once we have formed such a hypothesis, we need to determine whether it is true of other data and to see if it has other analytical consequences. A little further thought allows us to find support for the two-way distinction among nouns. For example, consider the usage of *much* and *many*:

(25) a. much evidence, much equipment, much information, much advice
b. *much clue, *much tool, *much armchair, *much bags

(26) a. *many evidence, *many equipment, *many information, *many advice
b. many clues, many tools, many suggestions, many armchairs

As observed here, plural count nouns can occur only with *many*, whereas mass nouns can combine with *much*. Similar support can be found in the usage of *little* and *few*:

(27) a. little evidence, little equipment, little advice, little information
b. *little clue, *little tool, *little suggestion, *little armchair

(28) a. *few evidence, *few equipment, *few furniture, *few advice, *few information
b. few clues, few tools, few suggestions, few armchairs

The word *little* can occur with mass nouns like *evidence*, yet *few* cannot. Meanwhile, *few* occurs only with count nouns.

Given these data, it appears that the two-way distinction is plausible and persuasive. We can now ask if this distinction into just two groups is really sufficient for the classification of nouns. Consider the following examples with *cake*:

(29) a. She makes very good cakes.
b. The president was hoping for a good cake.
c. The cake that Jones got was more delicious than the one that Smith got.

Similar behavior can be observed with a noun like *beer*:

(30) a. I like good, dark, full-flavored beers.
b. No one knows how to tell a good beer from a bad one.

These data show us that *cake* and *beer* can be classified as count nouns. However, observe the following:

(31) a. My pastor says I ate too much cake.
b. The students drank too much beer last night.

(32) a. We recommend that you eat less cake and pastry.
b. People now drink less beer.

The data indicate that *cake* and *beer* can also be used as mass nouns, since they can be used with *less* or *much*.

Step IV: Revising the Hypothesis. The examples in (31) and (32) imply that there is another group of nouns: those that can be used as both count nouns and mass nouns. This leads us to revise the hypothesis in (24) as follows:

(33) Revised Hypothesis:
There are at least three groups of nouns: Group 1 (count nouns), Group 2 (mass nouns), and Group 3 (count and mass nouns).

We can expect that context will determine whether a Group 3 noun is used as count or as mass.

As we have observed thus far, the process of discovering grammar rules crucially hinges on finding data, drawing generalizations, making a hypothesis, and revising this hypothesis with more data. In addition, we have noticed that grammatical generalizations may actually be generalizations about classes of words, like the class of count nouns.

1.4 Two Different Views of Generative Grammar

We have seen that a theory of English syntax seeks answers to questions like how we can produce an infinite array of grammatical sentences and why some sentences are grammatical (or acceptable) while others are not. To answer such questions, we can derive generalizations from the observations of examples or data under investigation, as we did when discovering the distinction between count nouns and mass nouns in English. Such reasoning is traditionally called **inductive reasoning**, which derives broad generalizations from specific observations. In inductive reasoning, then, the investigator draws a conclusion about what is probably the case given the evidence encountered thus far. By contrast, we can start out with a general statement, or hypothesis, and then test it against data to ascertain its validity. Such reasoning, called **deductive reasoning**, is often adopted in natural sciences like physics. In deductive inference, we thus investigate the consequences of a theory, asking ourselves what would have to be the case if the theory is valid.

Each of these two forms of reasoning is associated with a different view of generative grammar. Deductive reasoning is closely associated with the **transformational** or **movement-based** view, in which 'underlying structures,' which feature a transparent relationship between syntax and semantics, are altered by various operations that produce an array of 'surface realizations' of those patterns. One such relationship is that between a question and its declarative counterpart (e.g., *For whom will you vote?* vs. *You will vote for whom*). The array of acceptable sentences is determined by the conditions on the application of transformations: Some properties of a given underlying structure may block the application of a given movement rule. Transformational grammarians are engaged in a deductive enterprise because they follow the premise that human languages have all and only those (structural) properties that are expressible in the transformational formalism: a structure-building operation (in recent versions of this framework called *merge*) and a structural displacement operation (in recent implementations of the model called *move*). A transformational

grammar thus seeks maximally general analyses conforming to the categories, relations, and operations assumed to characterize the language faculty, as when, for example, a sentence's structural complexity is analyzed in terms of the number of operations required to derive that sentence from its underlying structure.

Inductive reasoning plays a key role in the development of the so-called **constraint-based** view. A constraint-based grammar simply enumerates all of the patterns that exist in a grammar from general to specific, without attempting to derive one from another. A sentence or phrase of the language is predicted to be permissible insofar as it conforms to one or more of the existing patterns. In what follows, we will consider some key ideas of these two frameworks.

1.4.1 Deductive Reasoning and the Nativist View

Deductive reasoning is the key mode of explanation in Chomsky's linguistic framework, which represents human grammatical competence as a 'generative engine' that can produce only the grammatical sentences of a language like English.[10] One key hypothesis accepted by proponents of Chomsky's theory of grammar is the **innateness hypothesis**, which is often called the *nativist view*.

The nativist view "takes as a basic assumption that children are 'hardwired' with linguistic knowledge that gives them access to structural representations in the absence of experience" (Thornton, 2016). These structural representations, as well as the computational mechanism that operates over them, are referred to collectively as '**universal grammar**' (UG) or the 'language faculty.' On this account, learning one's first language is simply a matter of determining, through

[10] The historical development of the Chomskyan view, also called Transformational Grammar, can be summarized as follows:

a. Standard Theory (1957–1965)
b. Extended Standard Theory (1965–1973)
c. Revised Extended Standard Theory (1973–1976)
d. GB (Government and Binding)/P&P (Principles and Parameters) Theory (1981–1990)
e. Minimalist Program (1990–present)

The Standard Theory, laid out by Chomsky (1957, 1965), is the original form of generative grammar, and introduces two representations for sentential structure: deep structure and surface structure. These two levels are linked by transformational rules. The next stage is the so-called Extended Standard Theory, where X-bar theory is introduced as a generalized model of phrase structures. The Revised Extended Standard Theory generalizes transformational rules as Move-α. These previous theories are radically revised in GB (Government and Binding)/P&P (Principles and Parameters) theory (1981–1990). GB theory, armed with subtheories like government and binding, is the first theory to be based on the principles and parameters model of language. The P&P framework also underlies the later developments of the MP (Minimalist Program), which tries to provide a conceptual framework for the development of linguistic theory (Chomsky, 1995).