

Chapter 6: *Syntax*

Klaus Abels

Key terms: constituent, principle, parameter, universal grammar, gloss, tree diagram, movement, c-command, grammar

Preview

This chapter examines the question of what limits there are in the way in which languages can differ from each other structurally. Whenever we utter a sentence in any language, the words come in a particular order and are grouped into phrases in a particular way. While it is obvious that words in a sentence are ordered, the organization into phrases is less obvious, often imperceptible. This chapter argues that the variation between languages is largely confined to perceptible properties of word order, while the imperceptible organization into phrases is the same – or very nearly so – in all languages. Whether this view is true and, if so, why, is at the heart of some of the most fundamental debates in linguistics with implications for all of cognitive science. The chapter starts by motivating the existence of abstract phrase structure and by outlining what kinds of facts the syntactic description of a language must account for. A sufficiently explicit discussion requires some technical tools and notions, which will be introduced. The chapter then explains the goals of a general syntactic theory: to delimit and explain the range of variation found in human languages. This is followed by a case study of the word order found in noun phrases across languages. The case study focuses on the idea that languages differ in word order but resemble each other in phrasal organization.

6.1 Introduction

This chapter introduces you to a problem that has driven research in theoretical linguistics since the middle of the twentieth century. This problem and its potential solutions have shaped the field of theoretical syntax, some of the most important debates within linguistics and with neighboring disciplines.

The problem arises from tension between two basic observations: First, human languages differ immensely in how words are arranged into sentences and ordered. Second, children pick up the language spoken in their environment quickly, at a young age; they end up with remarkably complex and largely similar **grammars** in the absence of explicit instruction.

The first observation suggests that grammars are very different from each other. The second observation suggests the exact opposite; if children acquiring a language had to sift through a vast space of radically different grammars in order to acquire their language, the process should be slow and error prone and it should require explicit instruction. Since it does not, grammars must be very similar to each other, despite

appearances. The problem then is to find the underlying unity among the superficial diversity of grammatical systems.

The chapter proceeds by first illustrating the kinds of facts that the grammatical description of a language must account for. Along the way it illustrates the remarkable complexity and abstractness of speakers' grammatical knowledge. To describe such facts adequately, tree graphs and some terminology about such graphs are introduced. The fact that languages vary across a broad range is illustrated on the basis of the variety of neutral word orders found in noun phrases across languages. It is then shown how the diversity of word orders can be factored into a set of very abstract rules that are true across all languages and a set of superficial and easily learnable rules, which vary across languages.

6.2 Syntax: the study of sentence structure

Syntax is the study of how phrases and sentences are constructed. The syntax of any given language is a precise and rigorous description of the rules that characterize the phrases and sentences of that language. More broadly, a general theory of syntax specifies the types of rules and rule systems found across all languages.

To see why linguists think about the syntax of languages in terms of rules, that is, in terms of ways of constructing sentences, rather than in terms of the sentences themselves, we need to consider an everyday fact about what it means to know a language: When you know a language, this knowledge enables you to understand and produce sentences you have never heard before. For example, you have never before heard or read the sentences in this chapter. Yet, you have no trouble determining that they are English and interpreting them. This usually happens so automatically and easily that we rarely stop to notice. Not only do we produce and understand sentences that we have never heard before, we can also judge whether a particular sentence is a sentence of our language or not:

- (1) a. Hyperintelligent jellyfish from Mars fly quickly.
- b. *Hyperintelligent from quickly fly Mars jellyfish.

The first example is recognizably a sentence of English; it has a clear enough, even if fairly odd meaning. The second example is not a sentence of English at all but merely a list of words. Knowing English, you know the difference. Example (1a) conforms to the rules by which English sentences are constructed but (1b) does not. Following standard practice in syntax, example (1b) has an asterisk (*) in front of it to indicate that it is felt to be deviant by native speakers of English.

Importantly, this judgment of deviance is not a judgment about meaning. Example (2a) is a grammatical, nonsensical sentence of English. Example (2b) on the other hand is ungrammatical in standard English though perfectly interpretable.

- (2) a. Complicated triangles are confusing for small dreams.
b. *Complicated games is confusing for small children.

Everybody who speaks English can easily make these judgments and many more like them. This indicates that to know English means to know, subconsciously, the rules of English sentence construction.

The grammatical rules that enable you to distinguish between the examples in (1) and (2a) depend in part on word classes, since the rules governing the placement of nouns in a clause are different from those governing the placement of verbs, adjectives or prepositions. Furthermore, the rules of the grammar must also ensure that the meaning of a sentence depends on its structure. Otherwise, all sentences that use the same words should mean the same thing, but this is clearly not true: *The cat caught a mouse.* does not mean the same thing as *The mouse caught a cat.*

Based on this, it is clear enough what it means to say that a sequence of words is a sentence of a particular language: We mean that native speakers of the language recognize that sequence of words as belonging to their language. Since the set of sentences that belong to any given language is infinite, linguists don't construe languages as sets of sentences but as rule systems, that is, grammars. For syntacticians this means that they study the rules and rule systems that allow the construction and interpretation of sentences.

Everybody who speaks a language tacitly knows the rules of the language, since it is the rules that allows them to produce and understand novel sentences. The grammar is therefore an abstract characterization of the computations going on in the speaker's and hearer's minds when producing and understanding sentences. The grammar makes claims about which sentences are part of the language and which ones aren't, what the structure is of the ones that are part of the language and what meaning they have. It does not make claims about how the grammar is implemented, what algorithms are used in the mind to compute these results, or how the computation unfolds in real time. Part of this task falls to the psycholinguistic modeling of real-time language production and perception. The grammar also does not model how the brain realizes these algorithms in its synaptic structure, of course.

This brief bird's eye characterization should be sufficient to place the following discussion of the rules governing noun phrase structure in various languages. We start by looking at some English examples. The examples have the recurrent sequence of words *those three green jellybeans* in them. Together these four words behave as a unit in the sentences; they form a phrase, also called a **constituent**. We can see that they behave as a unit because they can be replaced by a single word without changing the meaning. Thus in (3b) and (3c), *them* stands for *those three green jellybeans*.

- (3) a. Those three green jellybeans are tasty.
b. Rosa is looking at them.

The gloss informs you of the function and category of each word or morpheme.

We can now return to our discussion of the rules of English. In addition to the facts about word order mentioned above, the grammar of English has to capture the way words are grouped into constituents. We saw above that the words *those three green jellybeans* behave as a unit and, therefore, form a constituent. There is evidence for further organization within that constituent.

Consider the following example. The verb *eat* appears twice. The first time it takes the complete noun phrase *two green jellybeans* as its object. The second time the noun phrase is incomplete and consists only of the numeral *three*.

Underlining in the examples indicates which parts of the sentence to compare with each other and SMALL CAPS indicate stress.

(7) I'll eat TWO green jellybeans and you can eat THREE.

Think about the interpretation of the incomplete object in (7). It is natural to interpret it to mean *three green jellybeans*. Here, the missing part of the noun phrase means *green jellybeans*. *Green jellybeans* thus behaves as a unit and the words that form this unit can either be present or go missing together.

The same can be seen in the next example, where the incomplete object is naturally interpreted to mean *those green jellybeans* with the unit *green jellybeans* missing.

(8) I'll eat THESE green jellybeans and you can eat THOSE.

We have now established two facts: Example (3) shows that *those three green jellybeans* is a constituent and examples (7) and (8) demonstrate that *green jellybeans* is a constituent.

We can also show that *three green jellybeans* is a constituent. Example (9) makes this point. The incomplete noun phrase can mean *those three green jellybeans*, with the constituent *three green jellybeans* going missing as a unit.

(9) I'll eat THESE three green jellybeans and you can eat THOSE.

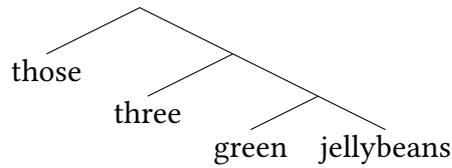
What we see is that *green jellybeans*, *three green jellybeans* and *those three green jellybeans* are all constituents. One can imagine the combining of words in terms of packing them together into boxes: Two things that are packed into one box form a unit. We thus have the following:

(10)

those	<table border="1" style="display: inline-table; vertical-align: middle;"><tr><td style="padding: 2px;">three</td><td style="padding: 2px;"><table border="1" style="display: inline-table; vertical-align: middle;"><tr><td style="padding: 2px;">green jellybeans</td></tr></table></td></tr></table>	three	<table border="1" style="display: inline-table; vertical-align: middle;"><tr><td style="padding: 2px;">green jellybeans</td></tr></table>	green jellybeans
three	<table border="1" style="display: inline-table; vertical-align: middle;"><tr><td style="padding: 2px;">green jellybeans</td></tr></table>	green jellybeans		
green jellybeans				

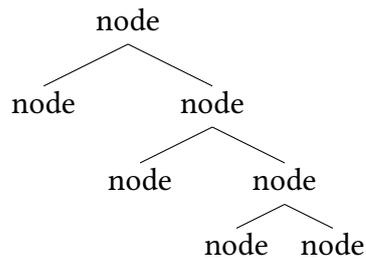
Instead of drawing boxes around groups of words, syntacticians usually express these facts using **tree diagrams**. The tree diagram below expresses the same information as (10).

(11)



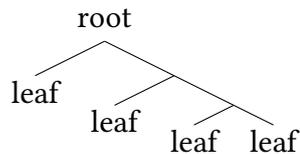
We will use such diagrams in the rest of this chapter, so it is worth introducing them in some detail. Every point in such tree diagrams where a line (called an edge) ends or where several edges meet is called a node.

(12)



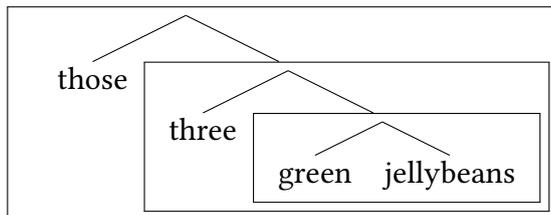
The nodes at the bottom and at the top of the tree are special. The ones at the bottom are called the leaves of the tree, (13), and this is where all the words that make up a sentence are drawn, as you can see in (11). The node at the very top of the tree is called the root.

(13)



If the leaves represent words, what do the remaining nodes represent? Each one of the remaining nodes in (11) corresponds to one of the boxes in (10). You can see this correspondence quite clearly in (14). Each node is the root of its own tree; each such tree corresponds to a structural unit, a constituent.

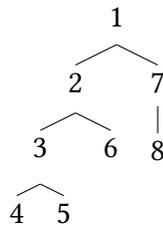
(14)



In the next tree, the nodes are numbered for ease of reference. We say that a node is the mother of all the nodes that it is connected to and which are immediately below it. Therefore, the root node 1 is the mother of nodes 2 and 7; node 2 is the mother of nodes 3 and 6; node 7 is the mother of node 8;.... The leaves (nodes 4, 5, 6, and 8) are not

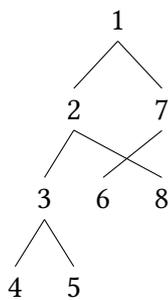
the mother of any other node. A node is the daughter of any node that it is connected to and that is directly above it. Thus, nodes 4 and 5 are the daughters of node 3; node 8 is the daughter of node 7; nodes 2 and 7 are the daughters of the root node 1. The root node is not the daughter of any other node. Finally, two nodes that have the same mother are called sisters. In the tree below nodes 4 and 5, nodes 3 and 6, and nodes 2 and 7 are sisters.

(15)



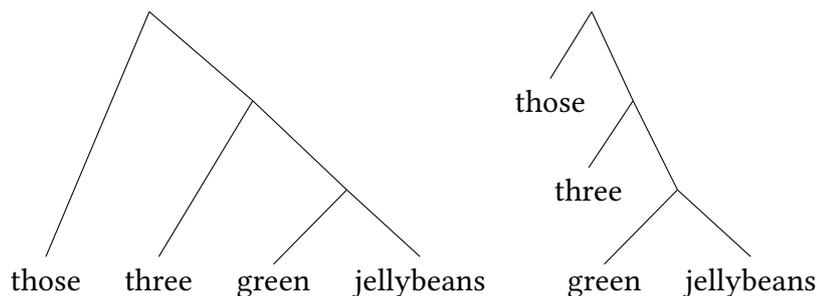
The edges of syntactic trees never cross. Diagram (16) is therefore not a permissible **syntactic tree**.

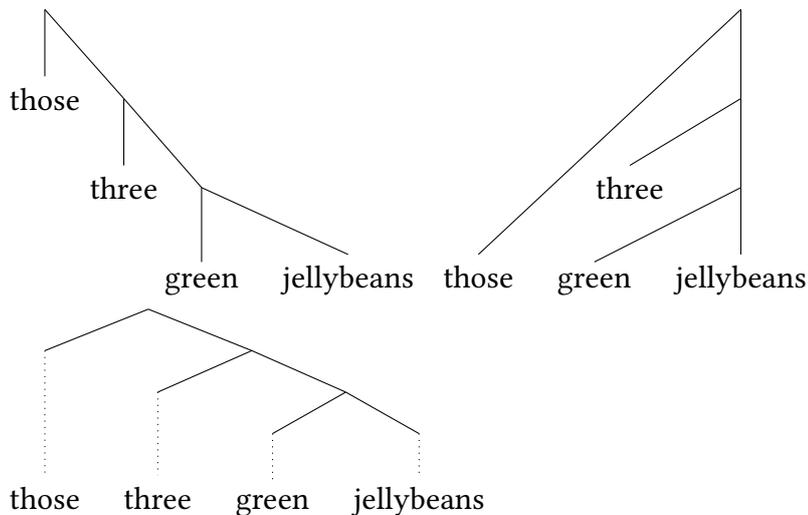
(16)



What matters for syntactic trees is which nodes are the mothers, daughters, and sisters of which other nodes and the order in which sisters come. The lengths or angles of edges play no role so long as branches don't cross. So all of the following trees mean the same as (11): the left-to-right order of the leaves represents the word order of the phrase.

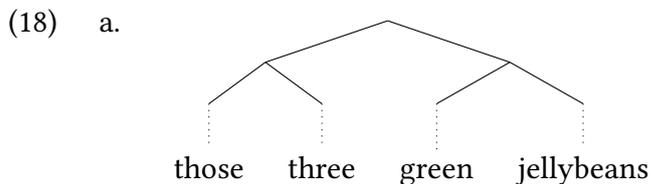
(17)

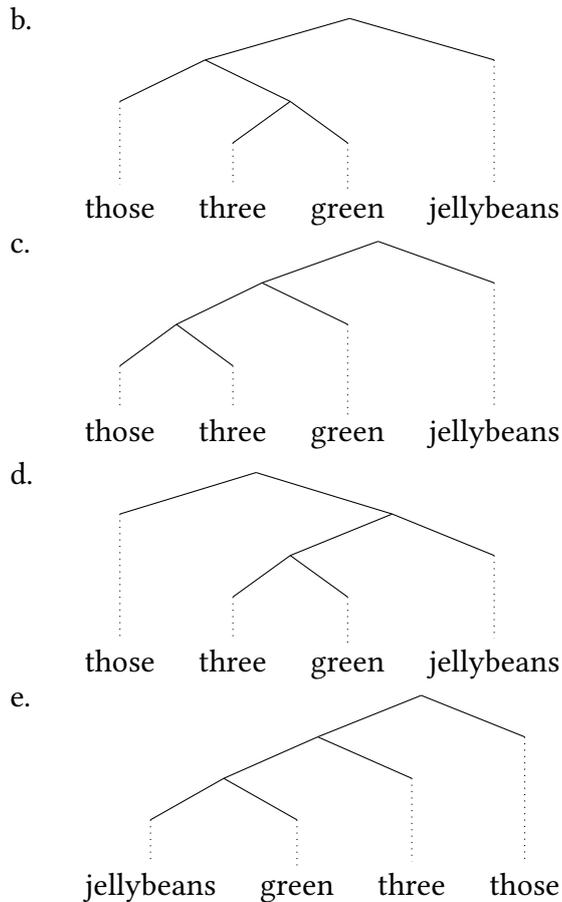




The final tree is particularly perspicuous, as it makes it easy to read off both the word order and the constituent structure of an expression. The dotted lines in the diagram have no significance, they are just a visual device connecting each leaf in the hierarchical structure of the tree to the word representing that leaf on the horizontal axis. Like edges in the tree, dotted lines may not cross. I will use trees like the final one from now on.

While the trees in (17) all look different but express exactly the same information, the following trees represent different information, because they either group or order the nodes differently. The structures in (11) and (17) all depict a constituent containing all and only the words *three green jellybeans*. We concluded that this is correct on the basis of example (9). None of the structures in (18a)–(18c) shows such a constituent. Instead, all three group the leaves *those* and *three* together. But there is no evidence that they form a constituent. The structures in (18b) and (18c) additionally group *those three green* together—a grouping that is not supported by the behavior of the noun phrase. (18b)–(18d) fail to group *green* with *jellybeans*—a constituent whose existence is supported by examples like (7). And (18d) claims that *three green* is a constituent—a claim for which there is no factual basis. Finally, (18e) is particularly interesting, since it gets the constituent structure exactly right—there is a constituent made up of *green* and *jellybeans*, one made up of *three*, *green* and *jellybeans*, and one made up of *those*, *three*, *green* and *jellybeans*—but the word order is wrong.



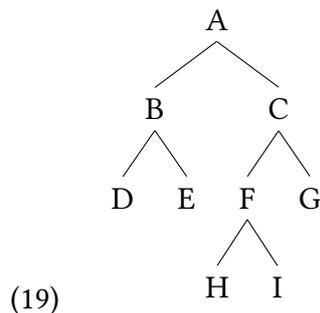


As we saw, the grammar of English dictates that the only way to combine a demonstrative, numeral, descriptive adjective and a noun into a single noun phrase is with the order and hierarchical structure in (11) and (17). (This is actually a simplification. Certain adjectives, like the words *visible* or *present*, can appear after the noun, others may appear after the noun only under special circumstances, yet others, like *other*, appear before the numeral. We will not deal with these complications here.) In the correct tree diagram for the English noun phrase, the demonstrative *those* is ‘higher’ than the other elements. What is meant by higher is that all the remaining words are contained in the sister of the demonstrative. Furthermore, the numeral is higher than the adjective and the noun. And the noun is (or, in more complex cases, is contained in) the adjective’s sister. Moreover, the demonstrative, numeral, and adjective all precede their respective sisters.

Since being ‘higher’ than another element is a very important concept in syntactic theory, syntacticians have coined a technical term for it. We say that one node in a tree **c-commands** another node if the second node is or is contained in the first node’s sister. Technically, we say that a node **dominates** all the nodes that it contains, or,

in other words, all those nodes that are below it and can be reached by following a continues path of edges that always go down and never up. Thus, in (19) B c-commands nodes C, F, G, H, and I and no other nodes. This is so, because C is B's sister, and F, G, H, and I are contained in — that is dominated by — C. B only has one sister, so the remaining nodes (A, B, D, and E) are not B's sisters. They are also not contained in B's sister because, in order to reach them from B's sister one would have to go up the tree at least once. Therefore, B does not c-command either A, B, D, or E.

Similarly, node C c-commands B, D, and E. D and E c-command each other and nothing but each other. Node A c-commands none of the nodes in the tree. Etc.



We can therefore rephrase our description of a correct English noun phrase by saying that the demonstrative must c-command any numerals, adjectives, and the noun within its noun phrase. The numeral must c-command any adjectives and the noun within its noun phrase. Adjectives must c-command the noun. Demonstratives, numerals, and adjectives precede their sister.

Key points: Syntax

- Linguists characterize grammars in terms of rules.
- Rules are necessary to explain why speakers can produce and understand novel sentences.
- The rules rely on abstract categories (nouns, verbs, adjectives,...).
- The rules characterize the word order and the abstract grouping of words into units.
- The structures of sentences formed according to such rules are usually represented by tree diagrams.

Exercise 6.1

Each of the examples below contains a clue that the underlined material is a constituent. Identify this clue and describe in what way the underlined material acts as a

unit. Provide a gloss for the underlined material and classify the examples according to the types of constituent that they identify.

- (20) He praised the student's fanciful suggestion but ignored the professor's.
- (21) These new products passed the test, but those failed.
- (22) You should read these two complicated papers because I cannot understand them.
- (23) I suggest we divide the work evenly so that I solve these two problems and you—those.
- (24) The customer preferred these red tiles, so we ordered them.

6.3 Syntactic theory and its goals

The approach taken in the previous section might seem foreign. The grammatical categories and tree diagrams introduced above have no direct analogue in your introspective everyday experience of language.

To explain this disconnect between your experience and those theoretical tools, we need to remind ourselves that linguistics is the scientific study of language. This means that we approach language with the same attitude of curiosity and detachment that is common to all of science: We should study language in the same way we investigate pendulums, chemical reactions, or the anatomy of beetles. Our introspective judgments and everyday experience of language can provide data but they should never be confused with the theory.

Precise descriptions of the rules that characterize the phrases and sentences of a given language provide answers to the question of what the grammar of that language is. They also lead to a number of further questions, all of which guide research in syntactic theory to a greater or lesser extent: What is the range of rule systems that we find in different languages? Are the rule systems similar or different? And why? How are these rule systems learned by children acquiring the language? How are the rules deployed by speakers in the real-time production and comprehension of utterances? To what extent are the rule systems shaped by the fact that they must be implemented in human brains and deployed in real time? To what extent are such rule systems shaped by the functions language has?

Different linguistic schools of thought emphasize some of these questions more strongly than others, but in some form, these questions define the broader goals of all current syntactic theorizing.

The answers to these questions are quite strongly interconnected. If it turned out that the rule systems of different languages varied in arbitrary and unpredictable ways, then children trying to learn a particular language would face a large and difficult task. If on the other hand it turned out that the rule systems of different languages

vary only in relatively minor ways or that the variation is constrained systematically, then children acquiring language face a much more narrowly circumscribed, and hence easier, task.

Noam Chomsky, the father of modern linguistic theory, has argued that the variation between languages must be fairly narrowly circumscribed. He reaches this conclusion on the basis of the assessment that language acquisition is relatively easy for children. This can be seen from a number of facts. First, language is acquired early by children – by the time they enter school, they have essentially acquired their native language. Secondly, language is acquired in much the same way by children with a rich and those with a poor linguistic environment and the ultimate level of attainment is comparable. Finally, children are able to acquire language on the basis of degenerate data; some of the sentences they hear are actually ungrammatical, because speakers make errors; some of the constructions that linguists have found to be highly revealing about the structure of sentences are very rare in everyday speech; children have no direct access to the intended meaning of sentences.

We have good reasons to believe that children get an important leg up in the process of acquiring language. Linguists generally believe that children are greatly assisted in their task by the fact that variation between languages is constrained and that there are grammatical properties shared across all languages. Grammatical properties that are shared across all languages are usually referred to as **principles** while points where we find restricted variation are called **parameters**. The system of principles and parameters together with a method for learning how the variable parameters are set for a given language is called **universal grammar**. In the next two sections we will discuss some of the principles and parameters that enter into the syntax of noun phrases.

The terminology of principles, parameters, and universal grammar introduced above is associated with the work of Noam Chomsky. Chomsky has sometimes suggested that universal grammar is innate and that it is task specific. The claim that universal grammar is innate is uncontroversial. Among other things, it explains why language disorders are heritable and therefore run in families (A. McMahon and R. McMahon, 2006) and why humans but no other species can acquire language with a syntactic organization. The question of whether universal grammar is a language specific capacity or whether the same mental mechanisms have more general application remains open at present (see for example Elsabbagh and Karmiloff-Smith, 2006). To answer the question, one needs to find out whether the principles and parameters discovered in language and the learning strategies used in language acquisition also characterize how we do other things like recognizing faces, planning movements, navigating in space, understanding social interaction, solving geometrical problems, solving ethical dilemmas, etc. Progress on this will come from constructing separate theories of the principles, parameters, and learning strategies involved in all of these different tasks and then to analyze carefully to what extent they can be unified. In what follows, we

will consider only the task of constructing an explicit account of the principles and parameters of linguistic structure.

There are a number of strategies researchers pursue to construct such an account. Conceptually the simplest approach is to construct grammars of different languages and to compare them carefully. We illustrate this method later on in this chapter using the noun phrase as an example.

A different, less obvious, strategy is to construct the grammar of only a single language and to assess to what extent its properties could be learned by a child hearing a representative sample of sentences of the language. If we discover properties that could not reasonably be learned that way and yet all native speakers end up acquiring these properties, we conjecture that those properties are in fact not learned but are given. Such properties have to form part of (or follow from) universal grammar.

The goals of syntactic theory are therefore to describe the grammars of different languages, to analyze which parts of the grammar are universal (the principles) and which parts are subject to variation across languages (the parameters), to create an account of how children acquire the grammar(s) of their native language(s), and to explain these findings in terms of language specific or more generally available cognitive constraints and processes.

The next two sections illustrate the search for linguistic universals.

Key points: Syntactic Theory

- The central goal of syntactic theory is to explain the way in which grammars are similar to each other and to circumscribe the ways in which they can differ from each other.
- Properties that characterize all grammars are called principles.
- Points of variation between grammars are called parameters.

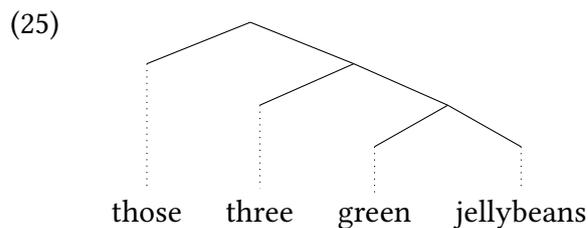
Exercise 6.2

Consider the following thought experiment: Suppose that biologists discovered a species of apes on an unexplored island who can communicate with each other but not with humans. The biologists study the apes' communication system and attempt to figure out its rules. Eventually, they come up with a grammar of the apes' language. The grammar is successful in capturing which utterances the apes accept as part of their language and which utterances they reject, as well as in correctly capturing the meaning the apes attribute to their utterances. Suppose furthermore that, to achieve this, the grammar has to posit complex rule systems and abstract entities (like the tree structures and syntactic categories of the previous section) of which the apes profess to have no introspective awareness.

Discuss to what extent the apes' lack of awareness of a complex rule system, of categories and abstract structures invalidates the biologists' findings and calls their truth into question.

6.4 Universal and language specific properties of phrase structure

Section 6.2 has shown that the grammar of English must represent the phrase *those three green jellybeans* as in (25).



We turn to noun phrases in Spanish now. In Spanish, the order of elements differs somewhat from English.

- (26) esos dos libros interesantes
 these two books interesting
 these two interesting books

The gloss indicates that the demonstrative and the numeral precede the noun, as they do in English, but the adjective follows it.

What structure should we assign to the Spanish noun phrase?

We know from (18a)–(18d) that a given sequence of words could logically be associated with a variety of structures. To decide which of them is correct, we used examples that reveal the grouping of words. Specifically, we considered the interpretation of incomplete noun phrases. Examples (7)–(9) showed that the adjective forms an exclusive unit with the noun and that the numeral forms an exclusive unit with the adjective-noun unit. We can use the same method to investigate Spanish.

The following example illustrates that the noun can be missing from a noun phrase in Spanish.

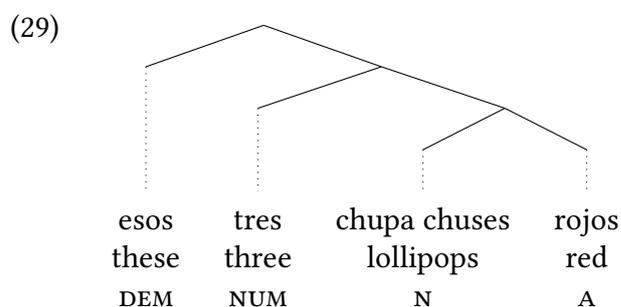
- (27) Yo voy a leer ESTE libro y tu - ESTE.
 I go to read this book and you - this
 I am going to read this book, and you are going to read that book.

With only a single word missing the example cannot tell us how words are grouped. We need to look at examples with several words missing. Relevant examples show

that the noun and the adjective form a unit, because they can be left out together: (28a). Likewise, numeral, noun, and adjective form a unit, because they can be left out together: (28b).

- (28) a. Yo voy a comer esos TRES chupa chuses rojos y tu - esos DOS
 I go to eat these three lollipops red and you - these two
 I am going to eat these three red lollipops and you are going to eat these two red lollipops.
- b. Yo voy a comer esos tres chupa chuses rojos y tu - ESTOS.
 I go to eat these three lollipops red and you - those
 I am going to eat these three red lollipops and you are going to eat those.

We depict the Spanish noun phrase by the following tree:



The diagram for the English noun phrase, (25), and that for the Spanish noun phrase, (29), differ in linear order but the hierarchical structure is identical.

Both in English and in Spanish the demonstrative c-commands any numerals, adjectives, and the noun; the numeral c-commands any adjectives and the noun; adjectives c-command the noun; demonstratives and numerals precede their sister. The only difference is the relative order the adjective and its sister. In English, the adjective precedes, in Spanish, it follows its sister.

A useful analogy might be that of a mobile. Depending on the way the arms of the mobile are turned, the words suspended at the bottom will appear in different orders, but its hierarchical organization in terms of what is attached to what remains constant. Phrased in the terminology of principles and parameters, we can say that the way an element is ordered with respect to its sister is a parameter while the hierarchical organization of the noun phrase, which is the same in Spanish and in English, is a candidate for a linguistic universal, a principle.

Of course, drawing any conclusions just on the basis of two languages would be rash. Nevertheless, we can consider the plausibility of the idea that the hierarchical organization of phrases is a principle. This idea makes good sense from the perspective of language acquisition: Word order is directly given to the child in the data she hears and is hence learnable, but it is impossible to infer the hierarchical structure of a phrase

solely on the basis of readily observable properties of sequences of words alone. Such considerations of learnability can provide an indirect argument for considering the hierarchical structure of noun phrases a linguistic principle.

In the remainder of this and the following sections, we will further explore the idea that the hierarchical structure but not the word order of the noun phrase represents a principle of universal grammar by looking at noun phrases from languages across the globe. The characterization of the facts builds on a large amount work comparing different languages, inspired by Greenberg 1963. The more specific claims come from Cinque, 2005 as re-interpreted in Abels and Neeleman, 2006. It is important to note that the only orders we consider are unmarked orders, that is, orders that can be used without putting special emphasis on any of the elements within the noun phrase. Special emphasis allows additional word order possibilities not discussed below.

We have already seen that in Spanish noun phrases have the same hierarchical organization as in English but the adjective follows its sister instead of preceding it. In other words, there is a parameter regulating the relative order of the adjective and its sister. We expect to find similar parameters regulating the order of demonstratives and numerals relative to their sisters. The following examples from the languages Dulong, a Tibeto-Burman language spoken in parts of China and Tibet, and Sranan, a creole language from Suriname, show that this expectation is borne out. In these languages word order in the noun phrase is like in English except that Dulong numerals and Sranan demonstratives follow the noun.

(30) Dulong simplified, based on LaPolla, 2003, p. 676

kō tái zǎjé ǎsūmpǎŋ

that big book three

Dem A N Num

those 3 big books

(31) Sranan simplified, based on Voorhoeve, 1962, pp. 33–36

den dri moi fooru dati

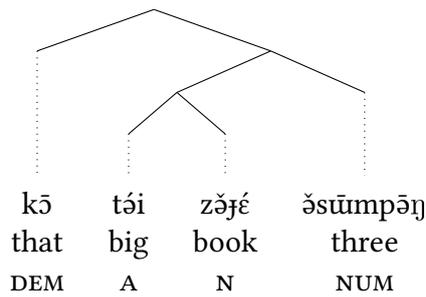
the three beautiful chickens these

Num A N Dem

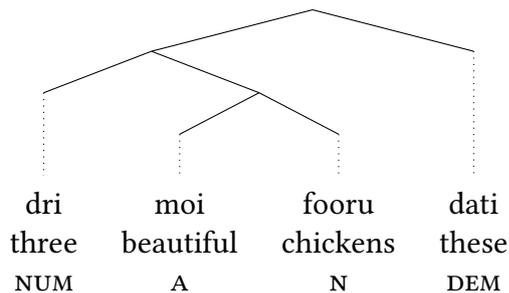
these 3 beautiful chickens

The noun phrases for these languages would be represented as follows:

(32) a. Dulong



b. Sranan—leaving out the article *den*—‘the’, since it does not concern us here



Herarchically these structures are identical to their English and Spanish counterparts. What sets Dulong and Sranan apart from English is only the position of the numeral and the demonstrative, respectively.

Of course, there is no reason to think that languages can set only a single ordering parameter differently from English. In many languages, all elements follow their sisters instead of preceding them. Gungbe, a Kwa language from Benin, is an example of this type. In this language, the order of demonstrative, numeral, adjective and noun is the mirror image of that found in English. This order type is very common among the languages of the world.

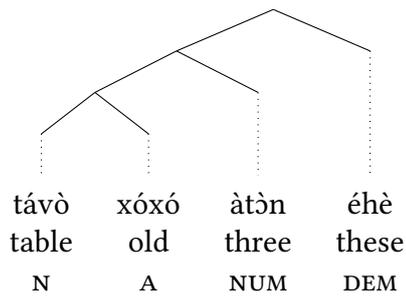
(33) Gungbe

Aboh, 2004, p. 78

távò dǎxó xóxó àtòn éhè ló lé
 table big old three DEM SPF.DF PL
 these specific three big old tables.

The tree structure for Gungbe is given in (34). This structure is simplified in that it ignores the morphemes *ló* and *lé* which have no direct correspondences in English. The glosses stand for specific definite and plural, respectively.

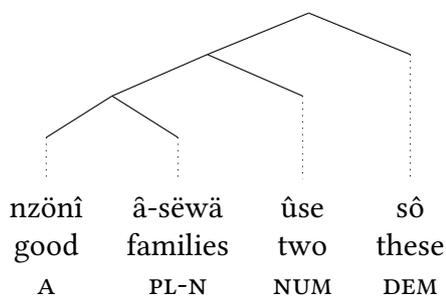
(34) Gungbe



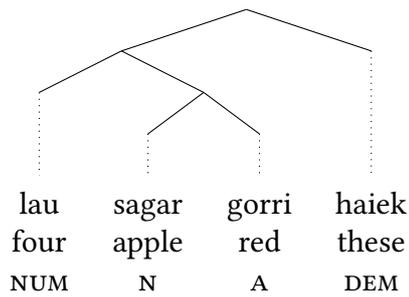
As before, this tree is identical to the structures we have seen above for English, Spanish, Dulong, and Sranan in hierarchical terms; only the order of elements is different.

There are three more logically possible orders that the single hierarchical structure we have been assuming produces without crossing any branches. The orders in question mirror those found in Spanish, Dulong, and Sranan: adjective-noun-numeral-demonstrative, numeral-noun-adjective-demonstrative, and demonstrative-noun-adjective-numeral, respectively. Indeed, languages with the word orders in question exist. The order adjective-noun-numeral-demonstrative is found in Sango, one of the official languages of the Central African Republic, numeral-noun-adjective-demonstrative is found in Basque, a language isolate spoken in regions of Spain and France, and demonstrative-noun-adjective-numeral – in Burmese, a Sino-Tibetan language spoken in Myanmar.

- (35) Sango: *these two good families* simplified from Thornell, 1997, p. 71 using Samarin, 1967, pp. 57–83

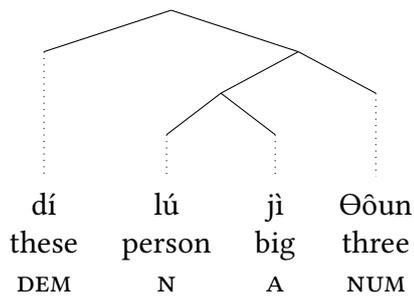


- (36) Basque: *these four red apples* Oyharçabal, 2012, p. 269



(37) Burmese: *these three big persons*

simplified from Jones, 1970, p. 5



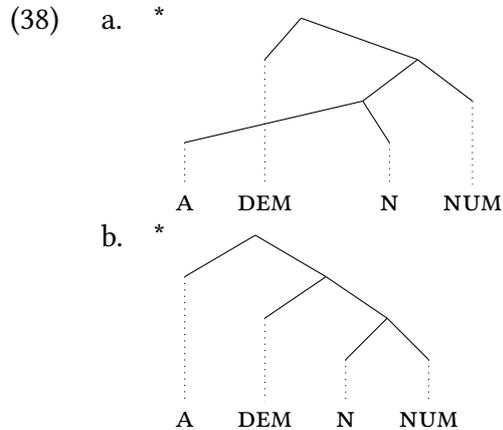
Again, the hierarchical structure in terms of c-command of the noun phrases above is identical to the now familiar structure, only the word order is different.

The idea that the hierarchical syntactic structure is a linguistic universal with the linear order being subject to parametric variation has produced the expectation that all of the above eight word orders should occur in some languages. As we saw, this expectation is borne out.

The proposed principle of universal grammar creates a second and far more important expectation. If the hierarchical structure of the noun phrase is universally fixed and the word order varies only in the way we have observed so far, then the eight orders given above should be the only orders we find since no further orders are possible without crossing any branches. To appreciate this point fully, we need to consider the fact that given four elements belonging to different classes, there are 24 logically possible orders ($4! = 4 \times 3 \times 2 \times 1 = 24$). We have now discussed eight of these orders and seen that, in line with our expectations, they occur as the unmarked word order in some languages. We have so far left the remaining 16 orders to the side. The proposed linguistic principle clearly predicts that none of them should occur.

According to Cinque, 2005, this expectation is partially correct; ten of the 16 remaining word orders conform with our expectations and never occur as the unmarked order. For example, no language employs the orders adjective-demonstrative-noun-numeral, adjective-numeral-noun-demonstrative, demonstrative-adjective-numeral-noun, and numeral-adjective-demonstrative-noun as the unmarked word order. If we try to construct trees corresponding to the unattested orders, such trees invariably either

have crossing branches, (38a), or violate the linguistic principle according to which the demonstrative c-commands everything in the noun phrase, the numeral c-commands everything but the demonstrative, and the adjective c-commands the noun, (38b).



This section has illustrated what linguistic universals, the principles of universal grammar, might look like and what the locus of language variation might be. We argued that, from the perspective of language learnability, the proposed structural principle of universal grammar makes sense as it simplifies the acquisition task greatly. At the same time, the variation resulting from the linear parameters are readily learnable. The system explored so far has the right key ingredients. It offers explicit and correct analyses for the languages discussed so far. It answers the question of which aspects of the grammar are universal and which aspects vary between languages. Finally, it provides a straightforward answer to the problem of language acquisition.

We have seen partial support for the proposed system of principles and parameters in so far as all expected word orders are attested as the unmarked order in some languages. Furthermore, some of the word orders that are expected to be impossible have indeed been claimed never to occur as the unmarked word order in any language (see Cinque, 2005 and references cited there). In the next section, we will discuss the six word orders that form counterexamples to the principles and parameters discussed so far and extend the system carefully so that it covers those orders but retains its pleasing properties of restrictiveness and easy learnability.

Key points: Phrase structure

- The neutral word order within the noun phrase in the vast majority of languages can be described by a set of abstract principles and superficial parameters as follows:
- principles
 - Within a given noun phrase, the demonstrative c-commands all other elements.
 - Within a given noun phrase, the numeral c-commands all other elements except for the demonstrative.
 - Within a given noun phrase, the adjective c-commands the noun.
- Parameters
 - The order of demonstrative, numeral, and adjective with respect to their respective sisters is subject to cross-linguistic variation.

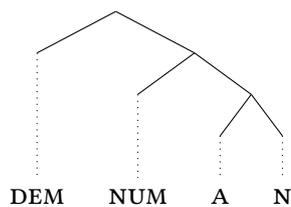
WEBSITE: Group exercise 1

Visit the website and, in a small group, investigate the question what motivation there is for the claim that demonstrative-numeral-adjective-noun is the neutral word order in English despite the fact that it is not, strictly speaking, the only possible word order.

Exercise 6.3

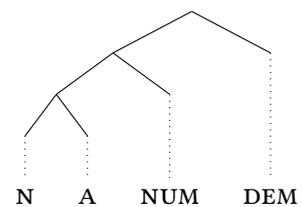
The word order and structures for the eight languages mentioned so far are repeated next to each other below.

(39) English

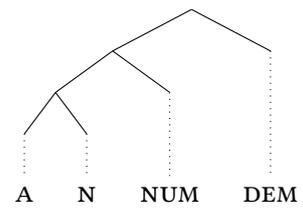
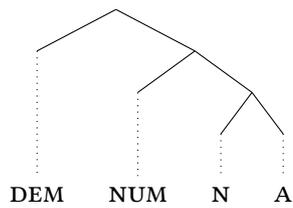


Spanish

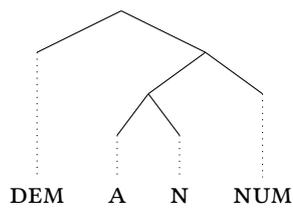
Gungbe



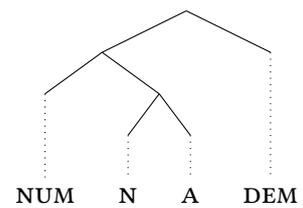
Sango



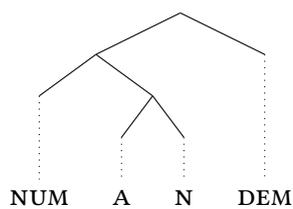
Dulong



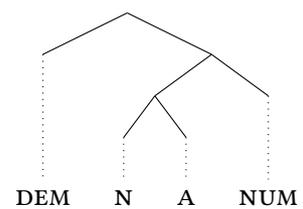
Basque



Sranan



Burmese



Verify for yourself that no further orders are possible without changing the structure and without illicitly crossing branches.

6.5 Universal and language specific properties of movement

Section 6.4 ended with a mixed conclusion. On the one hand, the idea that noun phrase structure but not the order is universally fixed finds support in the fact that all eight predicted orders are actually attested and that it solves the problem of learnability in the face of variation. Furthermore, some of the word orders that are not expected to occur are indeed unattested cross-linguistically. However, Cinque, 2005 documents six word order patterns that our current model does not predict. We turn to these orders now.

The first four new patterns are fairly simple variations on those already discussed.

The languages Maasai, a Nilo-Saharan language spoken in Kenya and Tanzania, Kĩitharaka, a Bantu language from Kenya, Kele, a nearly extinct Austronesian language

of Papua New Guinea, and Pitjantjatjara, an Australian language, illustrate four of the attested word orders not previously discussed.

- (40) Maasai Koopman, 2003

kù-n-dâ mésa-i àré sidân
 PL-F-that table-F.PL.ACC F.PL.two.ACC nice.PL.ACC
 DEM N NUM A

those two nice tables

- (41) Kĩtharaka Peter Muriungi (p. c.)

i-kombe bi-bi bi-tano bi-tune
 8-cup 8-this 8-five 8-red
 N DEM NUM A

these five red cups

- (42) Kele Ross, 2002, p. 132

pihin ha-mow il tóti
 woman one-CLF old this
 N NUM A DEM

this one old woman

- (43) Pitjantjatjara based on Eckert and Hudson, 1988, p. 89

Tjitji pala tjukutjuku kutjara
 child that small two
 N DEM A NUM

those two small children

These orders are beyond the reach of the theory developed so far. The following table compares the four new orders with attested orders from the previous section. Languages and orders already discussed are in unmarked rows, new orders are in rows marked with a double right arrow (\Rightarrow). The table is organized in such a way that all the ordering types that share relative ordering of demonstrative, numeral, and adjective are grouped together.

(44)	Language	Order				
	English	Dem	Num	A	N	
	Spanish	Dem	Num	N	A	
⇒	Maasai	Dem	N	Num	A	
⇒	Kiitharaka	N	Dem	Num	A	
	Sranan	Num	A	N	Dem	
	Basque	Num	N	A	Dem	
⇒	Kele	N	Num	A	Dem	
	Dulong	Dem	A	N	Num	
	Burmese	Dem	N	A	Num	
⇒	Pitjantjantjara	N	Dem	A	Num	

A simple generalization governs the table: In the four new patterns, the noun appears earlier in the string of words than our current theory allows. It seems as though the noun moves ‘to the left.’ Comparable movement of the noun ‘to the right’ never occurs: languages with the order shared in common by Sranan, Basque, and Kele or Dulong, Basque, and Pitjantjantjara but with the noun all the way at the end do not exist. This suggests that there is a principle banning postposing the noun.

The two remaining orders documented by Cinque (2005) are illustrated here using Aghem, a Niger-Congo language from Cameroon, and Banda-Linda, another Niger-Congo language from the Central African Republic.

(45) Aghem based on Hyman, 1979, p. 28

nwín fɪdú'ú fɪn fɪmɔ̃
 bird big this one
 N A DEM NUM

this one big bird

(46) Banda-Linda constructed based on Cloarec-Heiss, 1986, pp. 185–200

ōgbōrō yāfē sáyē bīfī
 tall women these two
 A N DEM NUM

these two tall women

Again it is useful to compare these orders to the ones we discussed in the previous section. The table groups languages together where the relative order of demonstrative and numeral on the one hand and of adjective and noun on the other hand is the same as in the languages with the new orders.

(47)

Language		Order		
	Spanish	Dem	Num	N A
	Burmese	Dem	N A	Num
⇒	Aghem	N A	Dem	Num
	English	Dem	Num	A N
	Dulong	Dem	A N	Num
⇒	Banda-Linda	A N	Dem	Num

Again there is a simple generalization. It appears that the constituent made up of noun and adjective in its language specific order appears earlier than our theory allows in both Aghem and Banda-Linda. (I should point out that Cloarec-Heiss, 1986, where the information on Banda-Linda word order comes from, does not analyze *ōgbōrō*—‘tall’ as an adjective but instead as a relational noun, which leads to an overall meaning more like *these two giants of women*. Though cross-linguistically rare, the order A N DEM NUM has also been reported for the languages Bai (Wiersma, 2003) and Koiari (Dutton, 1996).)

There are no further attested, unmarked word order patterns.

How can we account for this state of affairs? How do we allow the six new orders without threatening our conclusions about the universal structure of the noun phrase, which provided a solution to the learnability problem?

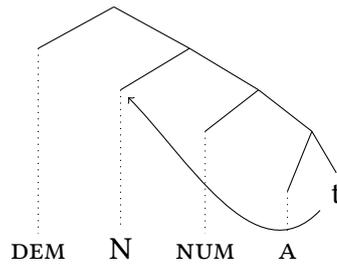
As a first step towards a solution we observe, that the new data from this section in some sense reinforces rather than challenges the significance of the universal hierarchical structure; this is so, because the additional orders can be given a very simple characterization in terms of this universal structure, but not otherwise: In the noun phrase, the noun or a constituent containing the noun may appear earlier than expected under the theory of noun phrase structure from section 6.4.

Since, for example, the numeral by itself is not a constituent containing the noun it never occurs earlier in the string of words than the theory from section 6.4 would predict: the order NUM DEM N A is unattested as an unmarked word order. Similarly, there is no constituent made up of noun and numeral to the exclusion of the adjective, even in those languages noun and numeral always occur right next to each other. Consequently, we do not find languages where just the numeral and the noun occur too early and hence, N NUM DEM A never occurs as the unmarked order.

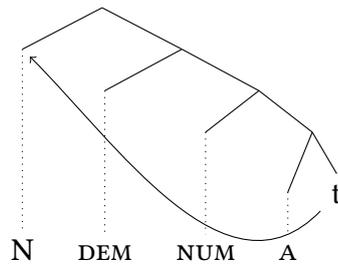
When an element appears earlier in the sequence of words than phrase structure theory would lead us to expect, syntacticians speak of movement. In English, the object of a verb usually appears immediately after the verb, but in questions it can be moved to the beginning of the clause. This accounts for the different placement of the object in *John read the book.* and *What did John read?*. If movement only impacted the word order without changing the structure, we would account for movement by allowing, in restricted circumstances, crossing branches in syntactic trees after all. However, there is strong evidence that movement changes order and structure.

In keeping with the conventional notation for movement structures in tree diagrams, the six new orders are drawn as shown in (48). The position indicated by *t* (mnemonic for 'trace') is an unpronounced, silent position. Its function is to indicate what the hierarchical position of the displaced element according to the basic phrase structure theory would be. There are also arrows connecting the trace and the movement element. These arrows differ from branches of the tree, since arrows are allowed to cross branches.

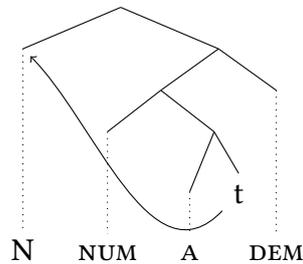
(48) a. Maasai



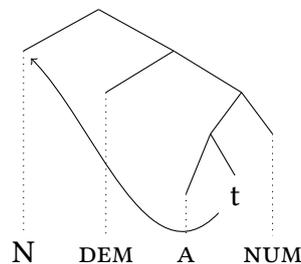
b. Kîitharaka



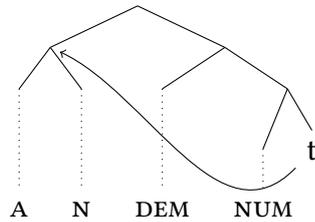
c. Kele



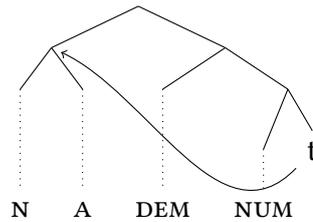
d. Pitjantjatjara - STRUCTURE CORRECTED COMPARED TO PUBLISHED VERSION



e. Banda-Linda



f. Aghem



We can summarize our exploration of the syntax of noun phrases as follows.

- (49) a. Universal principles of noun phrase construction
- (i) Ignoring the result of movement operations, noun phrases are hierarchically structured with any demonstrative c-commanding all other elements, any numeral c-commanding everything but the demonstrative, and any adjectives c-commanding the noun.
 - (ii) If **movement** occurs, the moved constituent occurs earlier in the sequence than it would have without movement, never later.
 - (iii) If movement occurs, the displaced constituent is or contains the noun.
- b. Language-particular parameters of noun phrase construction
- (i) The linear order of demonstrative, numeral, and adjective with respect to their sisters is subject to language-specific variation.
 - (ii) Whether movement happens or not is subject to language-specific variation.
 - (iii) If movement happens, the exact target position and the size of the moving constituent are subject to language-specific variation.

The system has a number of very pleasing properties. First, it derives the correct range of word order patterns. All and only the attested word order patterns can be derived within this system; we have seen how it derives the 14 attested word orders. The remaining ten logical possibilities cannot be derived without violating one of the universal principles.

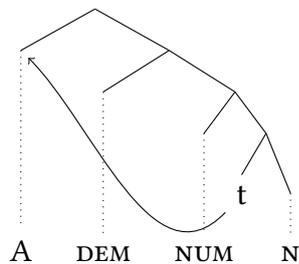
Exercise 6.4

Reconsider the order A DEM N NUM. The last section, in particular the discussion of the structures in (38), showed why the simple model of the noun phrase as a mobile could not produce this order. The order is unattested as an unmarked order and it remains outside the reach of the system even with the addition of movement.

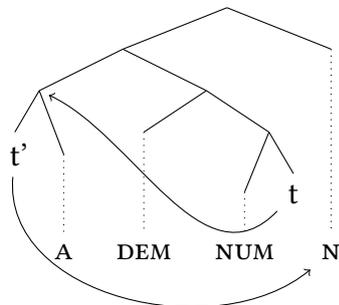
The structures in (50) illustrate this. Which order would have resulted from (50a) if movement had not happened? Which of the languages discussed in section 6.4 exemplified this order? Which of the elements moves in (50a)? Which of the conditions on movement in (49) is violated by this operation?

Now consider the structure in (50b), where movement has applied twice. Which order would have resulted if the second step of movement, that is, noun movement, hadn't occurred? Which language discussed in this section exemplifies this order? Which order would have resulted if neither of the movements had taken place? Which language from section 6.4 exemplifies this order? Which condition on movement stated in (49) is violated by (50b)?

(50) a.



b.



Exercise 6.5

Show that the even when the effects of movement are taken into account, the order A-NUM-DEM-N cannot be derived. (Hint: You will probably want to use the property that in (39) the English type of order is the only N-final one.)

In section 6.4, the idea was introduced that those properties of the noun phrase that vary across languages should be easily detectable in the input for the language learner. This idea underpinned our account of the noun phrase as a uniform hierarchical structure with varying linearization. The same idea emerges again from (49): Exactly which constituents move in a given language and where are easily detectable properties; they can be allowed to be parameterized. The effectively unlearnable hierarchical structure must be kept as a universal principle. The system now delivers precise analyses of all existing language types, rules out the inexistent ones, and solves the learnability problem. We have reached the end of our exploration of the noun phrase.

You may wonder at this point whether the principles and parameters discussed are valid outside of the narrow confines of the syntax of noun phrases. Indeed, many of them do. Consider the following example:

(51) Who are you talking to?

The example shows two things. First, the question word *who* moves to the beginning of the sentence to form a question. Second, the preposition does not move together with the question word—and in fact, in normal colloquial English it cannot. We saw above that in the noun phrase there is a parametric choice of whether movement happens to begin with, (49b-ii), and if so, whether the noun moves by itself or together with additional material, (49b-iii). The same type of parametric choice is also found in question formation. Some languages, like German, are very similar to English in that they, too, move their question words to the beginning of the sentence to form a question. However, unlike in English, prepositions have to front together with the question word.

(52) German

- a. Mit wem redest du?
with who talk you
Who are you talking with?
- b. *Wem redest du mit?
who talk you with

Other languages, like Chinese, are less similar to English and German in that they do not front their question words at all. This is shown in (53). Like English, Chinese generally shows the order subject-verb-object. This order is maintained in questions, (53). Leaving the question word at the end of the sentence in the English counterpart to this sentence would be impossible in a regular question.

(53) Chinese

Huang, 1982, 253 ex. 159

ni kanjian-le shei?
you see-ASP who
subject verb object

Who did you see?

These brief remarks illustrate that the system of parameters developed above has broader validity within syntax than just the narrow domain that we initially used to motivate them.

Key points: Movement

- Together with the structural principles and parameters developed in section 6.4, the following provides a comprehensive characterization of the neutral word order within the noun phrase across languages.
- Movement, which is invoked here, is regulated by a set of abstract, structural principles together with superficial, easy-to-learn parameters.
- principles of movement
 - To derive neutral word orders, only movement of the noun or a constituent containing the noun is licit.
 - To derive neutral word orders, only preposing movement is allowed.
- Parameters of movement
 - The exact landing site of movement is subject to cross-linguistic variation.
 - The exact size of the moving constituent is subject to cross-linguistic variation.

WEBSITE: Group exercise 2

Visit the website and, in a small group, investigate the question to what extent the layered, hierarchical structure for the noun phrase contributes to the explanation of cross-linguistic word order patterns discussed in this chapter.

6.6 Summary

This chapter characterized syntax as the study of how phrases and sentences are constructed. We analyzed English noun phrases and discovered the need for abstract structures and categories, which we represented with the help of tree diagrams. Section 6.3 introduced some of the broader, more ambitious goals of syntactic theory. These goals

are tied to the questions of language learnability, of the role played by nature and nurture in language acquisition and in variation between languages, and of the relative importance of language specific and domain general cognitive operations in shaping language.

The remaining two sections proceeded to develop an account of the neutral word order within noun phrases across the languages of the world. The account provides explicit analyses of noun phrases across languages, rules out unattested word order patterns, and provides a solution to the learnability problem. The presentation relied heavily on the tree notation from section 6.2, underscoring its usefulness.

Section 6.4 motivated and developed a model of the commonalities and differences between noun phrases in different languages. The model assumed that the hierarchical structure of noun phrases constitutes a principle of universal grammar and that the only variation between languages lies in the choice of the order in which two sister constituents appear. This model is partly successful but turns out to be slightly too restrictive.

Section 6.5 extended the model to achieve better cross-linguistic coverage without giving up the ability to solve the learnability problem. To this end further universal principles and a number of parameters of cross-linguistic variation in the realm of movement were posited. As a description, the system is very successful, since it allows all and only the actually attested word order patterns. The system remains sufficiently constrained to allow a neat solution to the learnability problem: those properties of noun phrase syntax that are very abstract are universally fixed and those that vary correspond to easily detectable properties of the input, namely, word order.

The theory of noun phrase syntax presented in this chapter illustrates the logic of the Principles and Parameters approach to language. The approach was first articulated by Noam Chomsky in the 1980s and developed into the Minimalist Program. All modern formal theories of syntax adopt the core idea of universal principles combined with a constrained set of parameters allowing cross-linguistic variation.

The success of this core idea comes from its descriptive strength and from its ability to explain the otherwise mysterious fact that children manage to acquire highly complex grammars of their native languages on the basis of poor and degenerate input.

The success of a theory is not only measured by the questions that it answers but also by the novel questions that it gives rise to.

After reading this chapter, you may wonder about the following questions, which would hardly have occurred to you before reading it and many of which couldn't be asked without the explicit account of noun phrase syntax provided here. Where do the principles that govern the hierarchy within the noun phrase come from? Is the syntactic mobile describing hierarchical aspects of the noun phrase learned or innately given? Does it characterize only language or other aspects of cognition, such as perception, as well? How would we find out? Does movement in the noun correlate with

other properties found in the languages that have it? Why do we find movement of the noun or a constituent containing it in some languages but not in others? Where does the restriction come from that the moved element has to move forwards in the sequence of words rather than backwards? Can the theory of word order in the noun phrase be generalized to other syntactic categories such as verb phrases or adjective phrases?

Some of these questions are syntactic by nature others can be addressed only through interdisciplinary work across the cognitive sciences. Linguists have long thought that the study of language can provide insights into the workings of the human mind. Such insights cannot be won without describing the rules characterizing different languages carefully or without delineating the boundary between possible and impossible linguistic systems. Likewise, the broad question whether the principles invoked in syntactic theory are specifically linguistic or characterize cognition more broadly can only be answered on the basis of the kind of articulated syntactic theory with a broad coverage of data hinted at in this chapter but not without such a theory.

WEBSITE: Syntax

Now go to the website and assess your knowledge of syntactic theory by completing the self-test questions!

Suggestions for further reading

James R Hurford (1994). *Grammar: A Student's Guide*. Cambridge, UK: Cambridge University Press

This book contains short articles on descriptive grammatical terminology. Among many other things, it treats adjectives, demonstratives, numerals, and nouns. The book contains exercises and is a useful guide to terminology.

Steven Pinker (1994). *The Language Instinct*. New York: Morrow

This book introduces the central problems and goals of syntactic theory. It is also a very readable polemic in defense of the position that many of the principles and parameters characterizing language are specifically linguistic and do not reduce to general cognitive mechanisms.

Richard K Larson (2010). *Grammar as Science*. Cambridge, MA: MIT Press

This book gives you an in depth introduction to syntactic concepts used in this chapter including phrase structure rules, tree relations, grammatical categories, and movement rules.

Answers to in-text exercises

Answer to 6.1:

In (20) ‘the professor’s’ is naturally interpreted as ‘the professor’s fanciful suggestion.’ ‘Fanciful suggestion’ behaves as a group, as the words go missing together. In (21) ‘those’ is naturally interpreted as ‘those new products.’ ‘New products’ behaves as a group, as the words go missing together. In (22) ‘them’ is naturally interpreted as ‘these two complicated papers,’ which behaves as a group, since it can be replaced by a single pro-form. In (23) ‘those’ is naturally interpreted as ‘those two problems.’ ‘Two problems’ acts as a group, since the words go missing together. Finally, in (24) ‘them’ is naturally interpreted as ‘these red tiles.’ ‘These red tiles’ acts as a group, as it can be replaced by a single pro-form: ‘them.’

Glossed examples, grouped according to the type of constituent that is picked out:

- (54) [A N]
 (20) fanciful suggestion
 A N
 (21) new products
 A N
- (55) [NUM N]
 (23) two problems
 NUM N
- (56) [DEM NUM A N]
 (22) these two complicated papers
 DEM NUM A N
- (57) [DEM A N]
 (24) these red tiles
 DEM A N

Answer to 6.2:

All of the elements DEM, NUM, and A can either precede or follow their sister. We start at the top left of (39) with the English order, where all of the above precede their sister. Moving clockwise, we invert only A for Spanish, only NUM for Dulong, and only DEM for Sranan. The languages in the right column represent the mirror image of those in the left column. Burmese inverts A and NUM, Basque—A and DEM, Sango—NUM and DEM, and Gungbe—all three. This exhausts all the options of what can be switched with what. Therefore, no further orders can be produced from this structure.

Answer to 6.3:

The biologists in our thought experiments are constructing an abstract, computational, symbolic theory of what is going on inside of the apes’ heads. They treat the apes as (complicated) information processing devices and come up with what Marr 1982 would have called a ‘computational theory’ of their language. Marr points out that a proper understanding of any information processing device necessarily includes

a computational theory. He give the example of a cash register. The computational theory of a cash register is simply provided by the mathematical function of addition. This is true, despite the fact that cash registers do not know that they are computing addition or how they are doing it. Likewise, the biologists computational theory of the apes' communication system can be true or false completely independently of the apes' beliefs about it.

Answer to 6.4:

In (50a), [A] moves. Without A-movement, the English order DEM-NUM-A-N results. Movement of [A] alone violates the condition that only constituents containing the noun may move.

Without N-movement, the Aghem order N-A-DEM-NUM results. Without both movements, the Spanish order DEM-NUM-N-A results. The second step in (50b) violates the condition that movement must result in an order where N appears earlier in the string than it would have had movement not taken place ((49a-ii)).

Answer to 6.5:

Derivations with movement necessarily start with one of the structures in (39). The target order of the exercise is A-NUM-DEM-N. This is a structure with N in final position. Since movement always has the effect that N is pronounced earlier than it would otherwise be, only N-final structures in (39) are possible source structures. The only such structure corresponds to the English type of order: DEM-NUM-A-N. Moving any of the constituents containing N leftward ([N], [A N], [NUM A N], [DEM NUM A N]) results in an order where N is no longer final. Deriving an N-final order from there would require illicit rightward movement. Therefore, A-NUM-DEM-N cannot be derived.

Glossary

- adjective: Adjectives are words that can function as modifiers within the noun phrase, where, in English, they typically appear between the article and the noun. This is called their attributive use. It is illustrated by the word *blue* in *I bought a blue balloon*. Adjectives also have a predicative use, exemplified by *blue* in *The balloon is blue*. This chapter talks only about the attributive use and concentrates on descriptive adjectives, that is, adjectives that pick out observer-independent, permanent, physical properties of an object.
- c-command: In a syntactic tree, a given node c-commands its sister and all nodes dominated by its sister.
- constituent: A group of words that occur next to each other in a sentence and act as a unit. In syntactic tree diagrams, constituents are represented by subtrees rooted at a particular node.

daughter: In a syntactic tree, the daughters of a given node (if any) are those nodes that can be reached by following a single edge of the tree downward.

demonstrative: Demonstratives are elements of the noun phrase. They often encode number distinctions and deictic distinctions having to do with proximity. English has four demonstratives, cross-classified as singular versus plural and proximal versus distal:

	proximal	distal
singular	this	that
plural	these	those

dominate: In a syntactic tree, a given node dominates all nodes that can be reached by following a connected strictly downward path through the tree.

gloss: A word-by-word or morpheme-by-morpheme translation of foreign language examples. Glosses are usually written below the foreign language examples and aligned with them. Linguistics now usually follow the Leipzig glossing rules: www.eva.mpg.de/lingua/resources/glossing-rules.php

grammar: A grammar is a characterization of all the rules (syntactic, morphological, phonological, semantic, etc.) needed to characterize a given language together with the lexicon for that language. More narrowly, the syntactic rules that characterize the grammatical sentences of the language with their meanings are sometimes referred to as its grammar. A grammar that is explicit is called a generative grammar.

mother: In a syntactic tree, the mother of a given node is the single node (if any) connected to it by a single edge going up.

movement: Movement is a rule of grammar that displaces a constituent from the position where we would expect it to appear based on other rules of the grammar. Movement affects the hierarchical structure of the sentence. It is indicated within tree diagrams by arrows, and the launching site of movement is conventionally indicated by a trace or a copy of the moved constituent.

noun: Nouns are members of an open class of words whose members occur as the main word of subjects and objects of sentences and of objects of prepositions. The British logician Peter Geach proposed a cross-linguistically stable semantic definition of nouns which is based on the fact that adjectives like *same* can modify nouns, but no other kinds of parts of speech. Not only that, but there also do not seem to be any other expressions with similar meaning that can modify verbs and adjectives. Consider the following examples.

- (58) a. John and Bill participated in the same fight.
b. *John and Bill samely fought.

Geach proposed that nouns and only nouns are predicates with identity criteria. Even in languages where something like (58b) is grammatical, it cannot mean the same as (58a). Instead of meaning that Bill and John participated in the same fight it means that they fought in the same way. The property of being a predicate with identity criteria can be used as a cross-linguistic diagnostic for being a noun.

numeral: Many languages distinguish between cardinal numerals (*one, two, three, four, five...* and ordinal numerals (*first, second, third, fourth, fifth,...*—'). Cardinal numerals are used to describe quantities, ordinal numerals to pick out a position on an explicit or implicit list. The discussion in this chapter is restricted to cardinals.

principle: A property or property of rules that characterizes all human languages.

parameter: A point of restricted variation characterizing the differences between the grammars of different languages.

sister: In a syntactic tree, any nodes that share the same mother are sisters.

tree diagram: A syntactic tree diagram is a picture of the structure of a sentence representing both word order and hierarchical organization into constituents. Standard trees also have the following properties: No node in a tree has more than one mother. Every tree contains exactly one node without a mother, the root. No edges (branches) cross. The words of the sentence do not dominate any further nodes; they are the leaves of the tree.

universal grammar: A complete description of the linguistics principles and parameters together with a learning procedure, capable of mapping a set of sentences to the corresponding grammar.

Index

- Adjective
- Aghem
- Banda-Linda
- Basque
- Burmese

- C-command
- Chomsky, Noam
- Demonstrative
- Dominance
- Dulong
- English
- Gloss
- Greenberg, Joseph H
- Gungbe
- Kele
- Kĩtharaka
- Linguistic Universal
- Maasai
- Mother
- Movement
- Noun
- Noun Phrase
- Numeral
- Parameter
- Phrase Structure
- Pitjantjatjara
- Principle
- Sango
- Sister

- Sranan
- Syntax
- Tree diagram
- Universal Grammar

References

- Abels, Klaus and Ad Neeleman (2006). "Universal 20 without the LCA". In: *GLOW Newsletter* 54.
- Aboh, Enoch Olade (2004). *The Morphosyntax of Complement-Head Sequences: Clause Structure and Word Order Patterns in Kwa*. Oxford Studies in Comparative Syntax. Oxford; New York: Oxford University Press.
- Brown, Keith, ed. (2006). *The encyclopedia of language and linguistics*. Elsevier North-Holland.
- Cinque, Guglielmo (2005). "Deriving Greenberg's Universal 20 and its Exceptions". In: *Linguistic Inquiry* 36.3, pp. 315–332.
- Cloarec-Heiss, France (1986). *Dynamique et équilibre d'une syntaxe: le banda-linda de Centrafrique*. Descriptions de Langues et Monographies Ethnolinguistiques. Cambridge & Paris: Cambridge University Press & Éditions de la maison des sciences de l'homme.
- Dutton, Tom E (1996). *Koiari*. Vol. 10. Languages of the World-Materials. München & Newcastle: LINCOM Europa.
- Eckert, Paul and Joyce Hudson (1988). *Wangka Wiru: A handbook for the Pitjantjatjara language learner*. Underdale: South Australian College of Advanced Education.
- Elsabbagh, M and Annette Karmiloff-Smith (2006). "Modularity of Mind and Language". In: *The encyclopedia of language and linguistics*. Ed. by Keith Brown. Elsevier North-Holland, pp. 218–224.
- Greenberg, Joseph (1963). "Some universals of grammar with particular reference to the order of meaningful elements". In: *Universals of Language*. Ed. by Joseph Greenberg. Cambridge, MA: MIT Press, pp. 73–113.
- Huang, C.-T. James (1982). "Logical relations in Chinese and the theory of grammar". Doctoral dissertation. MIT.
- Hurford, James R (1994). *Grammar: A Student's Guide*. Cambridge, UK: Cambridge University Press.
- Hyman, Larry (1979). "Aghem Grammatical Structure". In: *Southern California Occasional papers in Linguistics* 7.
- Jones, Robert B (1970). "Classifier constructions in Southeast Asia". In: *Journal of the American Oriental Society* 90.1, pp. 1–12.

- Koopman, Hilda (2003). *On the parallelism of DPs and clauses: Evidence from Kisongo Maasai*. Ms. UCLA.
- LaPolla, Randy J (2003). “Dulong”. In: *The Sinto-Tibetan Languages*. Ed. by Graham Thurgood and Randy J LaPolla. Routledge Language Family Series. Routledge, pp. 674–682.
- Larson, Richard K (2010). *Grammar as Science*. Cambridge, MA: MIT Press.
- Marr, David (1982). *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.
- McMahon, A and R McMahon (2006). “Genetics and Language”. In: *The encyclopedia of language and linguistics*. Ed. by Keith Brown. Elsevier North-Holland, pp. 21–24.
- Oyharçabal, Beñal (2012). “Word order in Basque determiner phrases”. In: *Noun Phrases and Nominalization in Basque*. Ed. by Urtzi Etxebarria, Ricardo Etxepare, and Myriam Uribe-Etxebarria. Vol. 187. *Linguistik Aktuell*. John Benjamins Publishing Company, pp. 267–282.
- Pinker, Steven (1994). *The Language Instinct*. New York: Morrow.
- Ross, Malcolm (2002). “Kele”. In: *The Oceanic Languages*. Ed. by John Lynch, Malcolm Ross, and Terry Crowley. Richmond (Surrey): Curzon Press, pp. 123–147.
- Samarin, William J (1967). *A Grammar of Sango*. *Janua Linguarum*. The Hague, Paris: Mouton & Co.
- Thornell, Christina (1997). *The Sango language and its lexicon*. *Travaux de l’Institut de Linguistique de Lund* 32. Lund: Lund University Press.
- Thurgood, Graham and Randy J LaPolla, eds. (2003). *The Sinto-Tibetan Languages*. Routledge Language Family Series. Routledge.
- Voorhoeve, Jan (1962). *Sranan Syntax*. Amsterdam: North-Holland.
- Wiersma, Grace (2003). “Yunnan Bai”. In: *The Sinto-Tibetan Languages*. Ed. by Graham Thurgood and Randy J LaPolla. Routledge Language Family Series. Routledge, pp. 651–673.

Biographical note

Klaus Abels received his Ph.D. in linguistics in 2003 from the University of Connecticut. He has held temporary appointments in Leipzig (Germany) and Tromsø (Norway) and has been at University College London since 2008, where he works as a reader in syntax. Klaus Abels is co-editor of the journal *Syntax*.