1 Introduction

Much has been written on what constitutes a clause cross-linguistically. In English, main clauses minimally contain a finite verb and a subject, whereas subordinate clauses may lack an independent subject or a finite verb. Before discussing the descriptive characteristics of the clause in Chapter 2, I first present a brief overview of Generative Grammar and some of the main issues of the moment: Universal Grammar, Principles and Parameters, and the status of phrase structure rules.

Chomsky (1965 and later) emphasizes descriptive and explanatory adequacy. By that, he means that our models need to describe what the grammar of a language is and how children learn it. In recent years, he has emphasized a third question, namely why the grammar is the way it is. This question is also known as going "beyond explanatory adequacy" (Chomsky 2004), and I will discuss it in Section 1.1 of this chapter, though it is not the focus of this book.

The chapter is organized as follows. In Section 1.1, the focus is on Universal Grammar, the early instantiations of phrase structure rules, and the Principles and Parameters approach that is still the basis for Minimalism. In Section 1.2, I sketch the changes from phrase structure rules through X-bar theory to bare phrase structure and to the current problematization of projection. I also outline the changes from transformations to Move-alpha to Agree and features. In Section 1.3, I introduce Cartography and, in Section 1.4, we look at word order issues. In Section 1.5, we return to what might be feature parameters of the clause.

1.1 GENERATIVE GRAMMAR AND UNIVERSAL GRAMMAR

In this section, we first examine Universal Grammar and how Universal Grammar interacts with Phrase Structure Rules, then the Principles and Parameters approach, and finally some recent issues.

1.1.1 Universal grammar

The first half of the twentieth century is, among other things, known for its precise descriptions of languages. Leonard Bloomfield and Zellig Harris are among the most influential linguists using a descriptive, structuralist approach. Psychology and learning theory are at that time dominated by behaviorism. Noam Chomsky and the generative model he develops offer an alternative, starting in the late 1950s, to descriptive linguistics and behaviorism and bring about a revolution in the fields of linguistics, psychology, and cognitive science. Chomsky continues to stress descriptive adequacy but adds explanatory adequacy to emphasize the interest in how the language faculty is represented in humans.

In the new model, the focus is on the mind of the language learner/user (the **competence**) and ceases to be on the structures present in the language produced (the **performance**). The input to language learning is seen as poor, a phenomenon known as the "poverty of the stimulus". The basis for this phenomenon is that speakers know so much more than what they have evidence for from the input. For instance, speakers of English have never been taught that sentences of the type in (1a) are grammatical but those in (1b) are not. Yet, they can judge this difference in grammaticality.

(1) a. Who did I hear **that** John met?
b. *Who did I hear **when** John met?

In (1a), *who* originates as the object of the verb *met* and is fronted to form the *wh*-question; in (1b), the same is true, but somehow changing *that* to *when* makes the sentence ungrammatical. We'll talk about *wh*-movement more in Chapter 2, where relative clauses are concerned, and more generally in Chapter 5.

How is acquisition of, for instance, (1) possible? We may never have heard (1a) and still know that it is grammatical, and we certainly haven't been taught that (1b) is ungrammatical. The answer to this problem of impoverished input, **Plato's problem** in Chomsky (1986a), is Universal Grammar, the initial state of the language faculty. This biologically innate organ helps the learner make sense of linguistic data and build an internal grammar (I-language), which then produces the sentences a speaker utters (E-language). See Figure 1.1.

The innate language faculty, when "stimulated by appropriate and continuing experience ... creates a grammar that creates sentences with formal and semantic properties," according to Chomsky (1975: 36). Thus, our innate language faculty (or Universal Grammar) enables

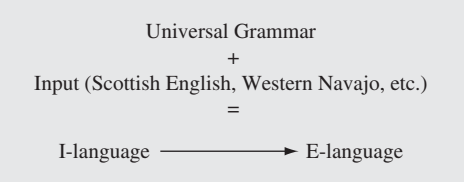


Figure 1.1 Model of language acquisition

us to create a set of rules, or grammar, by being exposed to (rather chaotic) language around us. The set of rules that we acquire enables us to produce sentences that we have never heard before. These sentences can also be infinitely long (if we have the time and energy).

Language acquisition, in this framework, is not imitation but an interaction between Universal Grammar and exposure to a particular language. "Learning is primarily a matter of filling in detail within a structure that is innate" (Chomsky 1975: 39). "A physical organ, say the heart, may vary from one person to the next in size or strength, but its basic structure and its function within human physiology are common to the species. Analogously, two individuals in the same speech community may acquire grammars that differ somewhat in scale and subtlety . . . These variations in structure are limited." (1975: 38).

As I mentioned, Chomsky, in various publications, has talked about descriptive and explanatory adequacy. With the first, he means that we have to be able to **describe** the grammar of a particular language; with the second, he means that we have to be able to **explain** how the child acquires its language. In the last decade, he has added that we have to look at why language is the way it is. This is also known as "Beyond Explanatory Adequacy," see e.g. Chomsky (2004). In the remainder of the book, I will not focus on the "why" question, but I will return to it in Chapter 7.

Having briefly introduced Universal grammar, I'll provide a brief overview of how Universal Grammar deals with clause structure.

1.1.2 Phrase Structure and transformations

The implementation of the model in Figure 1.1 has seen a number of major changes. Between the 1950s and 1970s, a generative transformational model was developed that uses recursive phrase structure rules (to derive the deep-structure) and transformations (to derive the surface structure). These rules use substantive and formal universals. The substantive universals concern universal categories (V, N, etc.)

and phonological features; the formal universals relate to the nature of rules.

Both phrase structure and transformational rules of the internalized grammar of the 1960s are very language-specific. For instance, two Phrase Structure rules for English are given in (2) and for Persian in (3). The Verb and NP are in reverse order in (2b) and (3b), since objects follow verbs in English but precede them in Persian.

Phrase structure rules:

(2) a. Sentence → NP + VP
b. VP → Verb + NP (Chomsky 1957: 27)

(3) a. Sentence → NP + VP
b. VP → NP + Verb

An example of a transformation is given in (4). This rule derives an English passive from an active sentence. It is again very language-specific.

Transformational rules:

(4) If S1 is a grammatical sentence of the form
NP1 - Aux - V - NP2,
then the corresponding string of the form
NP2 - Aux + be + en - V - by + NP1
is also a grammatical sentence. (Chomsky 1957: 43)

Universals in this system of deriving the clause structure are the categories used in the phrase structure rules and the way the rules are written. (Categories are currently no longer part of Universal Grammar, and the term "universal" is no longer used in this framework except in the phrase "Universal Grammar".)

The input and output of the Phrase Structure Rules are not constrained, and one could have a rule like that in (5), for instance. Such lack of restrictions is not desirable.

(5) VP → S PP N V PP

Therefore, to remedy this lack of constraints, the system in (2) to (4) becomes increasingly generalized and abstract, starting in the 1970s. Insights into phrase structure from Chomsky (1970) and Jackendoff (1977) replace such rules as (2) and (3) with X'-theory, again applicable cross-linguistically. We'll see more on X-bar in the next section. And after Ross' (1967) work discovers **islands**, domains from which movement cannot take place, such rules as (4) are replaced by "move

alpha" (= move anything anywhere). Such rules are applicable in any language. The consequence of all this is a system that invites cross-linguistic comparison, and, as a result, Universal Grammar comes to be seen in the late 1970s and early 1980s as consisting of Principles (true in all languages) and Parameters (choices to be made depending on the language).

1.1.3 Principles and Parameters

Principles can be pretty model-specific. I will list a few without elaborating on them further. They include the Subjacency Principle (Chomsky 1973), the Structure Preserving Hypothesis (Emonds 1976), the Head Movement Constraint (Travis 1984), Relativized Minimality (Rizzi 1990), the Phase Impenetrability Condition (Chomsky 2001), the Inclusiveness Condition (Chomsky 1995: 225–228), the Extension Condition (Chomsky 1995), No Tampering (Chomsky 2004: 117; 2008b: 138), the Head Preference Principle (van Gelderen 2004), and many others.

Many of the principles restrict how movement is constrained. For instance, Subjacency limits movement to crossing no more than two phrases of a particular kind; the Structure Preserving Hypothesis states that transformations, i.e. movement, can only move elements to positions that could be generated by means of phrase structure rules; the Head Movement Constraint says that heads only move to head positions; Relativized Minimality claims that heads are related to other local heads and phrases to other similar local phrases; and the Extension Condition requires that syntactic operations extend the tree at the root.

Some of these principles have been abandoned, e.g. Subjacency and Structure Preservation, although you will still see references to them in the literature. You will also see mention of some of the others in this book, e.g. head movement and Relativized Minimality, but not all will play a role. Principles continue to be valid to the present, although their location and level of specificity are debated. At the moment, the emphasis is on principles not specific to the language faculty (Universal Grammar), but to "general properties of organic systems" (Chomsky 2004: 105), "third factor principles" in Chomsky (2005; 2007). Thus, Chomsky identifies three factors crucial in the development of language.

Three factors

> (1) [G]enetic endowment, which sets limits on the attainable languages, thereby making language acquisition possible; (2) external data,

> converted to the experience that selects one or another language within a narrow range; (3) principles not specific to FL [the Faculty of Language]. Some of the third factor principles have the flavor of the constraints that enter into all facets of growth and evolution . . . Among these are principles of efficient computation (Chomsky 2007: 3)

The first factor is the traditional Universal Grammar with Principles and Parameters, and the second factor is the experience that we saw in Figure 1.1. The third factor marks a new emphasis but is somewhat related to the first factor. The third factor is favored above the language-specific first one (for reasons of simplicity) and can be divided into several types, including principles of efficient computation, which are "of particular significance in determining the nature of attainable languages" (Chomsky 2005: 6). Economy Principles are probably also part of more general cognitive principles, thus reducing the role of Universal Grammar even more.

Early examples of **parameters**, determined by the first factor, are pro-drop (Rizzi 1982), headedness (Stowell 1981), and movement of *wh*-elements (Huang 1982). **Pro-drop** is the cover term for a set of related phenomena, and there are many ways to account for having empty subjects. Not many linguists, however, believe that the phenomenon involves a +/– setting of an actual parameter called 'pro-drop'. Pro-drop is a collection of related phenomena: the absence of the subject of a finite verb, as in (6), subject-inversion, long *wh*-movement of the subject, etc. (see Chomsky 1981: 240).

(6) *tiene* *un* *bolígrafo* =pro-drop Spanish
have.3S a pen
'S/he has a pen.'

Headedness is a way to characterize a language, with Arabic and Irish being head-initial, and Japanese and Korean head-final. Following work by Kayne (1994), however, headedness has been abandoned as a formal parameter. In this framework, the basic word order is SVO, and other word orders come about through movement. I come back to issues of basic word order in Section 1.4. Likewise, the ***wh*-movement parameter** is now often seen as dependent on the character of the C or on an Extended Projection Principle (EPP) feature.[1]

[1] Setting the Binding Domain (Chomsky 1981: 225, fn 35) and finding the relevant barriers for Subjacency (Chomsky 1973) are two other early parameters. They are now part of a theory of phases or of an Economy Locality Principle.

Though most introductory generative syntax books continue to cite this set of three parameters, pro-drop/null subject, headedness, and *wh*-movement, these are often used in very descriptive ways to describe the typological characteristics of a language, not to explain what goes on in language acquisition. Since Chomsky (1995), a major question is how these parameters would have arisen in the brain. If the shift in humans from no language to language was immediate, it makes sense that there is one crucial change in the way the brain functions, and that change could have been the introduction of Merge. Complex parameters of the pro-drop variety don't fit in this non-gradual picture of evolution.

In addition, especially since Borer (1984), parameters consist of choices of feature specifications as the child acquires a lexicon (Chomsky 2004; 2007). The computational system of every language is seen as the same. Thus, all parameters are lexical, and they account for the variety of languages. If the child has evidence for gender in the language it hears, gender will be included; if not, it won't be. Baker, while disagreeing with this view of parameters, calls it the Borer-Chomsky-Conjecture.

(7) Borer-Chomsky-Conjecture

> All parameters of variation are attributable to differences in the features of particular items (e.g., the functional heads) in the lexicon (Baker 2008a: 156; 2008b: 3).

There have to of course be restrictions as to how much freedom the child has in selecting or ignoring the features and on how to bundle them. I will come back to this in Section 1.5.

The next section will examine changes in Generative Grammar in more detail and show how the move to (7) is possible.

1.2 FROM PHRASE STRUCTURE TO MINIMALIST FEATURES

In this section, I outline the system of the early 1980s (the Government and Binding framework, so called after the title of Chomsky's 1981 book) and subsequent changes up to and including Chomsky's Minimalist Program. In Section 1.2.1, I emphasize the changes in phrase structure rules, because this is crucial to the current Minimalist Program. In Section 1.2.2, we look at transformations

and how their character changes due to features. Section 1.2.3 finishes with a typical Minimalist derivation.

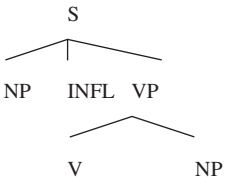
1.2.1 Phrase Structure and X-bar

The phrase structure rules of Chomsky (1981) do not differ crucially in form from those of Chomsky (1957). Thus, (2) above and (8) are quite similar.

(8) a. VP → V NP
b. S → NP INFL VP (Chomsky 1981: 25)

A tree using (8) would look as in (9).

(9)



Though the actual Phrase Structure Rules aren't very different, the rationale behind (8) had become more principled, more universally applicable. With Chomsky's (1970) paper "Remarks on Nominalization", X-bar theory had been introduced, which, at least for the lexical categories, did wonders to make the system less language-specific. Jackendoff (1977) refines this system even more. Crucial to X-bar theory is that all phrases look alike, with a head, a complement, a specifier, and possibly an adjunct. Rather than having separate rules for NP, VP, PP, etc., (10) generates rules for any lexical category.

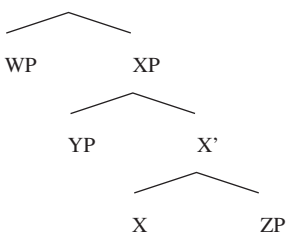
(10) a. XP → YP X' (YP = Specifier)
b. X' → X ZP (ZP = Complement)
c. XP → WP XP (WP = Adjunct)
(where X, W, Y, and Z stand for N, V, A, and P)

Chomsky had also worried about the redundancy between the lexicon and very specific phrase structure rules, such as (11), to accommodate such verbs as *give*. If *give* is in the lexicon with three arguments, that and such rules as (10) should be enough to project any space needed. The redundancy is then eliminated.

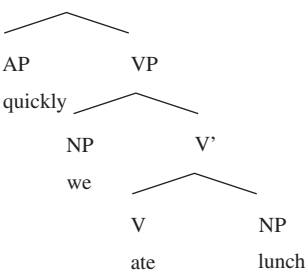
(11) VP → V NP (NP) (PP)

In tree form, (10) looks like (12a), where the X projects up and is the only head category. In (12b), an actual VP (X=V) is provided, although the subject doesn't appear as the specifier of the VP until the mid 1980s.

(12) a.

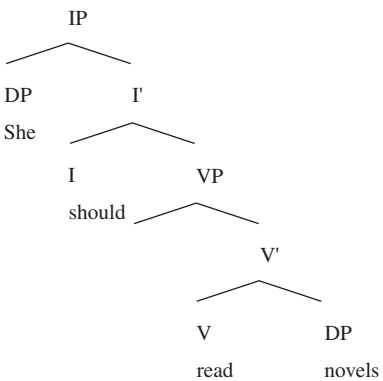


b.



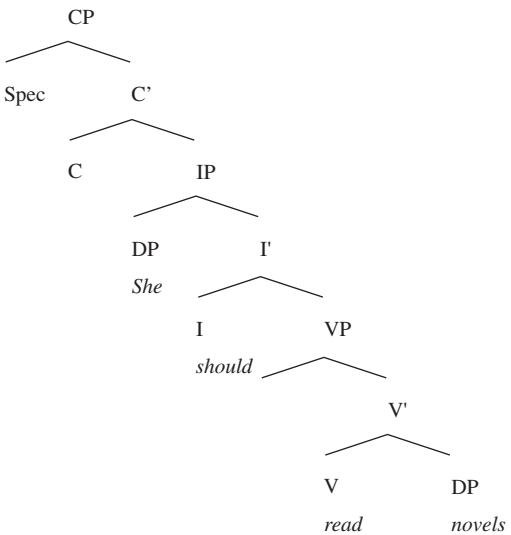
In the mid 1980s, work by den Besten (1983), Fukui and Speas (1986), Abney (1987), and others leads to a change where functional categories, C, INFL, and D are considered on a par with lexical categories and head their own projections. Taking the X-bar structure of functional categories into account as well, (8) is reformulated as (13), with INFL changed into I(nflection) and later into T(ense), and the NP argument always as part of a DP with a D head.

(13)



This structure has been the basic (non-expanded) structure of the clause since 1986, although most syntacticians assume a CP above the IP, as in (14), even if the clause does not include a *wh*-pronoun (typically in the Specifier of CP) or a complementizer (in the head of the CP).

(14)



Pollock (1989) suggests splitting up I into T and AGR(reement), known as the "Split Infl Hypothesis," and Chomsky (1989) ends up with an AGRs and an AGRo. English doesn't show much agreement, so we'll use Spanish to exemplify these two categories. AGRs is used for agreement with the subject, *-án* in (15), T for the future *-r*, and AGRo for the feminine plural *-das* on the V. (I have placed the passive auxiliary in T, but that has moved from a lower position.) The tree structure of (15) is given in (16).

(15) *Las casas* *se-r-**án*** *vendi-**das*** *(el mes próximo)* Spanish
the houses(F) be-FUT-3P sold-FP the month next
'The houses will be sold (next month).'