

# Principles, Parameters and Universal Grammar

The principles and parameters approach to syntax proposes that there is a set of universal principles shared by every human language, and that these are known by all human beings. Knowledge of a particular language, then, consists of knowledge of the settings of a finite number of parameters which define exactly how the universal principles need to be applied to construct grammatical sentences. If the parameters according to which languages may vary could all be found, then a given human language could be completely described by the values it assigns to each parameter; it would be *the* (only) human language with the parameters set in that way.

An interesting application which provides motivation for this theory of syntax is the idea of a computer system which uses it to perform translation between human languages. The "core" of the system is the knowledge of the universal principles. To this are added a number of "modules", one for each language the system is to work with. Each module includes, firstly, the lexical information for that language. (Admittedly this is a large volume of data, but it is inevitable; no universal theory will ever be able to derive the fact that the English word for hit is *hit*, that this verb's agent needs to be assigned nominative case and that its patient needs to be assigned accusative case, for example. Information like this must simply be individually encoded.) Secondly, each module includes the settings for each of the parameters of the core universal principles. Thus, combining the core universal principles with the data in the appropriate module, the system has a complete description of a given human language.

Suppose that a language-independent, abstract representation of the meaning of any sentence (in any language) could be agreed upon (probably making significant use of semantic roles, but the details of this structure are not the concern here), which would act as an interlingua. Given a sentence (sequence of words) in the source language (SL), it is possible that the system could construct the corresponding interlingua representation of its meaning, using the core principles and the module corresponding to SL; during this computation, the target language (TL) is completely irrelevant. Similarly, once the interlingua representation has been determined, it is possible that the system could construct the corresponding sentence in TL, using the relevant module; in this process, SL is completely irrelevant.

Such a system would be very scalable. The use of an independent interlingua would mean that having  $n$  language modules allows a number of SL-TL pairs in the order of  $n$ -squared; the use of the principles and parameters framework, to determine what needs to be stored in the language modules to describe the language's syntax, would ensure that only the minimum of language-specific information is stored in these modules, and that as much as possible is factored out and stored in the universal principles at the system's core. A minimum of information in the language modules means a minimum of work in adding a previously-unknown language to the system. For example, consider such a system which currently has 20 language modules provided; it can therefore perform translations involving 380 SL-TL pairs (20 possible SLs multiplied by 19 possible TLs). Simply adding a twenty-first language module would immediately increase this figure to 420 (21 times 20); in order to add this module, no knowledge is needed of the other 20 languages already present, and no

information will have to be entered which relates to human language in general rather than the language being considered. In other words, the work involved in adding this module would be kept to entering the bare minimum of information which is genuinely specific to the language. (The principles and parameters framework will in theory allow the data describing the language's syntax to be relatively small; of course the amount of lexical information required will be very significant, but as mentioned above this is inevitable.)

This essay aims to use Chomsky's theory of universal grammar to begin to determine what information can be stored in the system's core as universal principles, what parameters need to be given values in each added language module, and in what form the lexical information of each language module needs to be stored. Thus, clearly, the structure of such a system would closely parallel what Chomsky proposes as the structure of language knowledge in the human mind.

A feature found in all human languages is that there are structural relationships between components of a sentence, not only the linear relationships which are apparent from the sentence's spoken or written realisation as a sequence of sounds or words. For evidence of this, consider the following English statement and its corresponding question:

- (1) John will arrive tomorrow.
- (2) Will John arrive tomorrow?

Based on these two items, one could conclude that questions are formed in English by moving the second word of the corresponding statement to the front. However this can easily be disproven:

- (3) The letter will arrive tomorrow.
- (4) \*Letter the will arrive tomorrow?

Here the proposed rule does not work. The correct question form is:

- (5) Will the letter arrive tomorrow?

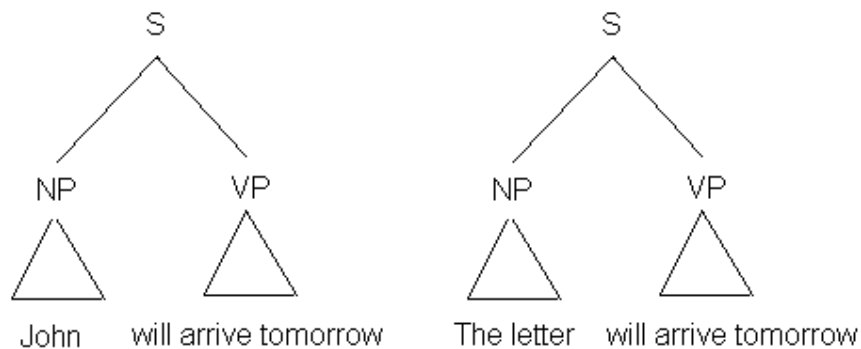
This involves moving the third word to the front. If speakers of English considered sentences *only* as sequences of words, how would they know that to form the correct question in (2) the second word must be moved, but to form the correct question in (5) the third word must be moved? The answer is that knowledge of the language depends on knowledge of a deeper internal structure to the sentence. In this case, English speakers know (at some level) that "John" in (1) and "The letter" in (3) are corresponding constituents, namely noun phrases, which allows them to correctly construct both questions. This is known as structure-dependency.

Analysing data from various human languages similarly shows that in fact structure-dependency is a feature of all languages. While it would seem very strange to human beings, it is perfectly conceivable that a language could, for example, always form questions by moving the second word of the corresponding statement to the front; but no such language exists. Structure-dependency, therefore, can be considered a property of human language in general, one of the general principles which apply

across all languages. (In terms of the translation system considered above, the fact that all languages are structure-dependent can be stored once in the core, and there is no need to specify this when describing a new language to be added to the system.)

In order to represent sentences, therefore, a kind of notation is needed which shows this hidden structure, not just the linear one evident in normal text. For this purpose, trees such as the following are used (triangles are used to simplify details which are not relevant for now):

(6)



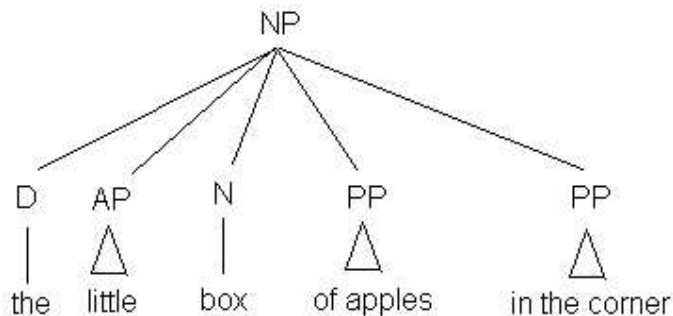
Now that the sentences are represented this way, it is clear that the movement required to form the two corresponding questions was in some sense the same transformation.

To construct trees like this, traditional parts of speech (noun, verb, adjective, preposition) and phrase types (NP, VP, AP, PP) were adopted as the constituents. This led to rules such as:

(7) NP → (D) (AP+) N (PP+)

and trees such as:

(8)

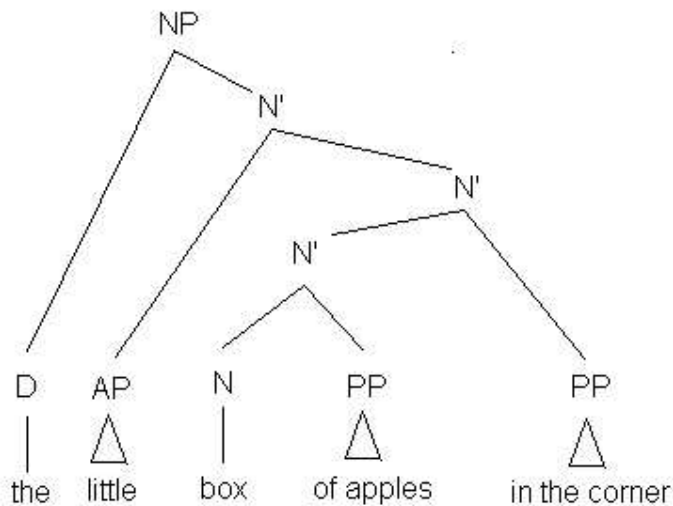


However it can be shown that, like the model of a sentence as simply a sequence of words, these structures are too "flat"; there is more hidden inside them. For example, the following sentence indicates that "box of apples" is in fact a constituent since it is replaced by "one", but it does not appear so in (8):

- (9) I saw the little box of apples in the corner, but not the big one on the shelf.

Thorough analysis of this noun phrase reveals that its structure is in fact as follows:

(10)



(The intermediate nodes, which did not exist before, have been labelled N' for lack of a better name.)

This tree leads to the following set of more articulated rules for describing the internal structure of NPs:

- (11) NP → (D) N'  
 (12) N' → (AP) N' or N' (PP)  
 (13) N' → N (PP)

Similar reasoning produces the following rules for English VPs, APs and PPs:

- (14) VP → V'  
 (15) V' → V' (PP)  
 (16) V' → V (NP)  
 (17) AP → A'  
 (18) A' → (AP) A'  
 (19) A' → A (PP)  
 (20) PP → P'  
 (21) P' → P' (PP)  
 (22) P' → P (NP)

Based on these 12 rules, there are significant similarities between the internal structures of NPs, VPs, APs and PPs. (Admittedly (14), (17) and (20) do seem somewhat meaningless at this stage, presented merely to create similarities, but they will have a use later.) X-bar theory is an attempt to capture the similarities between these rules. Using W, X, Y and Z as variables for the four categories N, V, A and P, rules (11)-(22) can be unified as follows:

- (23)  $XP \rightarrow (YP) X'$
- (24)  $X' \rightarrow (ZP) X' \text{ or } X' (ZP)$
- (25)  $X' \rightarrow X (WP)$

(The only anomaly is that the D in (11) is not a phrase, and therefore cannot be represented by YP. This is motivated by the fact that, in all other rules, the non-head component is phrasal; resolution of this problem is discussed below.)

The YP in (23), if present, is known as a specifier; the ZP in (24), if present, is known as an adjunct; and the WP in (25), if present, is known as a complement. The X in (25) is known as the head of the XP. Applying these general terms to the tree in (10), for example, shows that in that NP, the head is the N "box"; the PP "in the corner" and the AP "little" are adjuncts; and the D "the" is a specifier.

All the data used to derive these rules, however, were from English. They can not generate, for example, the following Japanese sentence:

- (26) Taro ga inu o mita.  
Taro SUB dog OBJ saw.  
Taro saw the dog.  
(data from Whaley, p.81)

This should include a VP "inu o mita". The complement of this VP is the NP "inu o", but this appears to the left of the head V, "mita". This VP could never be generated by the rule in (25).

So not all languages place their heads to the left of the complements, as English does; similarly not all languages place their specifiers to the left of the heads. However we want X-bar rules that apply to any language. The solution is to set these orderings to be one of the parameters of universal grammar which characterise each human language. The universal X-bar rules become:

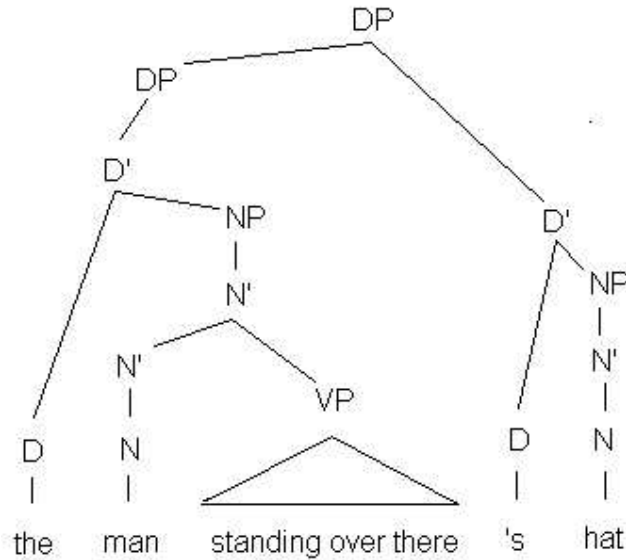
- (27)  $XP \rightarrow (YP) X' \text{ or } XP \rightarrow X' (YP)$  (Specifier Rule)
- (28)  $X' \rightarrow (ZP) X' \text{ or } X' \rightarrow X' (ZP)$  (Adjunct Rule)
- (29)  $X' \rightarrow (WP) X \text{ or } X' \rightarrow X (WP)$  (Complement Rule)

English (along with many other languages) has its Complement Rule parameter set to the first alternative in (29), such that the complement is on the right; Japanese has its Complement Rule parameter set such that the complement is on the left.

A problem with the theory so far, as mentioned above, is the fact that the D in rule (11) is not phrasal. This does not fit with the other rules, where all the non-head components are phrasal, and at present therefore conflicts with the universal Specifier

Rule. To fix this, the determiner can be considered as the head of a DP, of which the NP is the complement. As well as being an elegant theoretical solution, there is some evidence for this construction, for example the following tree:

(30)



(example from Carnie, p.145)

Without the idea of a DP, it would not be possible to construct a tree for this expression.

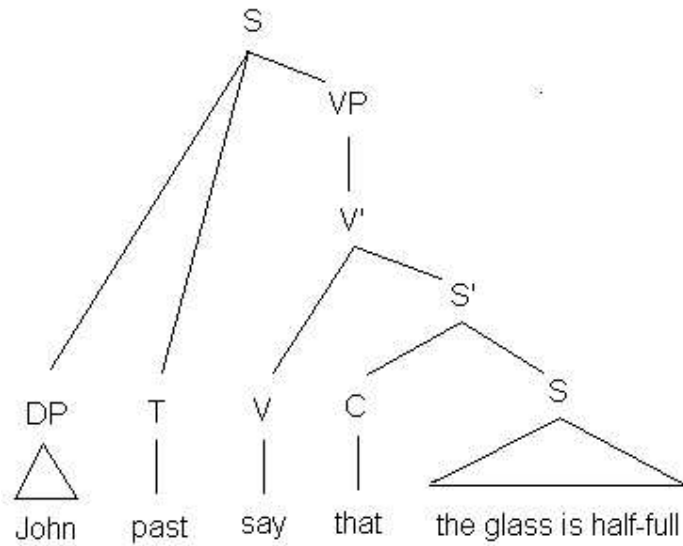
Now X-bar theory accounts quite well for the internals of NPs, VP, APs, PPs and DPs. These come together to form complete, meaningful sentences. The following basic rules can be proposed based on traditional English grammar to describe the composition of clauses:

(31)  $S' \rightarrow (C) S$

(32)  $S \rightarrow DP (T) VP$

Here C stands for complementiser, and T roughly for any tense inflections and auxiliary verbs. Here is an example of their use:

(33)

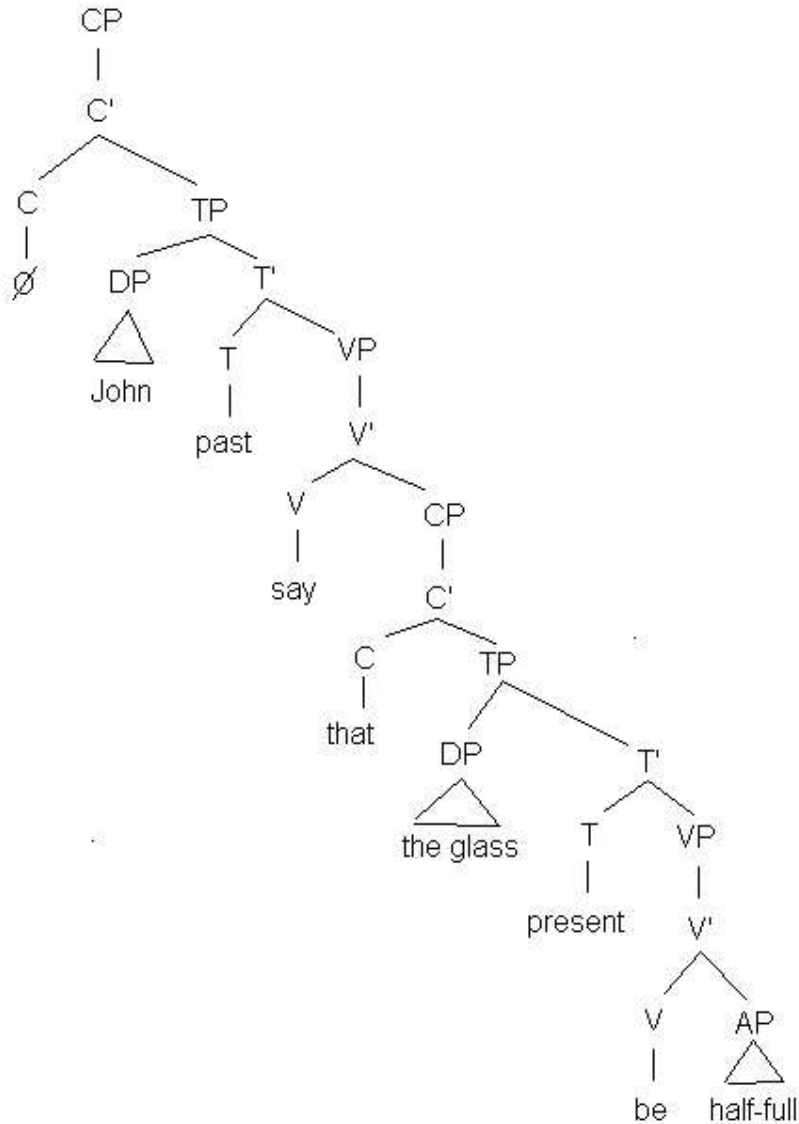


However these rules also do not fit with X-bar theory. In order to unify (31) with the theory, we can rename  $S'$  to  $CP$ , considering  $C$  as its head and  $S$  as its complement (and an empty specifier position). Similarly, to fix (32), we can rename  $S$  to  $TP$ , considering  $T$  as its head,  $VP$  as its complement, and  $DP$  as its specifier.

- (34)  $CP \rightarrow C'$
- (35)  $C' \rightarrow C TP$
- (36)  $TP \rightarrow DP T'$
- (37)  $T' \rightarrow T VP$

There is evidence for this structure, but for now these rearrangements can be considered as purely serving the purpose of unifying (31) and (32) with the theory. Here is the tree for the same sentence as in (33), drawn using these new rules:

(38)



Now X-bar theory provides a set of rules unifying the common properties of not only NPs, VPs, APs, PPs and DPs, but also the higher-level CPs and TPs, roughly corresponding to clauses. The common rules, combined with knowledge of the ordering parameters, describe the sentences which can be generated in a particular language. However, unfortunately they "overgenerate"; there exist sentences which satisfy the rules given so far but are not grammatical:

(39) Jane gave the cake to Mary.

(40) \*Jane ate the cake to Mary.

Clearly these two sentences have identically-constructed trees, yet one is grammatical and one is not. Something more than X-bar theory must be involved in determining the grammaticality of a sentence. The extra restriction comes from the lexicon; the

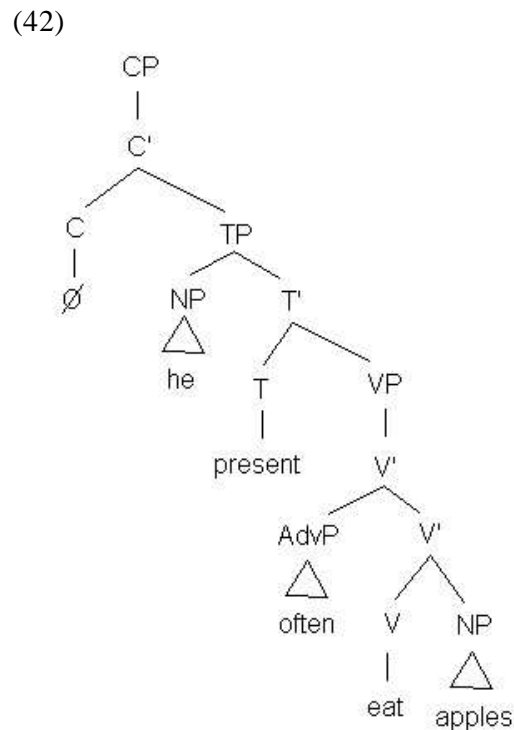


lexical entry for every verb specifies the number and types of arguments that can go with it. The verb "give" can take three arguments, two NPs ("Jane" and "the cake") and a PP ("to Mary"), so (39) is grammatical; the verb "eat" cannot take this combination, so (40) is ungrammatical. Thus a sentence produced according to the X-bar theory rules may still be "filtered out" as ungrammatical if the arguments do not match those accepted by the verb; if they do match, the sentence will be allowed. This criterion is known as the theta-criterion; the output from this combined process (X-bar theory and the theta-criterion) is known as a d-structure.

These d-structures are still not equivalent to grammatical sentences though. Observe that when all the details of the verb "say" are shown on the tree in (33), the correct form "said" is still not produced. The V "say" and the T "past" must combine to form this inflected verb. In (33), it is not difficult to predict the word order of the resulting sentence, because these two components appear together:

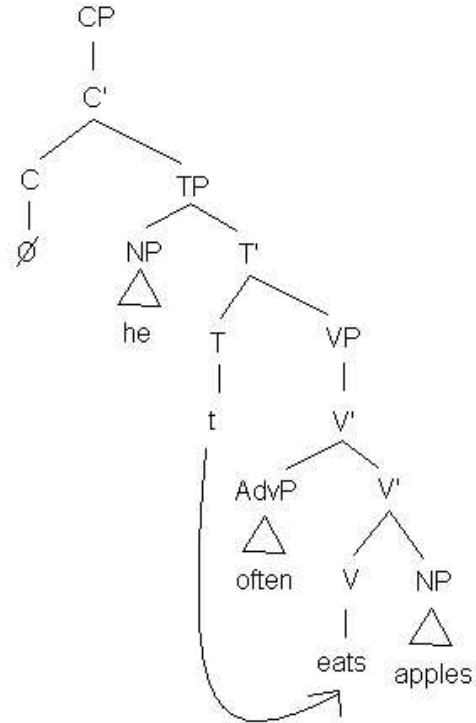
(41) John said that the glass is half-full.

However, consider the following tree:



Here, the AdvP "often" comes between the verb and the inflectional information it needs to combine with. The language must decide how this is going to be resolved, since it is obliged to produce a linear sequence of words. English decides to move the T to join the V, producing the word order we would expect, as shown in the following tree:

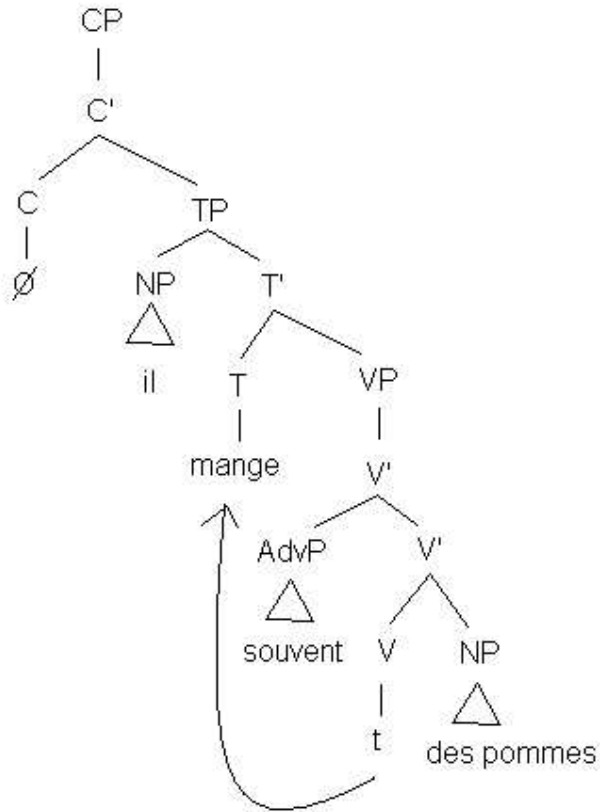
(43)



(The  $t$  indicates a "trace", which is left behind in a position out of which something has moved, and has no phonetic value.)

Other languages deal with this differently. French, for example, when confronted with a tree equivalent to that in (42), decides to move the V to join the T, producing a different word order from English:

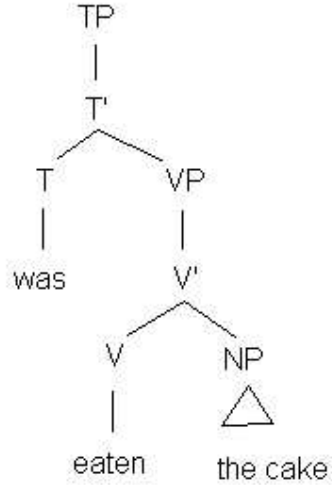
(44)



This choice is another of the parameters of languages, known as the "verb raising parameter"; a language is said to either raise the verb to T (as French does) or lower T to the verb (as English does).

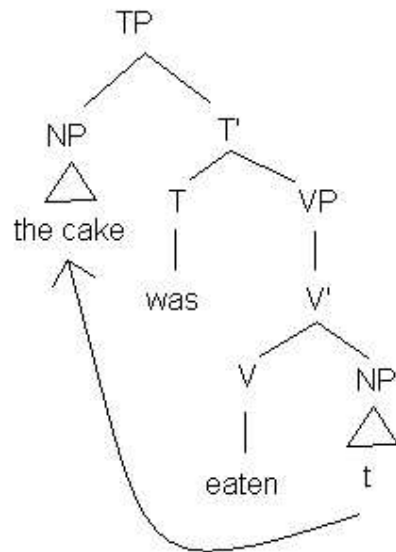
Movement of elements of the d-structure like this is often necessary to produce grammatical sentences, and is the source of many of the proposed parameters of languages. Another class of d-structures where it is necessary in English is that of passives. The following is a tree representing a passive:

(45)



The NP "the cake" must appear in this position of the tree for it to be accepted by the theta-criterion, because of a condition known as the locality condition. The correct surface word order has not yet been achieved though, so some movement must occur. To motivate this transformation, the EPP (Extended Projection Principle), which states that every clause must have a subject, can be used. In addition, Chomsky specifies a requirement that for an NP/DP to surface it must be given Case, and NP/DPs can only get Case in certain positions in the tree. To satisfy the EPP, the NP "the cake" (the only NP/DP in the tree) must move to the position where it can get nominative case; the positions where Case is assigned are further examples of parameters which are set for any language. In English, this parameter is set such that the nominative case is assigned to the specifier of a finite T, so the NP "the cake" must move there.

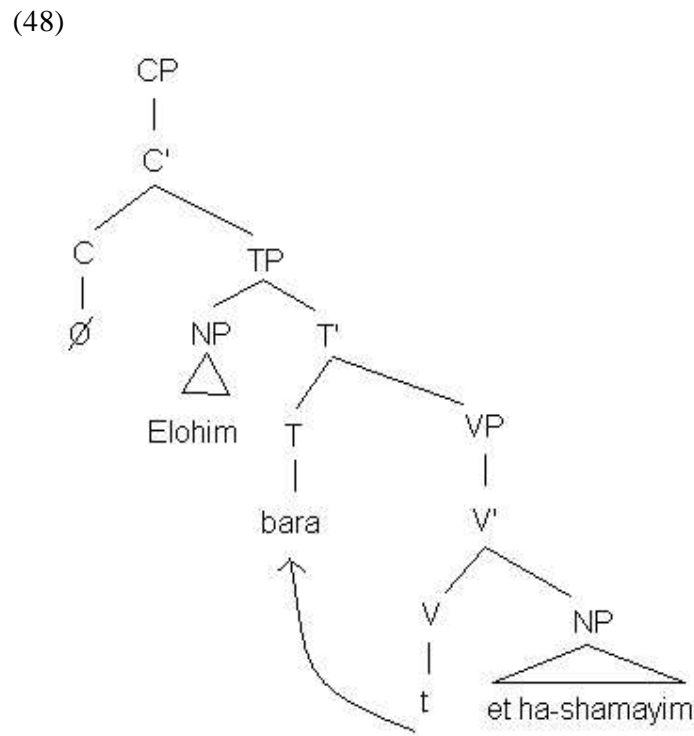
(46)



For another example of movement, consider the following sentence from Biblical Hebrew:

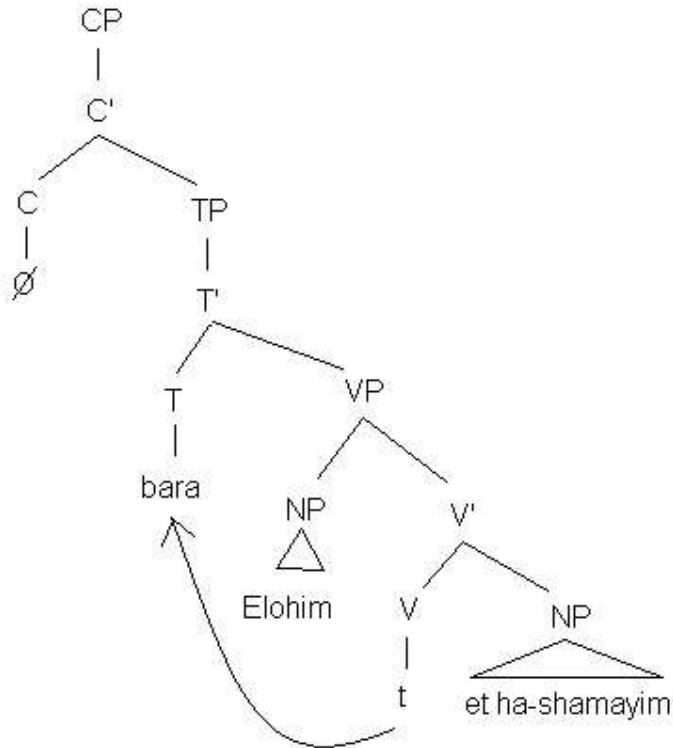
- (47) Bara Elohim et ha-shamayim.  
 created God OBJ ART-heavens  
 God created the heavens.  
 (data from Whaley, p.82)

This word order can never be generated from the X-bar rules because the verb and its object (its complement) are separated by the subject. No possible values of the ordering parameters will allow this, so some movement of constituents from their d-structure positions must take place. Firstly, Biblical Hebrew is a verb-raising language (indeed, without some movement to first separate the verb from its complement, it would be impossible for any movement to put the subject between them), so the following tree represents the tree after this verb movement has taken place:



Still, however, no possible modification of the X-bar ordering parameters will produce the surface VSO order. An alternative is simply that up until this point the subjects have been generated in the wrong position. The VP-internal subject hypothesis suggests that the subjects, rather than being specifiers of the TP as has been assumed, are in fact specifiers of the VP. This accounts perfectly for the VSO order:

(49)



As well as accounting for this word order, the hypothesis is supported by the behaviour of quantifiers (Carnie, p.241) and by the fact that it would allow subjects to be covered by the locality condition; it is therefore very appealing. If it is true, though, why does placing the subjects in the specifier positions of TPs produce the correct word order for other languages, such as English and French, as seen above? The solution is that the subjects in those languages were in fact generated in the VP-specifier position, and then moved to the TP-specifier position. A reason for an NP/DP in English to move to the TP-specifier position was mentioned above, and the same reasoning can cause the required movement here: the subject moves there to get Case. (In active and intransitives sentences it moves there from the specifier of VP, and in passives it moves there from the complement of V.) The reason the subject does not move to the TP-specifier position in Biblical Hebrew is that nominative case is assigned to a different position in the tree from English; this parameter is set to the VP-specifier position.

These recent examples have shown how the settings of just a small number of parameters can be used and interwoven to determine reasonably complex transformations which are applied to d-structures to create grammatical sentences. Many other transformations have been discovered which are similarly parameterised.

It should be noted, however, that stating that all languages have one fixed value for every parameter is a simplification. The correct setting which should be used for a particular construction can depend not only on the language, but on some property of the construction itself. For example, a well-known parameter of languages is which

case system they use: a nominative-accusative system, an ergative-absolutive system, or (very rarely) a tripartite system. German consistently uses a nominative-accusative system; Lakhota consistently uses an ergative-absolutive system. Yidin, however, uses a nominative-accusative system for pronouns and an ergative-absolutive system for other nominals; the decision of which system to use depends both on the language and on the NP which needs to be marked, or on two "dimensions". As another example, German uses SVO word ordering in main clauses but SOV ordering in subordinate clauses; at least one parameter must have a different setting for subordinate clauses to achieve this. The ordering parameter of the Adjunct Rule varies in English, seemingly according to the parts of speech involved. (The implication of this for the translation system considered earlier is simply that the core component must "show", for example, the Yidin module the NP which is to be marked before "asking" the module which case system to apply; the basis for the decision is language-specific and will rightly be stored in the module itself.)

Formalising these "two-dimensional" decisions within languages, and finding a set of parameters extensive enough to fully describe any human language, naturally remain goals of the principles and parameters approach. Considerable progress has already been made though, considering the enormous variation which at first appears to exist between human languages.

## **References:**

- Carnie, Andrew. 2002. *Syntax: A generative introduction*. Oxford: Blackwell.
- Chomsky, Noam. 1982. *Lectures on Government and Binding*. Dordrecht: Foris.
- Cook, V.J. 1988. *Chomsky's Universal Grammar: An introduction*. Oxford: Basil Blackwell.
- Dixon, R.M.W. 1977. *A grammar of Yidin*. New York: Cambridge University Press.
- Phillips, Colin. 2001. "Syntax" in *Encyclopedia of Cognitive Science*. Macmillan.
- Whaley, Lindsay J. 1997. *Introduction to Typology: The unity and diversity of language*. Thousand Oaks, California: Sage.