

Mapping the Terrain of Language Learning

Mark C. Baker

Rutgers University

1. Introduction: Learning and Typology

Language learning and language typology are often studied separately, and it is common for experts in one area to know rather little about the other. This is not merely an unfortunate historical coincidence; there are some powerful practical reasons why it is so. The detailed study of language learning typically involves the experimental investigation of groups of people who are at various stages in the learning process—i.e., children. Hence it prototypically takes place at university daycares in North America, where the children are usually learning English. In contrast, the study of typology is concerned with probing the full extent of the diversity that natural human languages can exhibit, and with finding and explaining any limits to that diversity that exist. As a result, it prototypically involves doing fieldwork with small numbers of fully competent adult speakers of less-studied languages. This fieldwork is made possible either by the researcher traveling to the often-remote areas of the world where these less-studied languages are spoken, or by finding speakers who have come nearer to the researcher through immigration. As a result, the contexts in which language acquisition can readily be studied and the contexts in which language typology can effectively be studied rarely overlap. Indeed, it is an unfortunate reality of the world as we have made it that a very large number of the “local languages” spoken by aboriginal populations, which are of great interest to typologists, have no children learning them at all (Hale, Krauss et al.

1992). Thus, it is impossible to study the acquisition of these languages by the most obvious methods. Conversely, even when people working on language learning overcome the practical challenges and manage to do a study in a daycare facility in Europe or East Asia, the knowledge gained is still only immediately relevant to two or three language families from among the dozens of language families and thousands of languages spoken in the world. This is a small and non-representative sample from the perspective of the typologist. As a result, acquisition experts and typology experts rarely have much direct contact with each other.

Although this disconnection between the two fields is easily understandable, there is nothing essential about it. On the contrary, one can argue that, from a theoretical and conceptual point of view, acquisition and typology should be closely related fields that constantly inform one another. There are at least two reasons why this should be true.

First, typological research can help show why language acquisition is possible. A big part of the typologist's job is to discover and understand the covariance among properties found in human languages. In the terminology of the trade (which originates with Greenberg (1963)), typologists seek and explain implicational universals—statements of the form “If language X has grammatical feature Y, then X also has grammatical feature Z.” Knowing implicational universals of this sort can help language acquisitionists with their fundamental problem, which includes explaining how human children take a finite amount of experience with a language and develop from it the capacity to create and understand an infinite variety of sentence structures in that language. The mismatch between the size of the input and the size of the potential output implies that there must be a great deal of indirect learning going on. In other words, for

some (many!) Xs and Ys, children's experience with linguistic domain X must enable them to draw conclusions about linguistic domain Y. The existence of implicational universals found by the typologists reveals one important way this can happen.

Second, typological research can be used to define a substantive lower bound for language learning. Since the beginning of modern generative linguistics about 50 years ago, there has been continual controversy over how much of our knowledge of language is innate. Generative linguists (i.e. Chomskyan linguists in the broadest sense) have argued that humans have a large innate endowment for language, and this is crucially what allows them to project the finite and idiosyncratic input they receive as young children into a rich and relatively invariant capacity to produce an unbounded variety of sentence structures. However, many language acquisitionists have been resistant to this idea, in part because it seems unparsimonious, and in part because it threatens to trivialize the whole notion of learning in a way that seems counterintuitive and counterproductive. This tension can manifest itself in discussions about whether such-and-such allegedly universal property of languages that the generative linguist claims to be innate could in fact be learned by the child from exposure to available input—discussions that (in my view) have in practice been relatively unproductive. But even the most extreme nativist must acknowledge that language cannot be entirely innate. If it were, then everyone in the world would speak essentially the same language, with at most the specific vocabulary items being different, and that is known to be false. So *at a minimum*, those grammatical features that distinguish one language from another must be learned. And typologists who study these differences have discovered that this is no small matter. While implicational universals are reasonably common, absolute universals—descriptive

statements that are true on the surface for all attested human languages—are extremely rare, and the few that are known (e.g. all languages have vowels; all languages have verbs) wouldn't get the learner very far by themselves. So it turns out that there is quite a bit that has to be learned on anyone's account. Suppose then that we concentrated on how these known differences among languages are learned. By doing so, we might learn some interesting and important results about the learning process that all parties could agree on. Then we could take what has been discovered about learning and see if the same processes or capacities could also be used to acquire the allegedly universal (and hence conceivably innate) aspects of language as well.

Given that there are great opportunities in principle for language acquisition and language typology to benefit from each other in these ways, in the remainder of this paper I as a formal typologist will try to communicate what my community (thinks it) knows about the nature of cross-linguistic variation in a way that can be relevant to people whose primary interest is language learning. I do this on several levels. First, I present some well-established examples of implicational universals and a theoretical understanding of them in terms of the notion of a parameter. These provide the opportunities for indirect learning, which projects knowledge of one linguistic domain into knowledge about another. Second, I'll present some relatively new and exploratory work on the logical relationships that exist among the known parameters. These have the result that an efficient learner should learn in a structured way, in which some parameters are entertained first while others are left for later. Finally, I discuss a few results from the (small) literature on the acquisition of parameter settings that suggest that the parameters are in fact learned in the order that one would expect based on our assessment of the

typological situation. This is offered as a kind of teaser, showing how knowledge of typology could fuel research into acquisition.

Overall, what typologists can offer acquisitionists could be something like a map of the territory they are concerned with. Based on our methods themselves, we cannot speak directly to the core learning question of what “mode of transportation” gets children from their starting point (the initial cognitive state) to their final destination (full knowledge of a particular human language, such as English). For all we know, it could be a sports car (a high tech, flexible device), or a subway car (a high tech inflexible [i.e., domain/path specific] device), or a skateboard (a low tech device)—or even some combination of these. But we can speak to the question of what “towns” and what crossroads lie on the path between the starting point and the final destination, of whether or not there are plausible alternative routes to the same destination, and of what some of the obstacles along the way are likely to be. This contribution should be of intrinsic interest in its own right. It might also give some baseline ideas of whether a car, a subway, or a skateboard would in principle be an effective method of getting from here to there. And this might in turn suggest better, more refined ways of testing which “modes of transportation” are used in practice.

2. Parameters and Indirect Learning

The first step is to see what kinds of implicational universals seem to be valid for natural languages, how they can be understood, and how they can feed into learning.

2.1 The basic word order parameter

The paradigm-defining examples of implicational universals come from research into word order. These were some of the first implicational universals to be discovered, and are some of the most robust and easy to understand. As a result, they are probably the most widely known of these phenomena. But they are worth reviewing, both because they are still not as well known as they could be, and because they present a clear example of a general way one might think about language.

The major constituents of a simple transitive sentence are (by definition) the subject, the object, and the verb. The subject is almost always the first element in such a sentence (with one major exception which I come to below); this is close to an absolute universal. However, the ordering of the verb and the object is quite variable. In languages that care about word order at all, the verb can come before the object, as in English, or the object can come before the verb, as in Japanese.

1

- (1) a. John ate a cheeseburger.
b. John-ga sushi-o tabe-ta. (Japanese)
John-NOM sushi-ACC eat-PAST

Moreover, roughly half of the languages of the world order words in the English way and half in the Japanese way (Tomlin 1986). In and of itself, this is not particularly remarkable; there are exactly two possibilities, and it doesn't matter which is chosen.

The interesting thing is that the choice that a language makes on this point seems to affect the way other words are ordered. Thus it was observed as far back as Greenberg's seminal (1963) paper that languages that put the verb before the object also put prepositions like 'in, on, under' before the noun phrase that they are associated with. In

contrast, languages that put the verb after the object also put the preposition after the associated noun (as a result, the words in question are more properly known as *postpositions*).

- (2) a. John went to the store with Mary.
b. John-ga Mary to Kobe ni it-ta. (Japanese, (Kuno 1973: 5))
John-NOM Mary with Kobe to go-PAST

There are further differences that go along with this as well, also discovered by Greenberg and his followers [Dryer, 1992 #548]. In English-style languages, when the tense marker is expressed as a separate word from the main verb, the tense marker or auxiliary verb comes before the main verb. In the Japanese-style languages, the tense marker or auxiliary comes after the main verb.

- (3) a. John will read a book; John is reading a book.
b. John-wa hon-o *yon-de* iru. (Japanese, (Kuno 1973: 141))
John-TOP book-ACC read-ING be
'John is reading a book.'

A fourth difference concerns the placement of the subordinating conjunctions that are found in complex sentences. In English, the subordinating conjunction comes before the embedded clause that it is associated with; in Japanese, the comparable element comes after it.

- (4) a. John thinks that Hiro showed the picture to Hanako.
b. Taro-ga Hiro-ga Hanako-ni syasin-o *mise-ta* to
Taro-NOM Hiro-NOM Hanako-to picture-ACC show-PAST that
omet-te iru.

Think-ING be

‘Taro is thinking that Hiro showed the picture to Hanako.’

As can be seen in (4), the cumulative effect of all these differences in word order can be quite striking.

A fifth difference comes inside noun phrases [Dryer, 1992 #548]. English has articles like *a* and *the* which go along with the main noun; they specify the exact meaning of the noun in approximately the same way that tense marker and auxiliaries specify the exact meaning of the verb. More specifically, in English these articles always come before the associated noun. Japanese happens not to have anything that corresponds to English articles, but Lakhota is a Native American language with similar word order; in this kind of language the article comes after the noun that it is associated with:

(5) John wowapi k’uhe oyuke ki ohlate iyeye. (Williamson 1984)

John letter that bed the under found

‘John found that letter under *the bed*.’

Putting this material together, we find that there is a whole family of implicational universals that can be stated something like this (cf. Lehmann {, 1973 #924}):

- (6) a. If a language puts the verb after the noun phrase it is associated with (i.e. the direct object), then it puts the preposition after the noun phrase it is associated with.
- b. If a language puts the verb before the noun phrase it is associated with (i.e. the direct object), then it puts the preposition before the noun phrase it is associated with.
- (7) a. If a language puts the verb after the direct object, then it puts the auxiliary

verb or tense particle after the corresponding main verb.

- b. If a language puts the verb before the direct object, then it puts the auxiliary verb or tense particle before the corresponding main verb.
- (8)
- a. If a language puts the verb after the direct object, then it puts the subordinating conjunction after the subordinate clause it is associated with.
 - b. If a language puts the verb before the direct object, then it puts the subordinating conjunction before the embedded clause it is associated with.
- (9)
- a. If a language puts the verb after the direct object, then it puts the article after the noun phrase it is associated with.
 - b. If a language puts the verb before the direct object, then it puts the article before the noun phrase it is associated with.

Note that exactly how these conditional statements are formulated is somewhat arbitrary. As far as we have seen so far, one could just as well reverse them, so that (for example) the position of the preposition is in the antecedent and the position of the direct object relative to the verb is in the consequent. Or one could state a conditional that relates the placement of articles directly to the placement of subordinating conjunctions. These details of the formulation are irrelevant, at least for present purposes.

The important truth behind these statements is that there is much less variety in how words are ordered in languages of the world than one might have thought. There are really only two primary stable word order patterns found in the world, and these two word order patterns show up over and over again in historically unrelated languages. For example, other languages with the English-style word order include the African language Edo, the Asian language Indonesian, and the Mesoamerican language Zapotec.²

- (10) a. Òzó tá wéé írèn ghá rhiè éfótò yè né!é ékpétìn. (Edo, Stewart p.c.)
 Ozo say that he will put photo in the box.
 ‘Ozo said that he will put a photo in the box.’
- b. Dia datang sehingga dia dapat ber-tjakap dengan Ali. (Indonesian)
 He come that he can INTR-talk with Ali
 ‘He came so that he could talk to Ali.’ [Chung, 1976 #633]
- c. Le Manuel w-ruu dxe ne w-et x-maa men. (Zapotec)
 ART Manuel C-leave day that C-die POSS-animal his
 ‘Manuel left the day that his animal died.’ (Black 1994: 83)

Other languages with the Japanese-style word order include the European language Basque, the African language Ijo, and the native American language Lakhota (see (5) above).

- (11) a. Jon-ek Miren-i egon-gelan liburu-a irakur-ri. (Basque)
 John-ERG Mary-to living.room-in book-the read-PAST
 ‘John read the book to Mary in living room.’
- b. a béle-bi-ò náma tua-a fa Nìmi. (Ijo)
 she pot-the-in meat put-NEG FA FUT
 ‘She will not put meat in the pot.’ (Carstens 2002:6)

The potential relevance of this discovery for language acquisition should be apparent. Suppose, for example, that a given language happens to have only one or two subordinating conjunctions, and these words are not that common. Then it could be that a child learning the language might not hear enough examples of the subordinating conjunctions to learn with confidence where they should be placed. (This is a

particularly plausible example, because children do not master multi-clause sentences until relatively late.) If, however, the child already knows some equivalent of the rules in (6)-(9), then indirect learning is possible. The child gets abundant evidence that (say) verbs come before direct objects in the ambient language, and that prepositions come before noun phrases. Thus, she can reliably “learn” that subordinating conjunctions come before embedded clauses in that language, without ever having heard one. She will, of course, have to have some direct experience in order to make use of this knowledge; for example, she would have to hear the rare word *tsi* before she could know that *tsi* is a subordinating conjunction in the Mohawk language. But she could learn this much more quickly by virtue of the rich and reliable expectations about where it should be—perhaps on as few as one or two exposures. Better yet, she may be able to use her indirectly acquired knowledge about word order to leverage knowledge about lexical items and their meanings by a form of syntactic bootstrapping (see Gleitman etc.). If she hears an unknown word coming before a subordinate clause, exactly where she would expect a subordinating conjunction to be, she might well conclude that this new word is in fact a subordinating conjunction. This in turn would give her a good idea of what kind of meaning the word should have—something that is difficult to observe directly for elements like subordinate conjunctions, whose meanings are quite abstract. (I assume that subordinating conjunctions in general have systematically different kinds of meanings from other words, for principled reasons; note that it is not easy to give a useful dictionary-style definition for them.)

Note that the implicational universals need not necessarily be 100% reliable for this kind of indirect learning to be effective. Typological studies suggest that in fact most

of them are not exceptionless, but are something like 95% reliable (see Dryer (1992) for some of the best statistics on this). For example, Hindi has Japanese-like word order for the most part, with verbs following objects, postpositions following noun phrases, and auxiliary verbs at the end of the sentence. Nevertheless, the subordinating conjunction *ki* ‘that’ comes before the embedded clause, not after it where one might expect it:

(12) Jaun soctaa hai ki laRkaa Merii-se baat karegii. (Dayal, p.c.)

John think be that boy Mary-with talk do

‘John is thinking that the boy will talk with Mary.’

This is a counterexample to the generalization in (8). But generalizations that are 95% accurate are still enough to facilitate indirect learning, one would suppose. 19 out of 20 times, the expectations generated by the implicational universals will be true and helpful. The benefits this confers might well more than outweigh the occasional confusion caused by the one time that the expectation turns out to be false.³ I return briefly to the possible theoretical implications of the fact that these “universals” tend not to be exceptionless at the end of this section.

Why do implicational universals like those in (6)-(9) hold? To some extent, linguists have different ideas about this (see note xx). But one attractive and plausible idea, which has its origins in Chomsky [, 1981 #1] and Stowell [, 1981 #101], is that all the observed correlations hold because they are the result from the operation of a single underlying mechanism. In this particular case, the underlying mechanism in question is the phrase-building component of the language faculty—the machine that Chomsky (1995) has dubbed “Merge”. This idea can be developed as follows. Notice that the ordering statements in (6)-(9) all have in common the fact that they define order between

two elements only if (i) one of them is a word level category (as opposed to a larger phrase), (ii) the two elements being ordered with respect to each other are adjacent, and (iii) the two elements form a unit with respect to meaning. For example, (6) implies that the preposition *to* comes before the noun phrase *the store* in a sentence like (2a). Here the preposition *to* is a single word (whereas the noun phrase *the store* need not be), it and the noun phrase are adjacent to each other, and it and the noun phrase form a meaning unit—in this case, they work together to define a location, as opposed to an object (the store itself). These are exactly the conditions under which we say that the preposition and the noun phrase together form a higher level phrase—more specifically, a prepositional phrase.⁴ The generalization in (6) is thus probably the result of the internal workings of the mechanism that builds this phrase. It can be restated as a phrase-building rule more or less as follows:

(13) English: When combining a preposition with a noun phrase to make a prepositional phrase, put the preposition *before* the noun phrase.

Japanese: ... put the preposition *after* ...

Now this rule can be generalized to cover other types of phrases by stating it as in (14), which then derives all the particular implicational universals in (6)-(9).

(14) *The Head Directionality Parameter*

English: When combining a word with a phrase in order to make a larger phrase, put the word *before* the phrase.

Japanese: ... put the word *after* ...

(13) is the special case of (14) when the word being combined happens to be a preposition. However, (14) applies equally well to the case when the word is a verb and

the phrase is a noun phrase; the result then is verb-direct object word order. We can also assume that tense markers and auxiliary verbs combine with the verb phrase (a phrase consisting of a verb and its objects) to form a phrase, traditionally known as the predicate. When this happens, the auxiliary/tense marker will come before the verb phrase as a whole in an English-type language; hence it comes before the verb in particular. Subordinate conjunctions combine with fully-formed clauses to create subordinate clauses; (14) implies that when this happens the subordinate conjunction will come first. The last kind of phrase to consider is the article+noun combination. The article *the* combines with a phrase like *pictures of Paris* to give the phrase *the pictures of Paris* (Abney 1987); (14) implies that the article comes first in this combination.

Languages like Japanese build the same kinds of phrases in essentially the same way; the only difference is that the phrase building machine in Japanese is configured in a slightly different way, so that the new word comes last and the pre-built phrase comes first. As a result, the word orders found in Japanese are exactly the opposite of those found in English over a wide range of syntactic structures. This is shown graphically in Figure 1, which presents two phrase structure diagrams of the kind that generative linguists are fond of. In such diagrams, each phrase is represented by a node in the tree diagram, and every word that belongs to a particular phrase is connected to the corresponding node (directly or indirectly) from below. For example, VP indicates the verb phrase, and the words *talking*, *with*, and *Pat*, all belong to this phrase. *With* and *Pat* also belong to a prepositional phrase (PP), but *talking* does not. Notice that the pattern of connections in the English version and the Japanese version are identical, and the differences in order are systematic; the order of words under each node (except the S

node, which stands for the clause as a whole) is the opposite in English from what it is in Japanese.

[INSERT FIGURE 1 ABOUT HERE]

The core idea behind this is that all the phrases in a particular language tend to have a common ordering for much the same reason that the china from a particular manufacturer tended to have similar coloration or the bullets shot from the same gun tend to have the same markings: they are all the result of a common process or machine. Choices like (14) which are embedded in the generative system of a language are known as *parameters*—a notion introduced by Chomsky (1981). Parameters are the theoretical explanation for why the implicational universals that can be observed by comparing languages hold. The various observable consequences of setting a parameter in one way as opposed to another is known as the *parametric cluster*.

There is, in fact, a significant descriptive gain that comes from looking at the matter in terms of parameters as opposed to explicit implicational universals. This stems from the fact that the statement in (14) is much more general than all the statements in (6)-(9) put together. For example, the statements in (6)-(9) do not, strictly speaking say anything about how a verb will be ordered with respect to phrases other than direct objects—prepositional phrases, for example, or embedded clauses. In contrast, (14) does: it automatically says that the verb should come before these sorts of phrases in English and after them in Japanese. And this is entirely correct, as one can see by looking at the examples in (2), (4), (5), (10), and (11). (14) also applies automatically to the formation of things like adjective phrases, which we have not even considered yet, in a way that is empirically warranted. Thus, we find adjective phrases like *smarter than Chris* in

English, with the adjective coming at the beginning, in contrast to adjective phrases like *Taro yori kasikoi* ‘Taro from smart(er)’ in Japanese, with the adjective at the end. So the parametric formulation makes possible much broader coverage than the initial formulation in terms of particular conditional statements.

The parametric approach to typology should make one feel that much more comfortable about using patterns like these to facilitate language acquisition via the use of indirect learning. It is always possible that a mere correlation holds for accidental reasons, and one does not want accidental correlations to drive the learning process. Fortunately, from the parametric point of view, these correlations are not at all accidental, but hold for a principled reason. There is also the question of what kind of prior knowledge the language learner needs to have to take advantage of indirect learning. While it is true that explicit conditionals like those in (6)-(9) could be used in indirect learning, one would have to ask where knowledge of those conditionals themselves came from. A child cannot learn them, because a child has access only to data from a small number of languages and these are necessarily (quasi-)universal statements. Conceivably they could be innately hard-wired into the child’s mind, but that may not seem so plausible in this format. It is plausible, however, that children’s minds come equipped with an innate phrase-building mechanism, which stands ready to group whatever words the child happens to learn into useful larger phrases. Setting the parameter in (14) is then simply a matter of tuning this one machine so that it functions in the right way. From this perspective it is not at all surprising that experience with (say) verbs and their direct objects should have repercussions for how subordinating conjunctions are ordered with respect to embedded clauses. The picture is analogous to calibrating one’s printer: if one

can get the colors on the test pattern to line up correctly, one expects the colors on all sorts of other graphics to line up correctly too—even graphics of a kind that you have never printed before.

Before going on, I must acknowledge that interesting theoretical questions are raised by the fact mentioned above that implicational universals tend to be 95% accurate rather than 100% accurate. What might this imply about exactly how parameters are used in the learning process?⁵ To make the issues more concrete, consider the following example taken from Nupe, a language spoken in Nigeria {Kandybowicz, 2003 #921}.

(15) Gana ta gaan Musa á tsi emi o.

Gana say that Musa PERF lie house LOC

‘Gana said that Musa has lain down in the house.’

For the most part, word order in Nupe is similar to word order in English: verbs come before objects, prepositional phrases, and embedded clauses; auxiliary-like elements (*á* PERF) come before the main verb; complementizers (*gaan* ‘that’) come before subordinate clauses, and so on. One thus wants to say that the Head Directionality Parameter is set to ‘before’ in Nupe, as in English. But there is an anomaly: the locative preposition *o* comes after the associated noun phrase *emi* ‘house’, not before it. What, then, does the existence of such anomalies show us about the nature of parameters and their role in acquisition?

At least three possibilities come to mind. The weakest possibility is that the parameters are heuristic guides to language learners, telling them what to expect but not what must be. This could be spelled out formally in a Bayesian learning framework, where learning how one phrase is ordered changes the prior probability for how another

phrase is ordered, but does not settle the matter. So Nupe children expect prepositional phrases to have the preposition first more strongly than Japanese children do, but they know that they must still learn a rule for prepositional phrase formation in particular, and they get compelling information that the preposition comes last that overrides the information received from the parameter.

A second, subtly different possibility is that parameters function as defaults. Unlike the heuristic view, on this view Nupe children set the Head Directionality Parameter as stated in (14), thereby acquiring a procedure that builds phrases of all kinds. Nevertheless, they also acquire a more specific rule that tells them something different about how to build prepositional phrases. This specific rule blocks the general rule from forming a prepositional phrase like *o emi* ‘LOC house’, by the general “Elsewhere” logic that is very familiar to (at least) linguists. This situation would be analogous to the way that the regular (default) rules of inflection are prevented from forming a word like *childs* in English by the existence of a more specific rule that gives the form *children* for this particular meaning.

The third and strongest view of parameters is to say that they (and the implicational universals they derive) are in fact exceptionless at the level to which they properly apply, but that level is an abstract one, and other factors influence how easily it can be observed. This view would say that all phrases in Nupe are indeed built by the same phrase-forming rule applying in the same way. Thus, even prepositional phrases are generated with the preposition coming before the noun phrase in Nupe, as in English. But there is more to syntax than just the building of phrase; there are also well-established “movement” processes that rearrange the initial structure in various ways and

for various reasons. Movement, for example, makes it possible to say *This book, I want to read* in English, even though direct objects should (and normally do) follow the verb that they are the object. Given this, it is perfectly possible that certain (yet-to-be-discovered) parameters of movement disrupt the original order of preposition and noun phrase in Nupe in an analogous way. This possibility is sketched graphically in (16).

(16) Gana ta gaan Musa á tsi emi [PP o <emi>].



Gana say that Musa PERF lie house LOC

Analyses of this kind have been made popular in the linguistics literature by Kayne [, 1995 #262]. On this view, parameters like (14) could guide learning in the strongest sense, such that children simply do not entertain any possibilities that are not consistent with the view that all phrases are initially built in the same way by the same syntactic engine. Parameters in this sense could play a very strong role in simplifying and reducing the search space that children need to find their way through. But of course what they learn when they acquire a setting for the Head Directionality Parameter is not a complete understanding of word order in their language; rather it is a partial understanding, to be supplemented by knowledge about (for example) the movement processes allowed in the language. These movement processes would presumably be constrained by parameters of their own that are relevant to this linguistic domain—although less is known about this.

Which of these three interpretations is correct? Are parameters best thought of as heuristic guides, as defaults, or as exceptionless principles in interaction with other factors? This is surely an empirical question, but it is a subtle one, and I do not believe

that we know enough to answer it yet. Therefore, I put it aside for now, identifying it as suitable topic for interaction and collaboration among learning theorists from different disciplines.⁶

2.2 Agreement parameters and their effects

Within the domain of word order parameters, the cluster of ideas I have presented may be familiar to a certain subset of readers. One reason for going through it carefully is that I want to build high on this foundation, so I need to be sure that it is well-laid. This parametric approach is, I believe, a powerful way of understanding the big picture of how languages differ from each other. In other words, the phrase structure parameter in (14) is not exceptional, but rather typical of language more generally. To illustrate this, I mention a few more implicational universals and the parameters that underlie them.

Apart from word order, another important grammatical phenomenon is agreement. English has a very modest amount of agreement in that the present tense verb takes different forms depending on whether its subject is singular or plural ((17a) vs. (17b)), and third person or first person ((17a) vs. (17c)):

- (17) a. The child *likes* spinach.
b. The children *like* spinach.
c. I *like* spinach.

Many other Indo-European languages have essentially the same agreement rule, but they manifest it more robustly. In Spanish, for example, there are different verb forms for all the different subject pronouns, and there is agreement in all the tenses, not just in present tense.

- (18) a. yo como 'I eat'
 b. tu comes 'you eat'
 c. el come 'he eats'
 d. nosotros comemos 'we eat'
 e. vosotros comeis 'you all eat'
 f. ellos comen 'they eat'

Now some languages have much less agreement than Indo-European languages.

Chinese, for example, has no agreement whatsoever; the verb looks exactly the same whatever subject it has:

- (19) a. wo kan-jian na ge ren.
 I see that CL person 'I see that person.'
 b. ni kan-jian na ge ren.
 you see that CL person 'You see that person.'
 c. ta kan-jian na ge ren.
 he see that CL person 'He see that person.'

In contrast, other languages have much more agreement than the Indo-European languages. Verbs in the Mohawk language, for example, show agreement with direct objects and subjects both (Lounsbury 1953; Postal 1979; Baker 1996). This is illustrated in (20), where the prefix on the verb changes in systematic ways if the subject is changed, but also if the direct object is changed.

- (20) *kenuhwe's* 'I like it.' *Rakenuhwe's* 'He likes me.'
senuhwe's 'You like it.' *Yanuhwe's* 'He likes you.'
ranuhwe's 'He likes it.' *Ronuhwe's* 'He likes him.'

yenuhwe's 'She likes it.' *Shakonuhwe*'s 'He likes her'
yakwanuhwe's 'We like it.' Etc. *Shukwanuhwe*'s 'He likes us.' Etc.

This is only a sample of the full Mohawk verb paradigm; in fact, there are 58 distinct verb prefixes in Mohawk, each of which expresses a different combination of subject and object.

Bantu languages like Chichewa (which is related to Swahili) are an interesting case that is intermediate between Mohawk and Spanish. Verbs in these languages show obligatory agreement with their subjects and optional agreement with their objects (Bresnan and Mchombo 1987). For example, the verbs in both (21a) and (21b) have the *zi-* prefix, showing agreement with the subject 'bees'. But only the verb in (21b) also has the *wa-* prefix, which agrees with the third person plural object 'hunters'. (21a) is similar to (18) in Spanish; (21b) is comparable to (20) in Mohawk.

- (21) a. Njuchi *zi-na-luma* alenje.
 Bees they-past-bit hunters
 'The bees stung the hunters.'
- b. Njuchi *zi-na-wa-luma* alenje
 Bees they-past-them-bit hunters.'
 'The bees stung the hunters.'

Although there is clearly a variety of different ways that agreement can show up in languages of the world, there are significant implicational universals at work too. For example, the following holds (see Croft (1990), among others):

- (22) No language shows more agreement with its objects than with its subjects.

Thus, there are languages like Spanish in which verbs agree with subjects but not objects, but there are no languages in which verbs agree with objects but not subjects. Similarly, there are languages like Chichewa in which verb agreement with subjects is obligatory and agreement with objects is optional, but there are no known languages in which verb agreement with objects is obligatory and agreement with subjects is optional.

The implicational universal in (22) seems to hold even at a finer degree of detail. For example, Mohawk verbs show different agreement if the subject is I+you as opposed to if it is I+others; basically, it has two different kinds of ‘we’. Mohawk verbs do not show this distinction for objects, however; there is only one kind of ‘us’, just as in English.

- (23) a. Tewa-nuhwe’s ‘We (I and you all) like it.’
 b. Yakwa-nuhwe’s ‘We (I and they, not you) like it.’
 c. Yukwa-nuhwe’s ‘It likes us (me and you, *or* me and them).’

 (Lounsbury 1953)

This can be seen as another manifestation of the generalization in (22).

Assuming that (22) is reliable and shows something about the nature of the grammar machine, it too can be used in indirect learning. If a child sees a clear manifestation of object agreement, for example, he can infer that the language also has subject agreement, even though some of the morphemes that express that agreement might be hard to recognize (perhaps it is largely concealed by phonological changes, for example). If he never sees a trace of subject agreement in simple intransitive sentence, then he knows that there is no point in investing time and energy looking for object agreement. And this logic probably works in more narrow domains, which are harder to

learn because there is less evidence available. For example, the child will not bother looking for a rare first person inclusive dual object agreement unless he has reason to believe that there is a first person inclusive dual subject agreement. On the other hand, if he happens to find an unusual object agreement marker, he knows to be on the look out for a comparable subject agreement marker.

In fact, knowledge about agreement can be used to leverage even more grammatical knowledge, due to a principle of grammar that can be stated crudely as follows (based on Chomsky [, 1981 #1] and Rizzi [, 1982 #85], implementing an idea originally due to Knut Tarald Taraldsen; see Jaeggli and Safir [, 1989 #568] for discussion and relevant qualifications):

- (24) If the verb shows [adequate] agreement with a given noun phrase, then the position of that noun phrase in the sentence is relatively free, and the noun phrase can readily be omitted from the sentence.

(Note that the converse of (24) is not true: there are languages that allow free word order and argument omission even when there is no agreement. Bidirectional implications are especially powerful for indirect learning, but simple implications have some value too.)

This generalization can be illustrated by comparing English and Spanish. We observed above that Spanish has much more robust agreement with the subject than English does. This correlates with the fact that in Spanish but not in English subject noun phrases can appear after the verb as well as before it, and Spanish subjects can be left out entirely.

- (25) a. Juan la vió.

John saw it

- b. La vió Juan

*Saw it John.

c. La vió. (He saw it)

*Saw it.

At an informal level, it is not so hard to see why this should be, although technical explanations differ. If there is enough subject agreement, word order is not needed to communicate which noun phrase is the subject; the subject must be the one noun phrase whose features match with those of the verb. Furthermore, if there is enough agreement on the verb, it is easy to see that an explicit subject pronoun could be redundant; hence it can be omitted without loss of information.

What makes (24) particularly interesting in the current context is it expands the effects of the agreement parameters alluded to above in ways that impact the syntax in a broader sense, creating more opportunities for indirect learning. For example, Mohawk verbs agree with their objects and Spanish verbs do not. Hence, by (24) it is no surprise that objects can appear on either side of the verb in Mohawk (but not in Spanish), and objects can be omitted entirely in Mohawk (but not in Spanish) [Baker, 1996 #137].

(26) a. Wa-shako-kv' ne eksa'a.

PAST-he/her-see the girl

'He saw the girl.'

b. Eksa'a wa-shako-kv'

girl PAST-he/her-see

'He saw the girl.'

c. Wa-shako-kv'

PAST-he/her-see

‘He saw her.’

(25) a. Vió la chica.

Saw the girl

b. #La chica vió.

The girl saw (bad as: ‘He saw the girl.’)

c. *Vió.

Saw (bad as: ‘He saw her.’)

More generally, direct objects have essentially the same grammatical behavior in Mohawk that subjects do, because they are agreed with in the same way that subjects are (Baker 1991; Baker 1996). That is not true in Spanish, where subjects and objects differ with respect to agreement, syntactic positioning, and other matters.

Here again the Bantu language Chichewa is an interesting intermediate case.

Subjects are always agreed with, can be freely ordered with respect to the verb, and can be omitted [Bresnan, 1987 #19]:

(28) a. Njuchi zi-na-lum-a alenje

Bees they-PAST-bit hunters

‘The bees bit the hunters.’

b. Zi-na-lum-a alenje njuchi

They-PAST-bit hunters bees

‘The bees bit the hunters.’

c. Zi-na-lum-a alenje

they-PAST-bit hunters

‘They bit the hunters.’

Objects are optionally agreed with. They show freedom of placement and omissibility if and only if the agreement with them is realized [Bresnan, 1987 #19]:

- (29) a. Zi-na-wa-lum-a alenje. (Or: Zi-na-luma alenje)
 they-PAST-them-bit hunters they-PAST-bit hunters
 ‘They bit the hunters.’
- b. Alenje zi-na-wa-lum-a (But not: *Alenje zi-na-luma.)
 hunters they-PAST-them-bit hunters they-PAST-bit
 ‘They bit the hunters.’
- c. Zi-na-wa-lum-a (But not: *Zi-na-luma.)
 they-PAST-them-bit they-PAST-bit
 ‘They bit them.’

Now recall that there is no language which agrees with objects more than subjects. When combined with the generalization in (24), this implies that there should be no language in which direct objects show greater freedom of placement than subjects or are easier to omit than subjects, and this seems to be true.

Overall, then we have the following agreement parameters:

(30) *The Agreement Parameters:*

Verbs in language x agree with:

- a. 0 NPs (Chinese, Japanese, Thai, etc.)
- b. 1 NP only (i.e., the subject, Spanish, Turkish, etc.)
- c. 1 or 2 NPs (Subjects and optionally objects, Chichewa)
- d. 2 NPs (Subjects and objects, Mohawk)

For convenience in what follows, rather than treating this as one parameter that can take several settings, I will treat this as four distinct parameters, each of which can be turned on or off. For example, Mohawk will have a positive setting for the 2-Agreement Parameter, whereas Spanish has a positive setting for the 1-Agreement Parameter. These simple parameters can then be used to leverage knowledge about a variety of other things, including verbal paradigms, and facts about the extent of free word order in the language. These inter-related properties are another example of a parametric cluster. Once again, languages are to a substantial degree interlocked wholes defined by a few simple choices in the sentence-building machines, and realizing this can simplify language learning considerably.

2.3 Two smaller-scale parameters

Parameters come in a variety of shapes and sizes. The word order parameter in (14) and the agreement parameters in (30) are two of the larger-scale examples, with an extensive impact on the “feel” of a language. Other parameters have more localized effects. Before going on, I briefly introduce two of the latter. In addition to giving further illustration of the range and scope of parametric phenomena, this will give us enough material so that we can consider how parameters interact with one another, and how these interactions might structure the learning process.

The first of these parameters was discovered by performing a close comparison of word order between English and French (Emonds 1978; Pollock 1989). English and French clearly have the same setting of the major word order parameters. Thus, subjects

come before predicates, auxiliary verbs come before verb phrases, and verbs come before objects in both languages:

- (31) a. John has often kissed Mary
b. Jean a souvent embrassé Marie.

In addition to the familiar elements already discussed, these sentences contain the optional adverbs *often* and *souvent*. We may assume that these are loosely attached to the beginning of the verb phrase without changing the intrinsic nature of that phrase. These adverbs do, however, reveal a subtle but systematic contrast between French and English. The difference shows up when there is no auxiliary verb or tense marker that is distinct from the main verb. In these circumstances, the tensed verb comes before the adverb in French and after the adverb in English:

- (32) a. John often kisses Mary. (Not: *John kisses often Mary.)
b. Jean embrasse souvent Marie (Not: *Jean souvent embrasse Marie.)

This difference has been understood in the following way. The tense element and the verb are, logically speaking, two distinct elements which make independent semantic contributions, and in some cases they show up as distinct words (e.g. ‘John *will kiss* Mary’). But in (32) these two conceptually distinct elements have fused into one. The difference between French and English comes from making different choices in exactly how this fusing is done:

(33) *The Verb Attraction Parameter*

Whenever tense is expressed by an affix rather than an independent word, either:

- a. The verb moves into the position originally occupied by the tense (French), or

- b. The tense attaches to the verb that is next to it, as long as no essential part of the sentence separates the two (English).

The crucial difference is in which of the two elements provides the fixed anchor for the combined word. In French, the verb moves and the tense marker is fixed; as a result the inflected verb appears before the adverb in (32b), the same position that the auxiliary verb occupies in (31b). In English, the tense marker moves and the verb stays fixed; thus the inflected verb appears after the adverb in (32a), the same position the main verb occupies in (29a). The difference can be shown graphically with the help of phrase structure diagrams and arrows to indicate movement, as in Figure Two.

[INSERT FIGURE TWO ABOUT HERE]

Emonds [, 1978 #680] and Pollock [, 1989 #77] demonstrate that this difference between French and English shows up also in negative sentences. In both languages the primary negative marker comes between the auxiliary and the main verb when the two are separate. In French, the simple inflected verb comes before the negative marker, as expected. In English, the negative marker blocks the tense from attaching to the verb, because it (unlike a simple adverb like *often*) counts as an essential part of the sentence, which intervenes between the two. The tense then needs to show up as an independent word, and a meaningless instance of the verb *do* appears for this purpose:

- (34) a. Jean (n') a pas mangé le gateau.
b. John has not eaten the cake.
c. Jean (ne) mange pas le gateau.
d. John did not eat the cake. (Not: *John not ate the cake.)

So (33) is a small-scale parameter, distinguishing two otherwise quite similar languages.

The last parameter I can discuss here can be seen by comparing French and English with the Celtic language Welsh. Words in Welsh are ordered in a way that is quite similar to how they are ordered in English and French: verbs come before objects, prepositions before NPs, auxiliaries before main verbs, subordinating conjunctions before embedded clauses, and articles before nouns (King 1993). These orders can be seen in the examples in (35). But there is one very striking difference, also observable in (35): subjects come after the verb in Welsh, not before the verb the way they do in French and English.

- (35) a. Bryn-odd y dyn gar.
 buy-PAST the man car.
 ‘The man bought a car.’
- b. Euthum i a Mair i’r sinema.
 went I with Mary to-the cinema
 ‘I went with Mary to the movies.’
- c. Disgwyliais i yr ennillai John
 expected I that would-win John
 ‘I expected that John would win.’

Indeed, this Verb-Subject-Object order is the third most common kind of word order in languages of the world, trailing only the Subject-Verb-Object word order of English/French/Edo/Indonesian and the Subject-Object-Verb word order of Japanese/Lakhota/Basque/Ijo. In fact, Zapotec also has this word order (see (10c)), as does Arabic and many languages of the Pacific. It is a minority type, but not an

inconsequential minority, being found in a little more than 10% of the languages of the world (Tomlin 1986).

However, in light of the discussion of French and English above, it is striking that it is only the finite, inflected verb that comes before the subject in Welsh. When the tense marker and the verb are distinct, the Welsh subject comes after the tense marker but before the main verb, as shown in (36).

- (36) Naeth y dyn brynu car.
did the man buy car
'The man did buy a car.'

Now compare (34) in Welsh with (29b) in French, and (33a) in Welsh with (30b) in French. The pattern of facts is essentially the same, except that the subject in Welsh appears in the same position that adverbs (and the negative marker) do in French. In particular, we can say that Welsh has the same setting for the Verb Attraction Parameter in (31) that French does: the inflected verb moves to the tense marker, and hence inherits its position relative to other sentence internal elements in Welsh as in French [Sproat, 1985 #675]. But there is another parameter that distinguishes the two languages; this has to do with where the subject appears [Koopman, 1991 #679]. This parameter can be stated as follows:

(37) *The Subject Placement Parameter*

- a. The subject noun phrase is merged together with the verb phrase
(a phrase consisting of the main verb and its objects, if any; Welsh), *or*
- b. The subject noun phrase is merged together with the auxiliary phrase
(a phrase consisting of the tense marker and the verb phrase; French, English).

In both types of language, the subject comes before the phrase it is merged with. Thus, in English and French, one first forms the predicate phrase [will [VP go to the store]] and then adds the subject to get [John + [will [VP go to the store]]]. In Welsh, one forms the verb phrase [VP go to the store], and adds the subject to that, giving [John + [VP go to the store]]. Then one forms the predicate phrase by adding the tense marker/auxiliary to this unit. By the Head Directionality Parameter in (14), the tense marker must come first, giving [will [John [VP go to the store]]]. Finally, if the tense marker is an affix, the verb abandons its original position, and moves to combine with the tense in accordance with (33a), giving the order [went [John [VP – to the store]]]. The difference between Welsh and English can also be represented in a phrase structure diagram, as shown in Figure 3.

[INSERT FIGURE 3 ABOUT HERE]

Thus, the Verb-Subject-Object order that is characteristic of Welsh and Zapotec is the result of three distinct parameters being set in a particular way: the Head Directionality Parameter is set to “before”, the Verb Attraction Parameter is set to “verb moves”, and the Subject Placement Parameter in (37) is set to “subject joins with verb phrase”. If each of these two-valued parameters has a 50-50 chance of being set in this way, we would expect approximately 1/8 of the languages of the world, or about 12.5% to be of this type, and that is quite close to the actual figure. That is another success for the parametric approach to cross-linguistic variation.⁷

3. The Parameter Hierarchy and Structuring the Learning Process

Thinking of language as a set of interlocked parameters, each of which influences a variety of particular constructions, simplifies the task of language learning in certain

respects. It means that learning does not necessarily have to proceed construction by construction. Rather, one can take advantage of indirect learning, where grammatical features learned in the context of one relatively common construction automatically apply to other constructions that are built by the same computational mechanism. This is a big advantage when it comes to learning the less-common constructions, of which there are many—probably infinitely many.

This way of looking at language does, however, raise another question, which is how the settings of the parameters themselves can be learned. This is a bit tricky, since it will often be the case that several parameters play into the formation of a single observable configuration. The clearest example of this we have seen is the explanation of the Verb-Subject-Object word order found in Welsh, where three different parameters all play a crucial role. If each parameter plays a role in shaping several observable constructions and each construction is the result of several parameter settings in interaction, it will often not be obvious how to untangle the threads to determine which observed structures provide reliable evidence for which parameter setting(s) (see, for example, Gibson and Wexler (1994) and Fodor (1998) for relevant discussion). In essence, this boils down to the question of which order the parameters should be learned in to achieve some kind of deterministic learning process and avoid errors from which it would be difficult to recover.

3.1 Ordering the Parameters

Some progress can be made on this issue by realizing that the kinds of parameters that are known for natural language syntax often enter into a particular kind of logical

relationship. I have already mentioned that some parameters have a much greater impact on the overall form of a language than others do. For example, the Head Directionality Parameter has a ubiquitous effect on phrases of all kinds. In contrast, the Verb Attraction Parameter in French as opposed to English has only a subtle effect on word order. It affects only one specific region of the clause, and even then the effect may only be detectable when an adverb or a negative particle is present.

In fact, some parameters may end up having *no* discernable effect on the way sentences are formed in the special case where some other parameter has been set in a way that makes the first parameter irrelevant in practice. This suggests the possibility of a useful way of ordering the parameters: parameters could be ordered by their power to affect one another, as well as their potentials for rendering each other irrelevant.⁸

To see a specific example of how this works, consider the logical relationship between the 2-Agreement Parameter (the Mohawk setting of (30)) and the Head Directionality Parameter. In terms of their actual statements, these parameters are concerned with different matters. The 2-Agreement Parameter determines how many of the participants of an action are represented by agreement on the verb that expresses the action, whereas the Head Directionality Parameter determines the order in which words are assembled into phrases. But there is an important relationship between them in practice. One of the consequences of setting the 2-Agreement Parameter positively is that agreement markers are included on the verb for both the subject and object. This means that the noun phrases expressing the subject and object notions can be freely ordered or omitted, in accordance with (24). The object noun phrase, in particular, does not combine with the verb to form verb phrases, as it does in other languages. The Head

Directionality Parameter is in practice irrelevant to the verb-object relationship in Mohawk. In this sense, the 2-Agreement Parameter “bleeds” the Head Directionality Parameter. Indeed, for the type of language that is most rich in agreement (so-called “head-marking languages”, in the terminology of Nichols (1986)), this reasoning carries over to other kinds of phrases. It so happens that agreement appears not just on verbs in languages like Mohawk, but also on nouns, prepositions, and adjectives, so word order is free in those phrases too [Baker, 1996 #137] . Because this is applied so consistently, the Head Directionality Parameter becomes irrelevant in practice in this sort of language, because the kinds of grammatical configurations it regulates never arise. As powerful as this parameter is in creating the extensive differences between English-style and Japanese-style word order, the Head Directionality Parameter is not the most powerful grammatical force; it can be rendered impotent by a particular setting of the Agreement Parameter. The 2-Agreement Parameter thus has a kind of logical priority, determining whether the Head Directionality Parameter can express itself or not. This is a principled reason for ordering the Agreement Parameter above the Head Directionality Parameter.

Now from the point of view of language learning, there would be good reason to translate this kind of logical priority into an acquisitional priority. In a well-designed system, it would make sense for the learner to try to establish the setting of the 2-Agreement Parameter before investing many resources into setting the Head Directionality Parameter, since the latter may turn out not even to be relevant. If the learner proceeds in this kind of structured fashion, the problem that many parameters influence the shape of many particular constructions will not be nearly so severe.

It turns out that it is not unusual for one parameter to have logical priority over another in this way. For example, we saw in section 2.3 that the Verb Attraction Parameter has important effects in ‘word-before’ languages such as English. It creates the Welsh and French variants of English-style word order, in which verbs come before adverbs and perhaps before subjects. But this same parameter has no appreciable effect on head-final languages like Japanese. It so happens that in head-final languages, the tense marker and the verb are already next to each other, so it doesn’t matter much whether the verb moves upward into the position of the tense marker or the tense marker attaches to the adjacent verb. Either way, the final word order will be Subject – Object – Inflected Verb. This can be seen in Figure 4.

[INSERT FIGURE 4 ABOUT HERE]

Even if an adverb were included, it would be attached to the verb phrase, coming between the subject and the object, so whether the verb moves or not would not make any difference for this either. This means that the Head Directionality Parameter has logical priority over the Verb Attraction Parameter, determining whether it gets a chance to express itself or not. In this case, a well-designed system would attempt to learn the Head Directionality Parameter before concerning itself with the Verb Attraction Parameter.

Stated in its most general form, the ordering principle I am proposing can be expressed as follows:

- (38) Parameter X is higher than parameter Y if and only if Y produces a difference in one type of language defined by X, but not in the other.

The corollary for acquisition is that the learner should seek to set Parameter X before Parameter Y whenever X is higher than Y.⁹

A little reflection on this matter shows that not all parameters will be ordered with respect to each other by these considerations. Sometimes one parameter will have an equivalent effect on both of the language types defined by another parameter, and vice versa. When this happens, the combination of two parameters that have two settings each gives a total of four distinct language types. The two parameters are logically independent—probably because they characterize non-interacting aspects of language. A case in point is the Head Directionality Parameter and the 1-2 Agreement Parameter, which says that the object (and similar constituents) can *optionally* be expressed as agreement on the verb that denotes the event. The 2-Agreement Parameter systematically bleeds the Head Directionality Parameter, because it completely prevents words from forming phrases of the relevant kind. The 1-2 Agreement Parameter is not so forceful: it creates possibilities but not requirements. A noun phrase may be omitted or freely ordered in such a language (when agreement is present), but it may also appear as a phrase together with the verb (when agreement is absent). Whenever this second option is taken, the Head Directionality Parameter determines whether the verb comes first or last in the phrase. In Chichewa, the verb comes before the object when there is no object agreement, as in English (see (29a,b)). In the Slave language (spoken in the Yukon Territory of Canada), on the other hand, object agreement is again optional, but the verb comes after the object when there is no object agreement, as shown in (39a) [Rice, 1989 #105].

(39) a. li ?ehkee ka-yihshu.

Dog boy it-bit.

‘The dog bit the boy.’

b. (?ehkee) li ka-ye-yihshu

boy dog it-him-bit.

‘The dog bit him (the boy).’

Thus all four of the logically possible ways of setting these two parameters generate clearly distinct looking-languages, as shown in Table 1. Therefore, these two parameters are not ordered with respect to each other.

[INSERT TABLE 1 ABOUT HERE]

The principle in (38) thus creates at best a partial ordering of the parameters. Note, however, that situations like this do not really make language learning more difficult. Since the two parameters are logically independent in this case, they can be learned in either order, and there is no complex interaction between them that affects what kind of evidence will be relevant to learning the setting of each parameter.

Consider next where the Subject Placement Parameter fits into the hierarchy being developed. This parameter concerns whether subject noun phrases are combined with the verb phrase or with the auxiliary phrase, working together with the Verb Attraction Parameter in ‘word-before’ languages to create Verb-Subject-Object order. It is clearly subordinate to the Head Directionality Parameter, because its effects are seen clearly only in the Subject-Verb-Object family of languages. In these languages, it determines whether the subject comes after an independent tense marker or inflected verb, as in Welsh, or before the tense marker or inflected verb, as in French and English. But in a Japanese-style, ‘word after’ language the tense marker and inflected verb always

come at the end anyway. Which phrase the subject merges with makes no easily detectable difference, as shown in Figure 5.

[INSERT FIGURE 5 ABOUT HERE]

The Subject Placement Parameter seems also to be subordinate to the Verb Attraction Parameter. I have discussed three types of languages that are created by these two parameters: the Welsh type, the French type, and the English type. Logically speaking, one might expect a fourth type of language to exist—one in which the subject comes between a tense auxiliary and the verb (as in Welsh), but tense affixes are attracted to the verb (as in English). Such a language would have sentences like those in (40) (where (40b) would be a simple declarative statement, not a yes-no question, which is the only interpretation it can have in English).

- (40) a. Chris buy+Past the car. (like English)
b. Will Chris buy the car. (like Welsh)

However, this combination of properties is not attested in the languages of the world (Baker 2002; Julien 2002). In Baker (2002) I argue that this is because the process that attaches the tense marker onto the verb in English is theoretically quite different from the process that combines the verb with the tense marker. The second process is a true movement, akin to passives or question movement. Like those well-studied phenomena, it always takes a linguistic element from a lower position to a higher position in the structure of the clause. The first process is more like a phonological restructuring, which takes two adjacent elements, one of which is “weak”, and packages them into a single word (this process is called “cliticization”). Since it is more phonological and less syntactic in nature, it (unlike true verb movement) is disrupted by having overt obligatory

sentence constituents appearing between the tense marker and the verb. The subject would be such a disrupting element in a “low subject” (Welsh-like) language. The upshot is that if the Verb Attraction Parameter is set in a particular way, the Subject Placement Parameter cannot be freely set to either value. Thus, the Verb Attraction Parameter can be ordered above the Subject Placement Parameter.

The last parameter to consider in my little sample is the 1-Agreement Parameter: the question of whether the verbs of a language agree robustly with one noun phrase (the subject), as in Spanish and Italian, or with zero noun phrases, as in Chinese. This parameter is closely related to the Pro-Drop parameter—historically the first parameter to be discussed (Chomsky 1981; Rizzi 1982). This says that in languages with “rich” subject agreement, the subject can be omitted and reordered. This is an affirmation of (24), together with the idea that languages with reduced, defective agreement paradigms like French (and English) do not really have subject agreement in the syntactically relevant sense. In its Pro-Drop guise, this parameter was originally presented as a matter internal to the Romance languages, distinguishing French (and English) from languages like Italian and Spanish. As such it would be relatively low in the hierarchy of parameters. Indeed, there is an important conjecture in the field that only a proper subset of the verb-attracting languages can be true Pro Drop languages (Alexiadou and Anagnostopoulou 1998). If this analysis is correct, then the Pro Drop Parameter would be ordered below the Head Attraction Parameter and probably below the Subject Placement Parameter as well. That is where I will put it.

Pulling all this together, the logical relationships among the parameters that we have considered can be diagrammed as in Figure 6. This diagram uses the following

conventions. If parameter X has logical priority over another parameter Y, then X is written higher than Y and is connected to Y by a downward slanting line. If two parameters are logically independent of each other, then they are written on the same line and are separated by a dash. For convenience, each parameter is assumed to have exactly two possible settings (although this will not necessarily always be true). If there is only one parameter at a level, then it has two branches going down from it, representing its two possible settings. If there are two independent parameters at a level, then there are four branches going out of the dash between them, representing the four possible combinations of settings for those two parameters. Since parameter Y is subordinate to another parameter X if and only if Y influences just one of the language types defined by X, it is natural to put Y at the end of the branch that represents the setting of X that Y influences. If there are no further parametric choices to be made given a particular setting of a parameter, then the branch ends in a terminal symbol *. Beneath this symbol I have (where possible) listed in italics two unrelated languages that have this combination of parameter settings.

[INSERT FIGURE 6 ABOUT HERE]

Figure 6 is a significant subpart of what I called the Parameter Hierarchy in Baker (2001), or, more picturesquely, the “Periodic Table of Languages”. In particular, it represents the deepest known part of the Hierarchy, which is the known parameters that are subservient to the “word-first” value of the Head Directionality Parameter. Presumably there are also smaller scale parameters that are subservient to the “word-last” value of the Head Directionality Parameter and to the 2 Agreement Parameter (see Baker (2001:ch.6) for some suggestions), but less is known about these

linguistically and much less is known in terms of languages acquisition, so I omit further discussion here.

3.2 Testing the Predictions: Initial Evidence

So far I have argued that one parameter often takes logical priority over another. I have also mentioned that a well-designed language acquisition system could take advantage of this fact to structure the learning process in an efficient way. Simply put, such a system should learn the settings for the higher ordered-parameters first, thereby constraining which of the narrow parameters need to be considered at all and fixing much of the linguistic background that is relevant to detecting the true setting of those parameters. The natural question to ask, then, is whether human children are in fact well-designed language learning systems in this sense. There is no doubt that they are efficient language-learners, given the way that syntactic capacities explode between 18 and 30 months. The question then is whether there is convergence between the typological studies I have reviewed and empirical studies of language development in this domain. To evaluate this possibility, we can consider what is involved in learning the grammar of French, as opposed to grossly similar languages like Spanish, Welsh, or English. The key question is whether the parameter settings that define French are acquired in the order that the Parameter Hierarchy predicts.

Unfortunately, nothing is known about the acquisition of the first parameter. The course of acquisition of what I have called the 2-Agreement Parameter has not been studied from this perspective, so far as I know. Once again, this is primarily for practical reasons: languages like Mohawk, for which the value of this parameter is crucial, happen

now to be relatively small and/or remote, with few children learning them. So there is an important gap in our knowledge here.

Consider the Head Directionality Parameter, then, which is the highest ordered parameter that remains. It is well-known that this parameter is learned extremely early. The earliest rigorous studies of language acquisition, such as Bloom (1970), Brown (1973), and Slobin's (1985) important crosslinguistic work, showed that children learning different languages show systematic differences in word order. Indeed, the very first two-word utterances of children, which appear between 18 and 22 months, show that they have already set the Head Directionality Parameter. An English-speaking one-year-old's first sentence might well be something like "Give cookie!" Japanese one-year-olds have comparable interests but their grammar is already different: they begin with "cookie give!" The prediction that this parameter would be one of the first ones set is thus correct.

After the setting of this major word order parameter has been established, the next one we expect children to attend to is the Verb Attraction Parameter. Déprez and Pierce (1993) have shown that this too is learned quite early. They studied the order of subjects, verbs, and negative adverbs in children learning French and English, starting at around 21 months. They found that children at this time already showed evidence of knowing that verbs move to the tense/auxiliary position in French but not in English. English-speaking children regularly utter sentences like *Not have coffee*, with the verb coming after the negative marker *not*. French children at the same stage say *Marche pas* ('works not', meaning 'it doesn't work') and *Veux pas lolo* ('want not milk'). In these examples, the verb has moved to tense and thus comes before the negative adverb *pas*. Interestingly, French children also use infinitive verb forms with negation. Since these

verbs have no tense value, we predict that they do not move to tense/auxiliary position, and hence should not come before *pas*. Strikingly, when French children use an infinitive, the verb consistently comes after *pas*, just as in the adult language. Examples are: *Pas casser* ('not to-break') and *Pas attraper papillon* ('not to-catch butterfly'). Children of 22 months have worked out that tensed verbs in French move forward, but French infinitival verbs and English verbs (whether tensed or not) do not.

According to the Parameter Hierarchy, the next matter to consider should be the Subject Placement Parameter, whether subjects are attached to Auxiliary Phrases or to Verb Phrases. Déprez and Pierce's (1993) study also bears on this question in an interesting way. They show that subject placement in French and English is not decided until slightly later, at about 24 months. As a result, both French and English children go through a stage in which their verbs are properly placed with respect to adverbs, but subjects are not properly placed. Thus, on their second birthday English children are wont to say sentences like:

(41) No I see truck.

No Leila have a turn.

Here the subject comes after the negative marker, not before it as in adult English.

(These examples also prove that children don't just mimic the orders they hear adults use in a superficial or conservative way.) French children of the same age use sentences in which the subject follows the tensed verb:

(42) a. Tombe Victor.

Falls Victor ('Victor falls.')

b. Veut encore Adrien du pain.

Wants more Adrien bread ('Adrien wants more bread.')

This can result in Verb-Subject-Object orders, as in (42b)—just as if the young French children were using a grammar like that of Welsh. At this point, the first three parameters in the hierarchy are set correctly, but the last two are not. English-speaking children begin to put the subject at the front of the sentence, attached to the Auxiliary Phrase, at an average age of 24.5 months. That, then, is roughly when they determine the correct setting for the Subject Placement Parameter.

At the bottom of the Parameter Hierarchy as I have given it is the Pro Drop Parameter, which determines whether clauses are required to have overt subjects or not. This is predicted to be learned relatively late, and that is correct (Hyams 1989). It has long been known that children learning English or French omit subjects as much as 50% of the time. They are much less likely to leave out direct objects. This results in utterances like the following:

- (43) a. Want go get it.
b. Not making muffins.

Similarly, French children of 26 months say *Est pas mort* ('Is not dead') for *Il est pas mort* ('It is not dead'). Hyams (1989) has observed that such sentences are grammatical in languages like Spanish and Italian. She concludes that these children have not yet determined the setting of the Pro Drop parameter. The average age for establishing this parameter value was 27 months in the children she studied—somewhat later than the setting of the Subject Placement Parameter—as predicted.

This information can be summarized in Table 2. This table repeats the Parameter Hierarchy from Figure 6, and adds estimates of when each parameter is learned, gleaned

from the (rather small) literature on acquisition that takes the notions of parameters seriously. Note that the times of acquisition do accurately mirror the logical relationships between the parameters.

[INSERT TABLE 2 ABOUT HERE]

I find this to be a very encouraging result. It should go without saying that the result is preliminary and tentative. Only a fraction of the true linguistic parameters have been identified by linguists, and only a fraction of those have been studied in terms of acquisition—and even those not so thoroughly. But as a kind of pilot inquiry I think that the results are striking enough to warrant further investigation along these lines.

4. Conclusions and Prospectus

The Parameter Hierarchy in Figure 6 is a kind of synthesis of what typologists have learned about how human languages vary. We have seen that this can provide a kind of roadmap of the terrain, which children can use in the process of language learning. There are still many parameters that need to be studied in this light. Also, methodologies need to be developed for timing these stages of acquisition accurately, since the pace of acquisition varies considerably from child to child, and many changes come very rapidly between 22 and 30 months. But if the pattern holds up under further scrutiny, this will be strong confirming evidence for the Parameter Hierarchy, as well as a useful tool in the study of language learning. The Hierarchy will not by itself answer the question of how children learn—not even how they learn the settings for parameters that they can already entertain. But it should be a great asset in this quest to know more exactly what children

are trying to learn and in what order, much as having a good roadmap will not get you to your destination, but it will be a big help to you in deciding how you will get there.

Acknowledgements

I thank Susan Goldin-Meadow and Lila Gleitman for encouraging me to write this article, and helping me to see what it could be. I also thank Jane Grimshaw and an anonymous reviewer for their comments on an earlier version, pushing me to clarify various points, discuss some related issues, improve the article in various respects. The articles remaining faults and errors are entirely my responsibility.

Many of the basic linguist topics and leading ideas touched on here are discussed more fully in my 2001 book, *The Atoms of Language*. Readers who become interested in a fuller exposition and additional references to the relevant linguistics literature are invited to look there.

Notes

¹Not all languages do care about word order. For example, any order of words is acceptable in many Australian languages, including Warlpiri (Hale 1983), and many Native American languages, including Mohawk) (Baker 1996). See section 2.2 below for some discussion of the latter case.

The following abbreviations are used in the glosses of linguistic examples: ACC, accusative case; ART, article; C, completive aspect; ERG, ergative case; FUT; future; INTR, intransitive marker; LOC, locative; NEG, negative; NOM, nominative case; PERF, perfect; POSS, possessive marker; TOP, topic marker. In many cases, I have suppressed detail in the glosses that would create needless confusion; for more, information, consult the original sources cited.

²While the implicational universals in (6)-(9) hold for Zapotec, it differs from the other languages mentioned in the placement of the subject. In English, Edo, and Indonesian, the subject comes before the verb (and the auxiliary, if applicable), whereas in Zapotec the subject often comes after the main verb. Note that the statements in (6)-(9) say nothing about the placement of the subject. I return to this difference when I discuss Welsh in section 2.3 below.

³ Jane Grimshaw points out that this is not necessarily true: sometimes a statistical generalization that produces the occasional wrong expectation can do more harm than good. Whether this is true or not depends on both the nature of the learning algorithm and the structure of what is being learned. My claim is only that sometimes probabilistic

knowledge is worth having. A more detailed investigation into these issues would go beyond both my space and my competence.

⁴ Note also that the preposition *with* comes after the noun phrase *the store* in (2a). This is not a counterexample to the generalization in (6b), because *with* is not associated semantically with *the store*; the two do not form a phrase. Rather *with* is grouped semantically with *Mary*, and it comes before *Mary*, as expected.

⁵ I thank an anonymous reviewer for pointing out this issue and pressing me to discuss it.

⁶ A related question—also raised by an anonymous reviewer—is how the sort of parametric explanation that I am sketching relates to functionalist accounts of linguistic universals, which often do not appeal to an innate universal grammar or explicit parameters. The issue is complex and goes beyond what I can discuss properly here, in part because there is a great diversity of functionalist explanations, so a short discussion is likely to do as much harm as good. But here are a few remarks that might help put the issues into context.

Two kinds of functionalist accounts can be distinguished. The first is intended to be on the same level of analysis as a parametric account, and constitutes an alternative to it. The second kind is intended as offering a deeper level of analysis, explaining why a particular parameter might be a useful part of the innate endowment for language. Some reasoning that is presented as an explanation of the first kind might well be reinterpreted as an explanation of the second kind, in my view. We can illustrate what this distinction might amount to by continuing with the example of word order.

The essence of the formal, parametric view as I have presented it is that the word order seen in phrases of different kinds tends to be uniform because all phrases are built

by a single rule or process, such as (14). This is a formal account in the sense that it explains why two linguistic facts correlate in terms of the internal structure of the language faculty (here the claim that there is only one phrase-building process). A functionalist explanation of the first kind might deny this. It could say that there are separate rules or processes that form each type of phrase, taking encouragement in this from the existence of mixed cases like Hindi and Nupe. Then the functionalist needs to explain why the uniform cases (like English and Japanese) are so much more common than the mixed cases, often by a ratio of 10 or 20 to 1. Typically they do this by appealing to something outside the language faculty per se that makes a uniform language better at the function of communicating than a mixed language is. For word order, the best developed proposal is Hawkins's [1990 #751; 1994 #752] claim that languages with uniform phrase structures are easier to parse. This kind of explanation is a genuine competitor to the parametric account I stated (but see below). In practice, the issues often come down to what to make of examples that seem to be less common or marked in some sense.

The other way a functionalist account could be pitched is that it could agree with the parametric account that all phrases are built by a single process, and then offer a functional explanation for why the language faculty is organized in that particular way. For example, one might offer a functional explanation—perhaps in terms of parsing and ease of language perception, à la Hawkins—for why there is a Head Directionality Parameter. There is no intrinsic conflict between parameters and functional explanation in this sense; rather the two can be complementary.

Finally, I note that in the end it seems almost certain that formal and functional explanations will be complementary in some respect, and that the two schools of linguistics present something of a false dichotomy. Consider again the functionalist alternative to the Head Directionality Parameter briefly sketched above. This assumes that how the various different kinds of phrases are formed is in principle independent. But this still accepts that all languages have consistent rules for how particular phrases are formed, and this is a kind of formal claim about the structure of the language faculty. Indeed, this view can be put in parametric terms: one could say that there are separate parameters for structuring the prepositional phrase, for structuring the verb phrase, for structuring the embedded clause, and so on. The question then is not really whether there are parameters or not, but what is the scale of those parameters—are there sweeping, large scale parameters, or only narrow, small scale parameters? That is an important (and debated) question, but not a question of formal vs. functional per se.

Newmeyer [, 1998 #627] is a useful source that discusses many of these questions in some depth (not that I agree with all of his conclusions).

⁷ Since the two parameters discussed in this section have a more localized effect on the language generated than the Head Directionality Parameter or the Agreement Parameter, they present fewer opportunities for the kind of indirect learning that is of interest for language learning. However, they may be more general than I have said. It has been known for a long time that the structure of noun phrases in English is quite parallel to the structure of sentences (see Chomsky [, 1970 #305], among many others). For example, compare the following:

- (i) a. Little Johnny will quickly destroy the new toy.

-
- b. The quick destruction of the toy; Johnny's quick destruction of the toy.
 - c. The/Johnny's big picture of Barney

In particular, the subject of the sentence in (ia) is parallel to the possessor of the noun phrase in (ib,c), the adverb in (ia) is parallel to the adjective in (ib,c), and the tense marker *will* in (ia) in the sentence is parallel to the articles *the* and *'s* in (ib,c). Note in particular that the subject-auxiliary-adverb-verb-other order in the clause is parallel to the possessor-determiner-adjective-noun-other order found in the noun phrase.

The order in noun phrases is different in French and Welsh, in a way that seems to be systematic. French has noun-adjective-other order in noun phrases, and this parallels the verb-adverb-object order found in clauses [Cinque, 1994 #544]. Welsh has Noun-Possessor-Adjective-Other order in noun phrases, and this parallels the verb-subject-adverb-object order found in clauses. These facts follow if the noun moves into the article position in French and Welsh, and if possessors are merged with the noun phrase rather than with the article-plus-noun phrase combination in Welsh [Ritter, 1991 #335]. If so, this suggests that the Verb Attraction Parameter applies to nouns too, and that the Subject Placement Parameter applies to possessors as well as to conventional subjects. In that case, there are opportunities for the child to learn about rare complex noun phrase structures indirectly, by transferring what they know about abundant clausal structures. I will not develop this possibility in detail here, partly for reasons of space, and partly because we have less detailed typological information about noun phrase structures than we do for clauses. (It has been known, however, since Greenberg (1963) that Verb-Subject-Object languages do consistently have Noun-Possessor order, so at least that part of the story is empirically well-grounded.)

⁸ This way of structuring the space of parameters was first discussed in Baker (2001), the work from which much of this material is drawn.

⁹ Jane Grimshaw raises the question of what happens if a clever theorist discovers some new technique or source of evidence for establishing whether the true structure of some Japanese-like language is in fact the verb movement variant shown in Figure 4 or the Tense movement variant. Would such a discovery show that the two parameters are really independent in the sense discussed just below, changing the predictions for language acquisition? I grant that this is a definite possibility; further research could very well reveal mistakes in how I have set things up here based on the current ‘state of the art’. But it might also be relevant whether the kind of data that the clever next-generation theorist appeals to in her argument is plausibly in the “primary linguistic data” that children reliably have access to. If not, then the hypothetical theoretical breakthrough would presumably not change the projected structure of the acquisition process. For these purposes, it shouldn’t matter so much whether two structures are distinguishable in theory as whether the child can reliably distinguish them in practice.

Jane Grimshaw also points out that one could have a symmetrical relationship between parameters, in which the right side of (38) holds, but so does its opposite, such that parameter X also produces a difference in one of the kinds of languages defined by parameter Y, but not in the other kind. This is a second way in which (38) could fail to get a determinate ordering for two parameters. I have no reason to think this is impossible, and it deserves more careful consideration. A minor amendment could be to change the left side of (38) to “Parameter X is not lower than parameter Y if and only

if...”. Then in this symmetrical case X and Y would receive the same ordering and should be learned at about the same time.

References

- Abney, Steven. 1987. The English noun phrase in its sentential aspect. Doctoral dissertation, MIT, Cambridge, Mass.
- Alexiadou, Artemis and Elena Anagnostopoulou. 1998. Parametrizing AGR: word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16: 491-539.
- Baker, Mark. 1991. On some subject/object non-asymmetries in Mohawk. *Natural Language and Linguistic Theory* 9: 537-576.
- Baker, Mark. 1996. *The polysynthesis parameter*. New York: Oxford University Press.
- Baker, Mark. 2001. *The atoms of language*. New York: Basic Books.
- Baker, Mark. 2002. Building and Merging, not Checking: the nonexistence of (Aux) SVO languages. *Linguistic Inquiry*.
- Black, Cheryl. 1994. Quiégolani Zapotec syntax. Doctoral dissertation, University of California, Santa Cruz, .
- Bloom, Lois. 1970. *Language development: form and function in emerging grammars*. Cambridge, Mass.: MIT Press.
- Bresnan, Joan and Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language* 63: 741-782.
- Brown, Roger. 1973. *A first language: the early stages*. Cambridge, Mass.: Harvard University Press.
- Carstens, Vicki. 2002. Antisymmetry and word order in serial constructions. *Language* 78: 3-50.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Croft, William. 1990. *Typology and universals*. Cambridge: Cambridge University Press.
- Deprez, Viviane and Amy Pierce. 1993. Negation and functional projections in early child grammar. *Linguistic Inquiry* 24: 25-68.
- Dryer, Matthew. 1992. The Greenbergian word order correlations. *Language* 68: 81-138.
- Emonds, Joseph. 1978. The verbal complex V'-V in French. *Linguistic Inquiry* 9: 151-175.
- Fodor, Janet Dean. 1998. Unambiguous triggers. *Linguistic Inquiry* 29: 1-37.
- Gibson, Edward and Kenneth Wexler. 1994. Triggers. *Linguistic Inquiry* 25: 407-454.
- Greenberg, Joseph. 1963. *Universals of language*. Cambridge, Mass.: MIT Press.
- Hale, Kenneth. 1983. Warlpiri and the grammar of nonconfigurational languages. *Natural Language and Linguistic Theory* 1: 5-49.
- Hale, Kenneth, Michael Krauss, Lucille Watahomigie, Akira Yamamoto, Colette Craig, Laverne Masayesva Jeanne and Nora England. 1992. Endangered languages. *Language* 68: 1-42.
- Hyams, Nina. 1989. The null subject parameter in language acquisition. In *The null subject parameter*, ed. Osvaldo Jaeggli and Kenneth Safir, 215-238. Dordrecht: Kluwer.
- Julien, Marit. 2002. *Syntactic heads and word formation*. New York: Oxford University Press.
- King, Gareth. 1993. *Modern Welsh: a comprehensive grammar*. London: Routledge.
- Kuno, Susumu. 1973. *The structure of the Japanese language*. Cambridge, Mass: MIT Press.

- Lounsbury, Floyd. 1953. *Oneida verb morphology*. New Haven, Conn.: Yale University Press.
- Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62: 56-119.
- Pollock, Jean-Yves. 1989. Verb movement, Universal Grammar, and the structure of IP. *Linguistic Inquiry* 20: 365-424.
- Postal, Paul. 1979. *Some syntactic rules of Mohawk*. New York: Garland.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris.
- Slobin, Dan, Ed. 1985. *The Crosslinguistic study of language acquisition*. Hillsdale, N.J.: L. Erlbaum Associates.
- Sproat, Richard. 1985. Welsh syntax and VSO structure. *Natural Language and Linguistic Theory* 3: 173-216.
- Tomlin, Russell. 1986. *Basic word order: functional principles*. London: Croom Helm.
- Williamson, Janis. 1984. *Studies in Lakhota grammar*. Ph.D. dissertation, UCSD.

Figure 1: English and Japanese Phrase Structure Compared

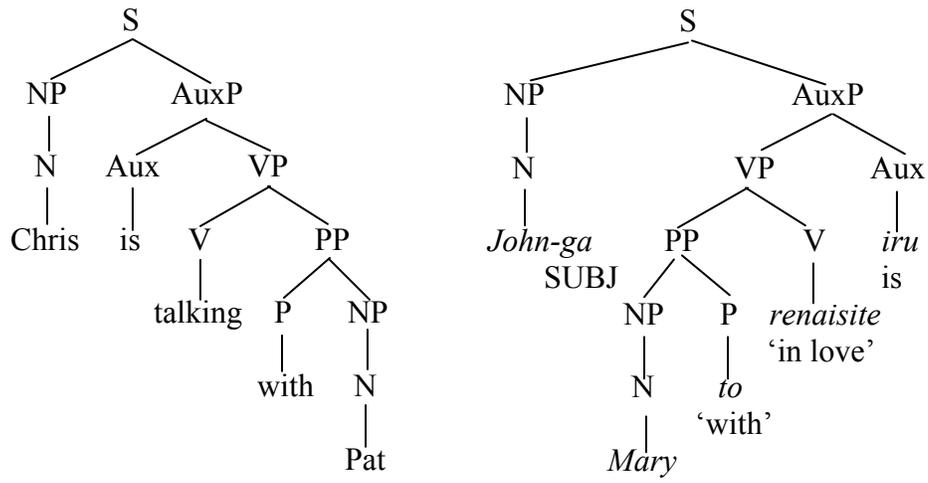
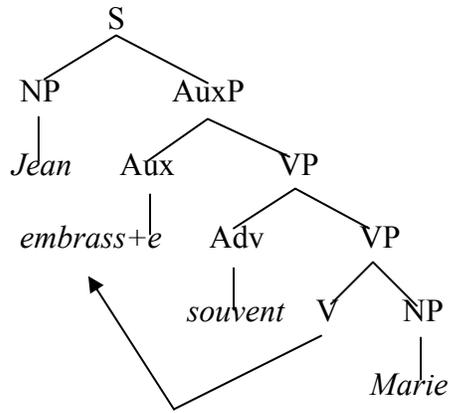


Figure 2: Verbs and Adverbs in French and English

French



English

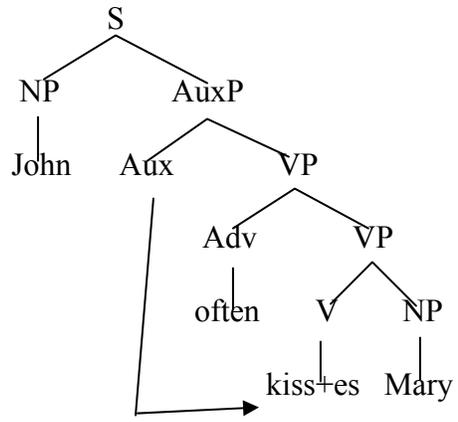


Figure 3: Subject placement in Welsh and English

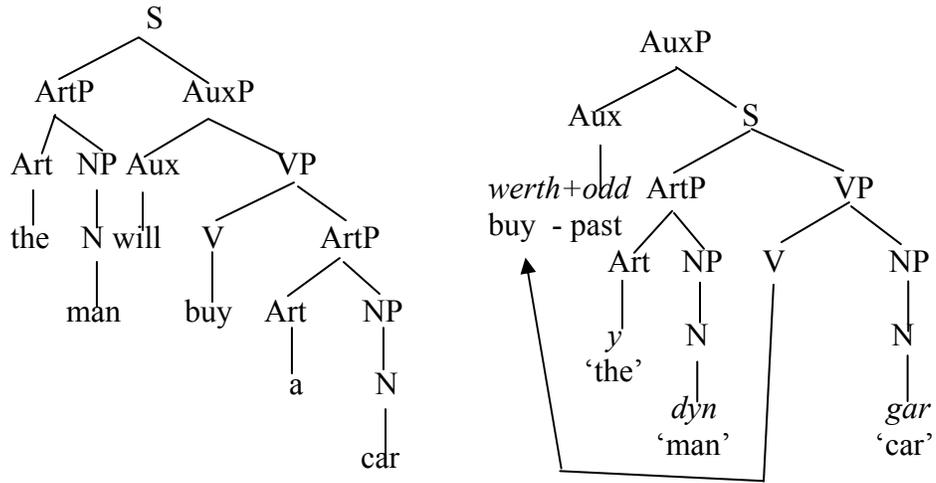


Figure 4: Verb and Auxiliary movement in a Japanese-style language

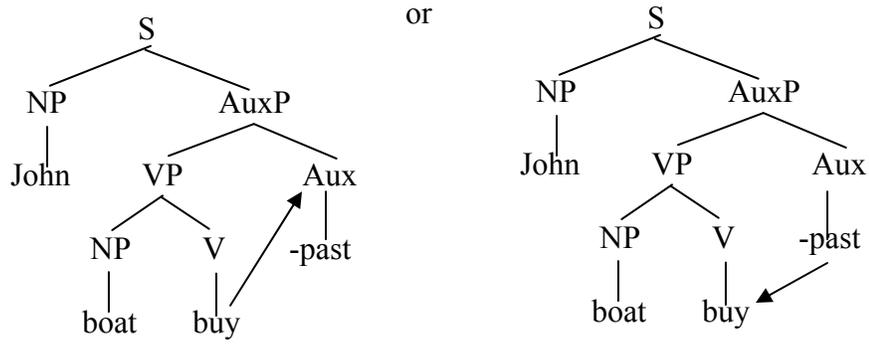


Figure 5: Subject Placement in a verb-final language

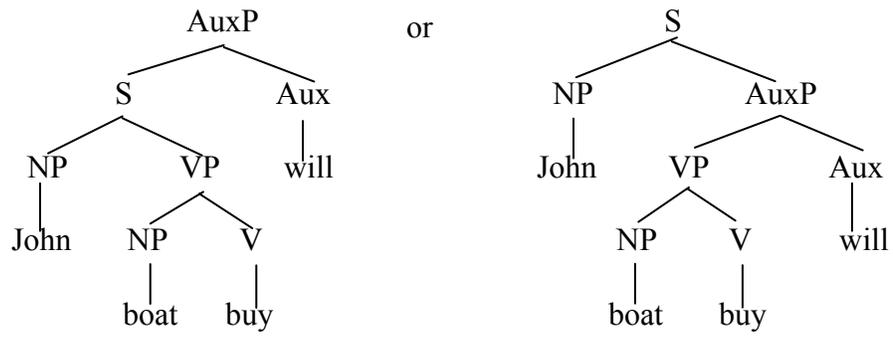


Figure 6: The Parameter Hierarchy (partial)

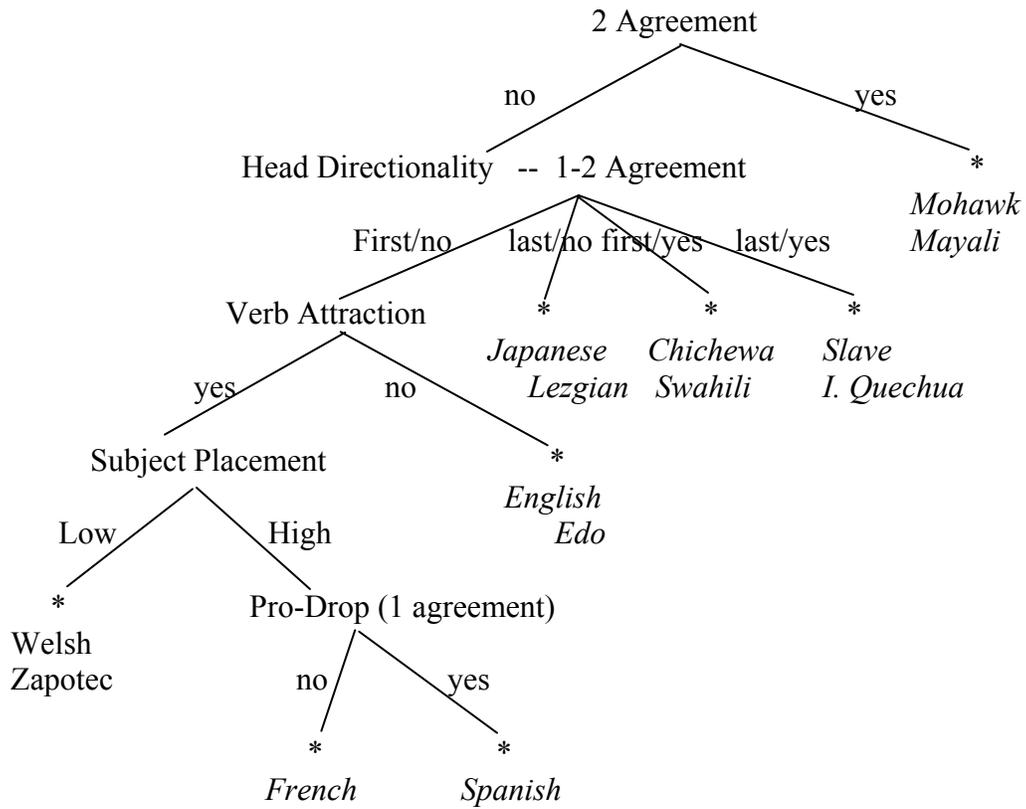


Table 1: Word Order and Agreement Interactions

	Word goes last in phrase	Word goes first in Phrase
1-2 Agreement	Slave	Chichewa
Not 1-2 Agreement	Japanese	English, Edo

Table 2: Time of Parameter Acquisition

Parameter Name	Ordering	Approx. Time of Acquisition
2 Agreement	1	Unknown
Head Directionality	2	By 18-21 months
Verb Attraction	3	By 22 months
Subject Placement	4	Around 24-25 months
Pro-Drop (1 agreement)	5	Around 27 months