

Mapping the Terrain of Language Learning

Mark C. Baker
Department of Linguistics
Rutgers University

This article explores how linguistic typology—the study of the variation found in human languages—can contribute to the study of language learning. All features that differentiate human languages must be acquired by their speakers, which suggests a concrete lower bound on what is learned. I discuss how the notion of the parameter explains observed implicational universals. I review some established parameters, and then I explore the idea that these parameters are logically ordered in a way that well-designed language learners can exploit to structure the learning process efficiently. Finally, I survey the literature on the acquisition of parameter settings by children, which suggests that children learn these settings in exactly the order predicted. Linguistic typology thus provides a precise linguistic map of the language-learning process.

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 but a circumstance based on some rather powerful practical reasons. The detailed study of language learning typically involves the experimental investigation of groups of people who are at various stages in the learning process—namely, children. Hence such research prototypically takes place at university day cares in North America where the children are 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. 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 traveling to areas of the world

where these such languages are spoken or by finding local speakers who happen to be in the vicinity say via 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 colossal number of the “local languages” spoken by aboriginal populations, which are of great importance to typologists, have no children learning them at all (Hale 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 day care 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 minute and nonrepresentative 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 each other, for at least two reasons.

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 originated 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 children take a finite amount of experience with a language and turn it into the capacity to create and understand an infinite variety of sentence structures. 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 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 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 have argued that humans have a large innate endowment for language that 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, which in my view have been relatively unproductive in practice. 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—a scenario that we know is not true. 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. Whereas 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) would not get the learner very far. 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.

Those in language learning and those in language typology have a great opportunity to benefit from one another. As a formal typologist, I try to communicate in the remainder of this article what my community knows (or 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 present some relatively new and exploratory work on the logical relationships that exist among the known parameters. These suggest that an efficient learner should learn in a structured way in which some parameters are entertained first whereas others are later. Finally, I discuss a few results from the available literature on the acquisition of parameter settings that suggests 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 is a quasi map of the linguistic territory. Given 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). For all we know, it could be a sports car (a high-tech flexible de-

vice), a subway car (a high-tech inflexible device, such as a domain- or path-specific device), a skateboard (a low-tech device), or even some combination. But we typologists can speak to the question of what towns and crossroads lie on the path between the starting point and the final destination, whether a destination has plausible alternate routes, and what potential obstacles are likely to be encountered along the way. Aside from its own intrinsic interest, this information 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, which in turn might suggest improved and refined ways of testing the modes of transportation used in practice.

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.

The Basic Word-Order Parameter

The paradigm-defining examples of implicational universals come from research into word order. These implicational universals were some of the first to be discovered and are the most robust and easy to understand. As a result, they are probably the most widely known, but they are worth reviewing 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 important exception, which I discuss in the section “Two Small-Scale Parameters”; 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.
 John-NOM sushi-ACC eat-PAST (Japanese)

¹Not all languages do care about word order. For example, any order of words is acceptable in many Australian languages, such as Warlpiri (Hale, 1983); and in many Native American languages, such as Mohawk (Baker, 1996). See the following section for 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.

Roughly half of the world's languages order words as English does and half as Japanese does (Tomlin, 1986). In and of itself, this is not particularly remarkable; there are exactly two possibilities, and it doesn't matter much which is chosen. The interesting observation 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 paper (1963) that languages that put the verb before the object also put prepositions such as *in*, *on*, and *under* before noun phrases. 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 in these languages are more properly known as *postpositions*.)

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

Further differences go along with this typology, also discovered by Greenberg and his followers (Dryer, 1992; Lehmann, 1973). 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.
 John-TOP book-ACC read-ING be
 'John is reading a book.' (Japanese; Kuno, 1973, p. 141)

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 Example 4, the cumulative effect of all these differences in word order can be quite striking.

A fifth difference comes inside noun phrases (Dryer, 1992). English has articles such as *a* and *the*, which go with the main noun; they specify the exact meaning of the

noun in approximately the same way that tense markers and auxiliaries specify the exact meaning of the verb. More specifically, these English 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 language, the article comes after the noun that it is associated with:

- (5) John wowapi k'uhe oyuke ki ohlate iyeye.
 John letter that bed the under found
 'John found that letter under the bed.' (Williamson, 1984)

Putting this material together, we find that there is a whole family of implicational universals that can be stated as follows (cf. Lehmann, 1973):

- (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, 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 éfòtò yé né!é ékpétin.
 Ozo say that he will put photo in the box.
 ‘Ozo said that he will put a photo in the box.’ (Edo; Stewart, personal communication)
- b. Dia datang sehingga dia dapat ber-tjakap dengan Ali.
 He come that he can INTR-talk with Ali
 ‘He came so that he could talk to Ali.’ (Indonesian; Chung, 1976)
- c. Le Manuel w-ruu dxe ne w-et x-maa men.
 ART Manuel C-leave day that C-die POSS-animal his
 ‘Manuel left the day that his animal died.’ (Zapotec; Black, 1994, p. 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 Example 5).

- (11) a. Jon-ek Miren-i egon-gelan liburu-a irakur-ri.
 John-ERG Mary-to living.room-in book-the read-PAST
 ‘John read the book to Mary in living room.’ (Basque)
- b. A béle-bi-ò náma tua-a fa Nìmi.
 she pot-the-in meat put-NEG FA FUT
 ‘She will not put meat in the pot.’ (Ijo; Carstens, 2002, p. 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 that these words are not that common. It could then be that a child learning the language might not hear enough examples of the subordinating conjunctions to learn with confidence where such words should be placed. If, however, the child already knows some equivalent of the rules in Items 6–9, then indirect learning is possible. The child gets abundant evidence that, say, verbs come

²Although the implicational universals in Items 6–9 hold for Zapotec, it differs from the other languages mentioned in its placement of the subject. In English, Edo, and Indonesian, the subject comes before the verb (and the auxiliary, if any), whereas in Zapotec the subject often comes after the main verb. Note that the statements in Items 6–9 say nothing about the placement of the subject. I return to this difference when I discuss Welsh in the section “Two Small-Scale Parameters.”

before direct objects in the ambient language and that prepositions come before noun phrases. Thus, the child can reliably “learn” that subordinating conjunctions come before embedded clauses, without ever having heard one. Of course, the child will have to have some direct experience to make use of this knowledge—for example, a child learning Mohawk would have to hear the rare word *tsi* before knowing that *tsi* is a subordinating conjunction. But the child could learn this rule quickly—perhaps as few as one or two exposures—by virtue of the rich and reliable expectations about where it should be. Better yet, the child may be able to use indirectly acquired knowledge about word order to leverage knowledge about lexical items and their meanings, via a form of syntactic bootstrapping (Fisher, Gleitman, & Gleitman, 1991; Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1994). If the child hears an unknown word coming before a subordinate clause—exactly where one would expect a subordinating conjunction to be—the child might well conclude that this new word is in fact a subordinating conjunction. This in turn would give the learner a good idea of what kind of meaning the word should have—something that is difficult to observe directly for elements such as 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; consider 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 issue). 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.
 John think be that boy Mary-with talk do
 ‘John is thinking that the boy will talk with Mary.’ (Veneeta Dayal, personal communication)

This is a counterexample to the generalization in Item 8. But generalizations that are 95% accurate are still enough to facilitate indirect learning, one would suppose: The expectations generated by the implicational universals will be true and helpful 19 out of 20 times. The benefits this confers might well more than outweigh the occasional confusion caused by a few occasions in which the expectation turns out to be false.³

³Jane Grimshaw has pointed 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 the nature of the learning algorithm and the structure of what is being learned. My claim is simply that, sometimes, probabilistic knowledge is worth having. A more detailed investigation into these issues would go beyond my space and competence.

At the end of this section, I return briefly to possible theoretical implications of the fact that these “universals” tend not to be exceptionless.

Why do implicational universals such as those in Items 6–9 hold? To some extent, linguists have different ideas about this question (see footnote 6). But one attractive and plausible idea, which has its origins in Chomsky (1981) and Stowell (1981), is that all the observed correlations hold because they 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 Items 6–9 all have in common the fact that they define order between two elements only if (a) one of them is a word-level category (as opposed to a larger phrase), (b) the two elements being ordered with respect to each other are adjacent, and (c) the two elements form a unit with respect to meaning. For example, Item 6 implies that the preposition *to* comes before the noun phrase *the store* in a sentence such as that in Example 2a. Here the preposition *to* is a single word, whereas the noun phrase *the store* need not be; the preposition and the noun phrase are adjacent to each other; and the preposition and the noun phrase form a meaning unit—in this case, they work together to define a location, as opposed to an ordinary object (the store itself). These are exactly the conditions under which the preposition and the noun phrase form a larger phrase—more specifically, a prepositional phrase.⁴ The generalization in Item 6 is probably the result of the internal workings of the mechanism that builds this phrase. Part of Item 6 can be restated as a more or less phrase-building rule:

- (13) When combining a preposition with a noun phrase to make a prepositional phrase, either
- a. put the preposition *before* the noun phrase (English); or
 - b. put the preposition *after* the noun phrase (Japanese).

This rule can then be generalized to cover other types of phrases by stating it as follows, which derives all the particular implicational universals in Items 6–9.

- (14) *Head directionality parameter*
When combining a word with a phrase to make a larger phrase, either
- a. put the word *before* the phrase (English); or
 - b. put the word *after* the phrase (Japanese).

Item 13 is the special case of Item 14, when the word being combined happens to be a preposition. However, the latter of the two applies equally well to the case when the

⁴Note also that the preposition *with* comes after the noun phrase *the store* in Example 2a. This is not a counterexample to the generalization in Item 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.

word is a verb and the phrase is a noun phrase; the result then is a 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 larger phrase, what is traditionally known as the *predicate*. When this happens, the auxiliary or 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; Item 14 implies that when this happens, the subordinate conjunction comes first. The last kind of phrase to consider is the article–noun combination. The article combines with a phrase such as *pictures of Paris* to give the phrase *the pictures of Paris* (Abney, 1987); Item 14 implies that the article comes first in this combination.

Languages such as 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 prebuilt phrase comes first. As a result, the word orders found in Japanese are exactly the opposite of those found in English and cover 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. Likewise, *Chris*, *Pat*, and *Mary* all belong to noun phrases (NPs). Notice that the pattern of connections in the English version and the Japanese version are identical and that 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 that in Japanese.

The core idea behind this analysis 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 tends to have similar coloration or that the bullets shot from the same gun tend to have the same markings: They are all the result of a common process. Choices such as those in Item 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 are known as the *parametric cluster*.

There is 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 Item 14 is much more general than are all the statements in Items 6–9 put together. For instance, the statements in Items 6–9 do not say any-

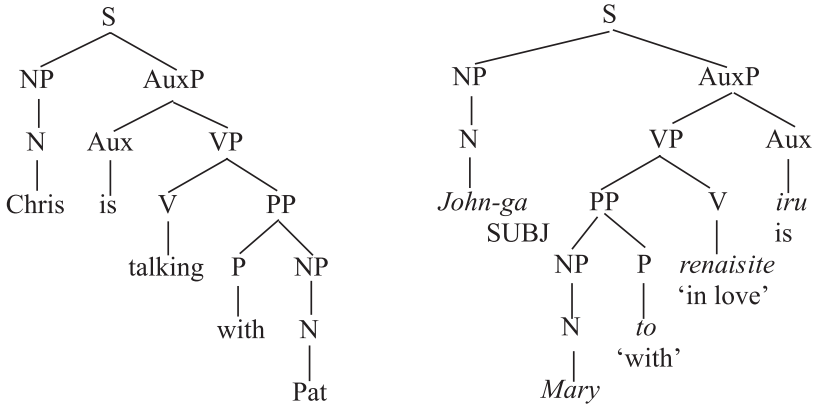


FIGURE 1 English and Japanese phrase structure compared.

thing about how a verb is ordered with respect to phrases other than direct objects—for example, prepositional phrases or embedded clauses. In contrast, Item 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 Examples 2, 4, 5, 10, and 11. Item 14 also applies automatically to the formation of adjective phrases, which we have not considered yet, in a way that is empirically warranted. Thus, we find adjective phrases such as *smarter than Taro* in English, with the adjective coming at the beginning; and adjective phrases such as *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 does the preliminary formulation in terms of particular conditional statements.

The parametric approach to typology should also make one feel comfortable about using patterns such as 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 rather hold for a principled reason. There is also the question of what kind of knowledge the language learner needs to have to take advantage of indirect learning. Although it is true that explicit conditionals, such as those in Items 6–9 could be used in indirect learning, one would have to ask where knowledge of those conditionals themselves comes from. A child cannot learn them, because a child has access only to data from a small number of languages and because these are necessarily quasi-universal statements. They could conceivably be innately hardwired into the child’s mind, but that idea may not seem so plausible in this for-

mat. It is plausible, however, to think 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 Item 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 earlier 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 & Baker, 2003).

- (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; subordinating conjunctions (*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 it is 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 paradigm, 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 the preposition to come first in the prepositional phrase more strongly than Japanese children do; but the former know that they must still learn a rule for prepositional-phrase formation in particular, and they get compelling information from direct observation that the preposition comes last. This overrides the information received from the parameter in Item 14.

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

A second, subtly different possibility is that parameters function as defaults. Unlike the heuristic view, this view dictates that Nupe children set the head directionality parameter, 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 quite familiar to linguists. This situation would be analogous to the way that the default rules of inflection are prevented from forming a word such as *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 that they are exceptionless at the level to which they properly apply, but that level is an abstract one; 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 applied in the same way. Thus, prepositional phrases are generated with the preposition coming before the noun phrase in Nupe, exactly as in English. But there is more to syntax than just the building of phrases; 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 of which 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 in the following example.

(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). On this view, parameters such as that in Item 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 powerful 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; it is a partial understanding, to be supplemented by knowledge about say the movement processes allowed in the language. These movement processes would presumably be constrained by pa-

rameters of their own that are relevant to this linguistic domain—although less is known about this issue.

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

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 one another. In other words, the phrase structure parameter in Item 14 is not exceptional but rather typical of language more generally. To illustrate, what follows are a few more implicational universals and the parameters that underlie them.

⁶A 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 to explicit parameters. The issue is complex and goes beyond what I can discuss 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 it constitutes an alternative to such an account. The second kind is intended as offering a deeper level of analysis, explaining why a particular parameter might be a useful aspect of human language. Furthermore, 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 that in Item 14. This is a formal account in 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 fact. 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 such as Hindi and Nupe. Then the functionalist needs to explain why the uniform cases (such as English and Japanese) are so much more common than the mixed cases, often by a ratio of 10 or 20 to 1. Typically, one does so 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 would be. For word order, the best developed proposal is Hawkins's claim (1990, 1994) that languages with uniform phrase structures are easier to parse. This kind of explanation is a genuine competitor to the parametric account I stated. In practice, the issue often comes to a head when presented with examples that seem to be less common or marked in some sense.

Apart from word order, another important grammatical phenomenon is agreement. English has a 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. Spanish, for example, has a verb form for every subject pronoun, and it also has agreement in every tense, not just in the 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 do Indo-European languages. Chinese, for example, has no agreement whatsoever; the verb looks exactly the same, regardless of its subject:

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

Finally, I note that in the end it seems almost certain that formal and functional explanations will be complementary in some respects 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, which assumes that the various kinds of phrases formed are done so, in principle, independently. But this notion still accepts that all languages have consistent rules for how particular phrases are formed—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 versus functional per se.

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

- (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.
s/he see that CL person
'He/she sees that person.'

In contrast, other languages have much more agreement than do the Indo-European languages. Verbs in the Mohawk language, for example, show agreement with direct objects as well as subjects (Baker, 1996; Lounsbury, 1953; Postal, 1979). This is illustrated in Example 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) | <i>kenuhwe's</i> | 'I like it.' | <i>Rakenuhwe's</i> | 'He likes me.' |
| | <i>senuhwe's</i> | 'You like it.' | <i>Yanuhwe's</i> | 'He likes you.' |
| | <i>ranuhwe's</i> | 'He likes it.' | <i>Ronuhwe's</i> | 'He likes him.' |
| | <i>yenuhwe's</i> | 'She likes it.' | <i>Shakonushe's</i> | 'He likes her.' |
| | <i>yakwanuhwe's</i> | 'We like it.' | <i>Shukwanuhwe's</i> | 'He likes us.' |

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

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

Thus, there are languages such as 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 such as 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 Item 22 also seems to hold 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, Mohawk has two different kinds of *we*. Mohawk verbs do not, however, show this distinction for objects; there is only one kind of *us*, just as in English.

- (23) a. Tewa-nuhwe's 'We like it.' (I and you all)
 b. Yakwa-nuhwe's 'We like it.' (I and they, not including you)
 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 Item 22.

Assuming that Item 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, the child 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 they are largely concealed by phonological changes). If the child never sees a trace of subject agreement in simple intransitive sentence, then the child knows that there is no point in investing much 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 there is reason to believe that a first-person-inclusive dual-subject agreement exists. However, if the child happens to find an unusual object agreement marker, then the child knows to be on the lookout for a comparable subject agreement marker.

In fact, knowledge about agreement can be used to leverage even more grammatical knowledge, owing to a principle of grammar that can be stated crudely as follows (based on Chomsky, 1981, and Rizzi, 1982, implementing an idea originally due to Knut Tarald Taraldsen; see Jaeggli and Safir, 1989, 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 this statement 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 value too. This generalization can be illustrated by comparing English and Spanish. We observed earlier 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 that 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 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 those of the verb. Furthermore, if there is enough agreement on the verb, an explicit subject pronoun is redundant; hence, it can be omitted without loss of information.

What makes Item 24 particularly interesting in the current context is that it expands the effects of the agreement parameters alluded to earlier in ways that affect the syntax, thereby creating opportunities for indirect learning. For example, Mohawk verbs agree with their objects, and Spanish verbs do not. Hence, by Item 24 it is no surprise that objects can appear on either side of the verb in Mohawk (but not in Spanish) and that objects can be omitted entirely in Mohawk (but not in Spanish; Baker, 1996).

- (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.' (Mohawk)
- (27) 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.'; Spanish)

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

Once again, the Bantu language Chichewa is an interesting intermediate case. In this language, subjects do not always have agreement; they can be freely ordered with respect to the verb; and they can be omitted (Bresnan & Mchombo, 1987):

- (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 have optional agreement: They show freedom of placement and omissibility if and only if the agreement with them is realized (Bresnan & Mchombo, 1987):

- (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 that agrees with objects more than subjects. When combined with the generalization in Item 24, this implies that there should be no language in which direct objects show greater freedom of placement than do subjects or are easier to omit than are subjects. This conclusion seems to be true.

Overall, then we have the following agreement parameters:

- (30) *The agreement parameters*
 Verbs can agree with
 a. 0 NPs (Chinese, Japanese, Thai, etc.)
 b. 1 NP only (the subject; Spanish, Turkish, etc.)
 c. 1 or 2 NPs (subjects and optionally objects; Chichewa)
 d. 2 NPs (subjects and objects; Mohawk)

Rather than treat this as one parameter that can take several settings, I regard this as four distinct parameters, each of which can be turned on or off. For example, Mo-

hawk has a positive setting for the two-agreement parameter, whereas Spanish has a positive setting for the one-agreement parameter. These simple parameters can then be used to leverage knowledge about a variety of other matters, including verbal paradigms and facts about the extent of free word order in the language. These interrelated properties are another example of a parametric cluster. We see once again that languages are to a substantial degree interlocked wholes defined by a few simple choices in the sentence-building machines; realizing this can simplify language learning considerably.

Two Small-Scale Parameters

Parameters come in a variety of shapes and sizes. The word-order parameter in Item 14 and the agreement parameters in Item 30 are two of the large-scale examples that have an extensive impact on the overall “feel” of a language. Other parameters have 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, doing so gives 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 settings for the major word-order parameters. Thus, in both languages, subjects come before predicates, auxiliary verbs before verb phrases, and verbs before objects:

- (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*. One 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, semantically speaking, two distinct elements that make independent

contributions to the meaning; and in some cases, they show up as distinct words (e.g., *John will kiss Mary*). But in Example 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 Example 32b, the same position that the auxiliary verb occupies in Example 31b. In English, the tense marker moves and the verb stays put; thus, the inflected verb appears after the adverb in Example 32a, the same position that the main verb occupies in Example 29a. The difference can be shown graphically with the help of phrase structure diagrams and arrows to indicate movement, as in Figure 2.

Emonds (1978) and Pollock (1989) have demonstrated that this difference between French and English shows up also in negative sentences. In both languages the primary negative marker comes between the auxiliary verb 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

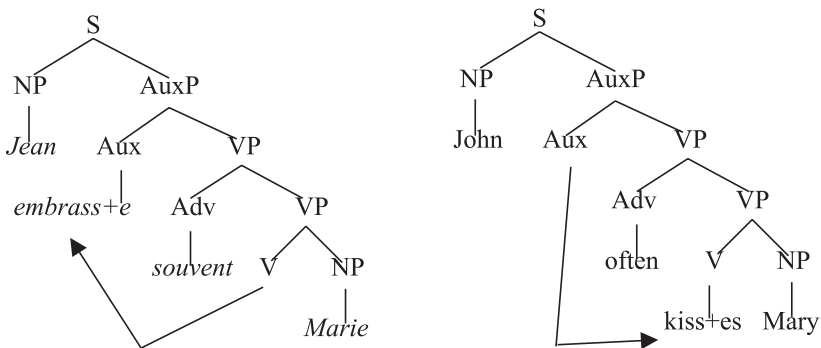


FIGURE 2 Verbs and adverbs in French and English.

from attaching to the verb because it (unlike a simple adverb, such as *often*) counts as an essential part of the sentence and thus 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 gâteau.
 b. John has *not eaten* the cake.
 c. Jean (ne) *mange pas* le gâteau.
 d. John *did not eat* the cake. (not *John not ate the cake.)

Example 33 then is a small-scale parameter, distinguishing two otherwise quite similar languages.

The last parameter I 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 before objects, prepositions before noun phrases, auxiliaries before main verbs, subordinating conjunctions before embedded clauses, and articles before nouns (King, 1993)—orders that can be seen in the following example. But there is one striking difference: In Welsh, subjects come after the verb, not before, as 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 characteristic of English, French, Edo, and Indonesian; and the subject–object–verb word order characteristic of Japanese, Lakhota, Basque, and Ijo. Zapotec also has this word order (see Example 10c), as do 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, it is striking that only the finite inflected verb 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 Example 36.

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

Now compare Example 34 in Welsh with Example 29b in French, and Example 33a in Welsh with Example 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 (Item 33) that French does: The inflected verb moves to the tense marker; hence, it inherits its position relative to other sentence-internal elements in Welsh as in French (Sproat, 1985). But there is another parameter that distinguishes the two languages; it has to do with where the subject appears (Koopman & Sportiche, 1991). This parameter can be stated as follows:

- (37) *The subject placement parameter*
 The subject noun phrase is merged with
- a. the verb phrase, a phrase consisting of the main verb and its objects (Welsh); or
 - b. 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 or auxiliary to this unit. By the head directionality parameter in Item 14, the tense marker must come first, giving [will [John [VP go to the store]]]. Finally, if the tense marker is an affix, then the verb abandons its original position and moves to combine with the tense in accordance with Item 33a, giving the order [went [John [VP to the store]]]. The difference between Welsh and English can be represented in a phrase structure diagram, as shown in Figure 3.

All in all, 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 is set to “subject joins with verb phrase.” If each of these two-valued parameters has a 50–50 chance of being set either way, we would expect approximately one eighth of the languages of the

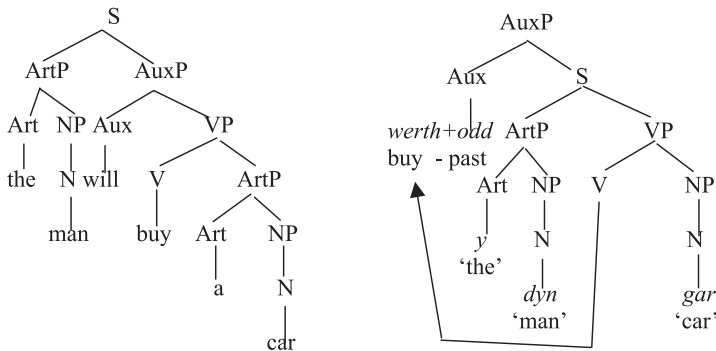


FIGURE 3 Subject placement in Welsh and English.

world, or about 12.5%, to be of this type—and that is quite close to the actual figure, thus another success for the parametric approach to cross-linguistic variation.⁷

THE PARAMETER HIERARCHY AND THE STRUCTURING OF THE LEARNING PROCESS

Thinking of language as a set of interlocked parameters, each of which influences a variety of particular constructions and 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 com-

⁷Since 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 indicated so far. 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, among many others). For example, compare the following:

1. Little Johnny will quickly destroy the new toy.
2. The quick destruction of the toy,
Johnny's quick destruction of the toy.
3. The/Johnny's big picture of Barney

The subject in Sentence 1 is parallel to the possessor of the noun phrase in Sentences 2 and 3; the adverb in Sentence 1 is parallel to the adjective in Sentences 2 and 3; and the tense marker *will* in Sentences 1 is parallel to the articles *the* and *'s* in Sentences 2 and 3. 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.

putational 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 turns out to be a bit tricky, since it will often be the case that several parameters play into the formation of an observable configuration. The clearest example of this is the explanation of the verb–subject–object word order found in Welsh, where three parameters play a crucial role. If each parameter plays a role in shaping several observable constructions and if each construction is the result of several parameter settings in interaction, it will often be unapparent how to untangle the threads to determine which observed structures provide reliable evidence for which parameter settings (for relevant discussion, see Gibson and Wexler, 1994; Fodor, 1998). For the learner, this boils down to the question of which order the parameters should be learned to achieve some kind of deterministic learning process and to avoid errors from which it would be difficult to recover.

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 that in English, has only a subtle effect on word order. It affects only one specific region of the clause, and even there 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 another parameter has been set in a way that makes the first one irrelevant in practice. This scenario suggests the possi-

Word 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). Welsh has noun–possessor–adjective–other order in noun phrases, which 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–noun phrase combination in Welsh (Ritter, 1991). If so, this suggests that the verb attraction parameter also applies to nouns 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 relatively rare, complex noun–phrase structures indirectly by transferring what they know about abundant clausal structures. I do 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 analysis is empirically well grounded.)

bility of a useful way of ordering the parameters, perhaps by their power to affect one another and by their potentials for rendering one another irrelevant.⁸

To see a specific example of how this ordering works, consider the logical relationship between the two-agreement parameter (the Mohawk setting of Item 30) and the head directionality parameter. In terms of their actual statements, these parameters are concerned with different matters. The two-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 two-agreement parameter positively is that agreement markers are included on the verb for the subject and object. This construction means that the noun phrases expressing the subject and object notions can be freely ordered or omitted, in accordance with Item 24. The object noun phrase, in particular, does not combine with the verb to form a verb phrase the way it does in other languages. The head directionality parameter is in practice irrelevant to the verb–object relationship in Mohawk. In this sense, the two-agreement parameter “bleeds” the head directionality parameter. Indeed, for the type of language that is most rich in agreement (so-called head-marking languages; Nichols, 1986), this reasoning carries over to other kinds of phrases. It so happens that agreement appears not just on verbs in languages such as Mohawk but also on nouns, prepositions, and adjectives; so, word order is free in those phrases too (Baker, 1996). Because agreement applies 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 two-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 two-agreement parameter before investing many resources into setting the head directionality parameter, since the latter may be irrelevant. If the learner proceeds in this kind of structure, the problem of parameters’ influencing the shape of particular constructions will not be so severe.

It turns out that it is not unusual for one parameter to have logical priority over another in this way. For example, in the section “Two Small-Scale Parameters” we

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

saw 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 such as Japanese. It so happens that in head-final languages, the tense marker and the verb are already next to each other; so, it does not matter much whether the verb moves upward into the position of the tense marker or whether the tense marker attaches to the adjacent verb. Either way, the resulting word order will be subject–object–inflected verb. This construction can be seen in Figure 4.

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 still not make any difference. 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 one another by these considerations. Sometimes one parameter will have

⁹Jane Grimshaw raised the question of what happens if a clever theorist discovers some new technique or source of evidence for establishing whether the true structure of a particular 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, thereby changing the predictions for language acquisition? This is a definite possibility, and further research could very well reveal mistakes in how I have set things up here. But it might also be relevant to consider whether the clever next-generation theorist’s data plausibly reside within the “primary linguistic data” of which children have reliable access. If not, then the hypothetical theoretical breakthrough would presumably not change the projected structure of the acquisition process. For these purposes, it should not matter so much whether two structures are distinguishable in theory; what should matter is whether the child can distinguish them in practice.

Jane Grimshaw has also pointed out that one could have a symmetrical relationship between parameters in which the right side of Item 38 holds, but so does its opposite such that parameter X 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 Item 38 could fail to define a determinate ordering for two parameters. I have no reason to think it impossible; the case certainly deserves careful consideration. A minor amendment could be to change the left side of Item 38 to “Parameter X is not lower than parameter Y if and only if ...” Then, in the symmetrical case, X and Y would receive the same ordering. The prediction would thus be that the two parameters should be learned at about the same time.

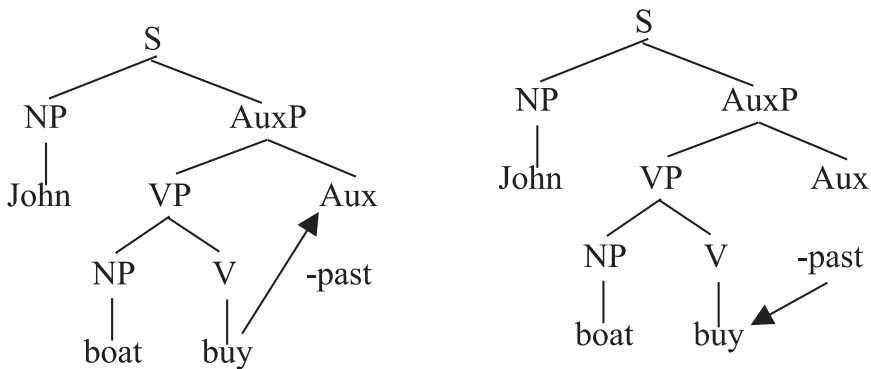


FIGURE 4 Verb and auxiliary movement in Japanese-style language.

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 noninteracting aspects of language. A case in point is the head directionality parameter and the one/two-agreement parameter, which says that the object (and similar constituents) can optionally be expressed as agreement on the verb that denotes the event. The two-agreement parameter systematically bleeds the head directionality parameter because it completely prevents words from forming phrases of the relevant kind. The one/two-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 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 Example 29a and 29b). In contrast, in the Slave language (spoken in the Yukon Territory of Canada), object agreement is again optional; but the verb comes after the object when there is no object agreement, as shown in 39a (Rice, 1989).

- (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.

The principle in Item 38 thus creates at best a partial ordering of the parameters. Note, however, that situations such as 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. 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 an 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 the inflected verb always come at the end anyway. Which phrase the subject merges with makes no easily detectable difference, as shown in Figure 5.

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. 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

TABLE 1
Word Order and Agreement Interactions

	<i>Word Goes Last in Phrase</i>	<i>Word Goes First in Phrase</i>
1–2 Agreement	Slave	Chichewa
Not 1–2 Agreement	Japanese	English, Edo

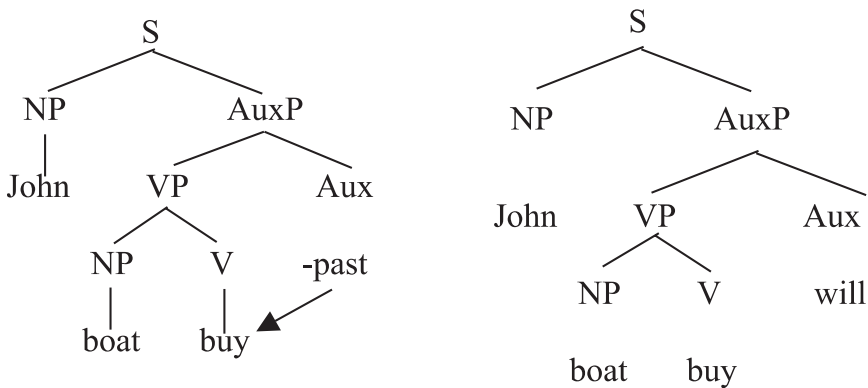


FIGURE 5 Subject placement in a verb-final language.

to the verb, as in English. Such a language would have sentences like those in Example 40, where Example 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 explanation, I have argued (Baker, 2002) that 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 like a phonological restructuring, which takes two adjacent elements (one of which is “weak”) and packages them into a single word. Since it is more phonological and less syntactic in nature, it (unlike true verb movement) is disrupted by 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 one-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), which says that in languages with “rich” subject agreement the subject can be omitted and reordered. This is an affirmation of Item 24, with the idea that languages with reduced, defective agreement paradigms—such as those in 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 such as 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 & Anagnostopoulou, 1998). If this analysis is correct, then the pro-drop parameter would be ordered below the verb attraction parameter and probably below the subject placement parameter as well—and that is where I will put it.

In sum, 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 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 terminal symbol *. Beneath this symbol I have listed (where possible) two unrelated languages in italics that have this combination of parameter settings.

Figure 6 is a significant subpart of what I have called the *parameter hierarchy* (Baker 2001), or more picturesquely the *periodic table of languages*. In particular, it represents the deepest known part of the hierarchy, which comprise the known parameters that are subservient to the “word first” value of the head directionality parameter. Presumably there are also parameters that are subservient to the “word last” value of the head directionality parameter and to the “yes” value of the

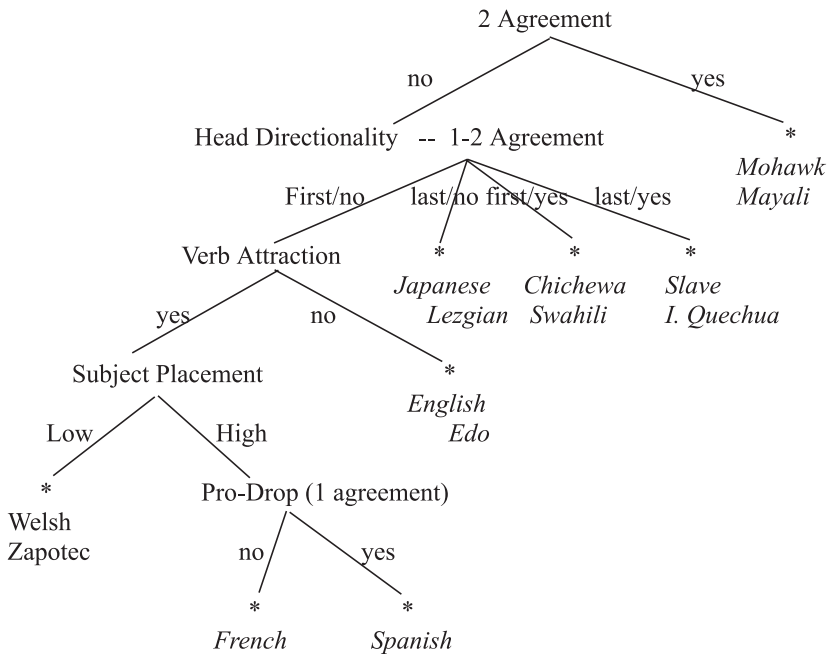


FIGURE 6 The parameter hierarchy (partial).

two-agreement parameter (see Baker, 2001, chap. 6, for some suggestions); but less is known about these linguistically and in terms of language acquisition, so I omit further discussion here.

Testing the Predictions: Initial Evidence

We have seen 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 is whether there is convergence between the typological studies I have reviewed and the 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 that of the grossly similar languages Spanish, Welsh, and 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 development for what I have called the two-agreement parameter has not been studied from this perspective, so far as I know. Once again, this is primarily for practical reasons: Languages such as Mohawk, for which the value of this parameter is crucial, happen to constitute relatively small or remote languages, with few children learning them; so, there is an important gap in our knowledge here.

Consider, then, the head directionality parameter, which is the highest ordered parameter that remains. It is well known that this parameter is learned quite early. The earliest rigorous studies of language acquisition—such as that by Bloom (1970) and Brown (1973) and the important crosslinguistic work reported in Slobin (1985)—showed that children learning different languages show systematic differences in word order. Indeed, the very first two-word utterances of children, appearing between 18 and 22 months, show that the head directionality parameter is already set. A typical early sentence from an English-speaking 1-year-old is *Read book* (Gia, age 20 months and 2 weeks) or *Eating cereal* (Eric, 20 months and 2 weeks; Bloom 1970, pp. 94, 108). Japanese 1-year-olds have comparable interests, but their grammar is already different: they begin with sentences like *Ame choodai yo* ‘candy give EPH’ (meaning ‘Give [me] candy’; Clancy, 1985; see also, Aksu-Koc & Slobin, 1985, on Turkish). The prediction that this parameter would 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 who were learning French and English, starting at around 21 months. They found that children at this age 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 the 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 include *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 that French infinitival verbs and English verbs do not (whether tensed or otherwise).

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 issue in an interesting way. They showed that subject placement in French and English is not decided until slightly later, at about 24 months. As a result, French and English children go through a stage in which their verbs are properly placed with respect to adverbs and negation but in which their subjects are not properly placed. Thus, on their second birthday, English children are wont to say sentences like the following:

- (41) a. No I see truck.
b. 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 do not 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 Example 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 prediction is correct (Hyams, 1989). It has long been known that children learning English or French omit subjects as much as 50% of the time; however, 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 such as Spanish and Italian. She has concluded that these children have not yet determined the setting of the pro-drop parameter. In the children she studied, the average age for establishing this parameter value was as predicted 27 months—somewhat later than the setting of the subject placement parameter.

These results can be summarized in Table 2. This table repeats the parameter hierarchy from Figure 6 and suggests when each parameter is learned, estimations gleaned from the rather small literature on acquisition that takes the notion of parameters seriously. Note that the times of acquisition do accurately mirror the logical relationships between the parameters.

I find this to be a seriously encouraging result, though certainly one that 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 coming from a sort of pilot inquiry, the results are striking enough to warrant further investigation along these lines.

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 linguistic map of the terrain, which children use in the process of language learning. There are still many parameters that need to be studied in this light. Furthermore, methodologies need to be developed for timing these stages of acquisi-

TABLE 2
Time of Parameter Acquisition

<i>Parameter Name</i>	<i>Order</i>	<i>Time of Acquisition</i>
2 agreement	1	Unknown
Head directionality	2	18–21 months
Verb attraction	3	22 months
Subject placement	4	24–25 months
Pro-drop (1 agreement)	5	27 months

tion accurately, since the pace of acquisition varies considerably from child to child and since many changes come quite rapidly between 22 and 30 months. But if the pattern holds under further scrutiny, it will serve as 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. It should, however, be an asset in this quest to know exactly what children are trying to learn and in what order, much as having a good road map will not get you to your destination but will assist you in deciding how to get there.

ACKNOWLEDGMENTS

The author thanks Susan Goldin-Meadow and Lila Gleitman for encouraging the writing of this article and for helping to see what it could be. Thanks are also extended to Jane Grimshaw and an anonymous reviewer for their comments on an earlier version, pushing me to clarify various points, discuss some related issues, and improve the article in various respects.

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

REFERENCES

- Abney, S. (1987). *The English noun phrase in its sentential aspect*. Unpublished doctoral dissertation, MIT, Cambridge, Massachusetts.
- Aksu-Koc, A., & Slobin, D. (1985). The acquisition of Turkish. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 1* (pp. 839–880). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Alexiadou, A., & Anagnostopoulou, E. (1998). Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory*, 16(3), 491–539.

- Baker, M. (1991). On some subject/object non-asymmetries in Mohawk. *Natural Language and Linguistic Theory*, 9(4), 537–576.
- Baker, M. (1996). *The polysynthesis parameter*. New York: Oxford University Press.
- Baker, M. (2001). *The atoms of language*. New York: Basic Books.
- Baker, M. (2002). Building and merging, not checking: The nonexistence of (aux) SVO languages. *Linguistic Inquiry*, 33(2), 321–328.
- Black, C. (1994). *Quiébolani Zapotec syntax*. Unpublished doctoral dissertation, University of California, Santa Cruz.
- Bloom, L. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT Press.
- Bresnan, J., & Mchombo, S. (1987). Topic, pronoun, and agreement in Chichewa. *Language*, 63(4), 741–782.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Carstens, V. (2002). Antisymmetry and word order in serial constructions. *Language*, 78(1), 3–50.
- Chomsky, N. (1970). Remarks on nominalization. In R. Jacobs & P. Rosenbaum (Eds.), *Readings in English transformational grammar* (pp. 184–221). Waltham, MA: Ginn.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, Netherlands: Foris.
- Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.
- Chung, S. (1976). An object creating rule in Bahasa Indonesian. *Linguistic Inquiry*, 7, 41–87.
- Cinque, G. (1994). On the evidence for partial N-movement in the romance DP. In G. Cinque, J. Koster, J.-Y. Pollock, L. Rizzi, & R. Zanuttini (Eds.), *Paths towards universal grammar: Studies in honor of Richard S. Kayne* (pp. 85–110). Washington, DC: Georgetown University Press.
- Clancy, P. (1985). The acquisition of Japanese. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition: Vol. 1* (pp. 373–524). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Croft, W. (1990). *Typology and universals*. Cambridge, MA: Cambridge University Press.
- Deprez, V., & Pierce, A. (1993). Negation and functional projections in early child grammar. *Linguistic Inquiry*, 24, 25–68.
- Dryer, M. (1992). The Greenbergian word order correlations. *Language*, 68(1), 81–138.
- Emonds, J. (1978). The verbal complex V'-V in French. *Linguistic Inquiry*, 9, 151–175.
- Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology*, 23(3), 331–392.
- Fodor, J. D. (1998). Unambiguous triggers. *Linguistic Inquiry*, 29(1), 1–37.
- Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry*, 25(3), 407–454.
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, 73(2), 135–176.
- Gleitman, L. R. (1994). Words, words, words. *Philosophical Transactions of the Royal Society of London, Series B*, 346(1315), 71–77.
- Greenberg, J. (1963). *Universals of language*. Cambridge, MA: MIT Press.
- Hale, K. (1983). Warlpiri and the grammar of nonconfigurational languages. *Natural Language and Linguistic Theory*, 1, 5–49.
- Hale, K., Krauss, M., Watahomigie, L., Yamamoto, A., Craig, C., Masayesva Jeanne, L., et al. (1992). Endangered languages. *Language*, 68(1), 1–42.
- Hawkins, J. (1990). A parsing theory of word order universals. *Linguistic Inquiry*, 21, 223–261.
- Hawkins, J. (1994). *A performance theory of order and constituency*. Cambridge, MA: Cambridge University Press.
- Hyams, N. (1989). The null subject parameter in language acquisition. In O. Jaeggli & K. Safir (Eds.), *The null subject parameter* (pp. 215–238). Dordrecht, Netherlands: Kluwer.
- Jaeggli, O., & Safir, K. (1989). The null subject parameter and parametric theory. In O. Jaeggli & K. Safir (Eds.), *The null subject parameter* (pp. 1–44). Dordrecht, Netherlands: Kluwer.
- Julien, M. (2002). *Syntactic heads and word formation*. New York: Oxford University Press.

- Kandybowicz, J., & Baker, M. (2003). On directionality and the structure of the verb phrase: Evidence from Nupe. *Syntax*, 6(2), 115–155.
- Kayne, R. (1995). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- King, G. (1993). *Modern Welsh: A comprehensive grammar*. London: Routledge.
- Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua*, 85, 211–258.
- Kuno, S. (1973). *The structure of the Japanese language*. Cambridge, MA: MIT Press.
- Lehmann, W. (1973). A structural principle of language and its implications. *Language*, 49, 47–66.
- Lounsbury, F. (1953). *Oneida verb morphology*. New Haven, CT: Yale University Press.
- Newmeyer, F. (1998). *Language form and language function*. Cambridge, MA: MIT Press.
- Nichols, J. (1986). Head-marking and dependent-marking grammar. *Language*, 62(1), 56–119.
- Pollock, J.-Y. (1989). Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry*, 20(3), 365–424.
- Postal, P. (1979). *Some syntactic rules of Mohawk*. New York: Garland.
- Rice, K. (1989). *A grammar of Slave*. Berlin: Mouton de Gruyter.
- Ritter, E. (1991). Two functional categories in noun phrases: Evidence from Modern Hebrew. In S. Rothstein (Ed.), *Syntax and Semantics: Vol. 26* (pp. 37–62). San Diego, CA: Academic Press.
- Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht, Netherlands: Foris.
- Slobin, D. (Ed.). (1985). *The crosslinguistic study of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Sproat, R. (1985). Welsh syntax and VSO structure. *Natural Language and Linguistic Theory*, 3, 173–216.
- Stowell, T. (1981). *Origins of phrase structure*. Unpublished doctoral dissertation, MIT, Cambridge, Massachusetts.
- Tomlin, R. (1986). *Basic word order: Functional principles*. London: Croom Helm.
- Williamson, J. (1984). *Studies in Lakhota grammar*. Unpublished doctoral dissertation, University of California, San Diego.