What You Can Say Without Syntax: A Hierarchy of Grammatical Complexity

Ray Jackendoff and Eva Wittenberg, Tufts University[1]

ABSTRACT. We propose a hierarchy of grammars that offer a range of resources for expressing meaning, all the way from the grammatically simplest languages to fully complex languages. The hierarchy can be understood in two ways: First, it provides a formal tool for describing communication systems. Second, it enables us to make an empirical claim about the human language faculty: it is a palimpsest, consisting of layers of different degrees of complexity, in which various grammatical phenomena fall into different layers. Several linguistic phenomena illustrate this claim.

[1] For invaluable comments on many iterations of this material, we are deeply indebted to the members of the Tufts linguistics research seminar, past and present: Neil Cohn, Ari Goldberg, Anita Peti-Stantić, Rob Truswell, Stephanie Gottwald, Chelsey Ott, Naomi Berlove, Geert Booij, and Rabia Ergin-Somer. We also wish to thank Jesse Snedeker and her lab for important discussion, and David Gil, Dan Everett, Annie Senghas, Mark Aronoff, Susan Goldin-Meadow, Dea Hunsicker, Marie Coppola, Tom Bever, and Gina Kuperberg for answering our many questions about issues in their work that we are addressing here. Thanks also to Daniel Ross, Lila Gleitman, David Gil, Ljiljana Progovac, Fritz Newmeyer, and Laurel Preston for comments on a near-final version of this chapter. We are grateful for the opportunity to present this material not only at the University of Washington workshop on which the present volume is based, but also at the workshop on grammatical complexity at the University of Konstanz (2012), the University of North Carolina's 2012 Linguistics Colloquium, and the Tufts Conference on Language and Representation 2012.

1. Introduction

We present here the preliminary results of a large-scale thought experiment: How much and what kind of thought can be expressed without resorting to the tools provided by fully complex syntax? What can semantic structure alone accomplish before syntax is needed to regiment the message? How much expressive complexity can be created with minimal tools?

These questions are more than just a mind game. We address them by developing a hierarchy of successively more complex grammars, beginning with trivial grammars that permit just one word per utterance, and culminating in fully complex grammars of natural languages, such as Dutch, Turkish, and Mohawk. The goal of the hierarchy is to provide a framework that accounts for differences in syntactic complexity across languages and language-like systems. But we find that it also can be used to describe certain subsystems of fully complex languages, stages of language acquisition, aspects of language processing, and, crucially, the interplay of language with other cognitive systems.

Our model focuses on the mapping from meaning to sound and on the range of messages that different grammatical systems can convey. We assume that any differences in intelligence are negligible for people across the planet. They think the same thoughts, no matter what kind of grammatical system they use; and they express the same kinds of thoughts, regardless of the grammatical tools they have: past, present and future events, cause-and-effect relationships, social relationships, hypothetical questions, and so forth. Yet languages obviously differ in terms of their syntactic complexity.

Our idea is that the simpler grammars in our hierarchy put more responsibility for comprehension on pragmatics and discourse context. For instance, to understand a child's one-word utterance, one needs to rely heavily on inferences about what the child might have in mind. As the child's grammar acquires more grammatical devices, it provides more resources for making complex thoughts explicit, reducing the workload on the hearer. One could say that the syntactic complexity of a maturing speaker is gradually easing the semantic and pragmatic burden on the hearer. For another example of rich semantic and poor syntactic structure, consider Riau Indonesian, as described by David Gil (this volume). If a Riau speaker says *chicken eat*, the hearer has to figure out from context that the intended message is 'someone is eating that chicken over there' as opposed to, say, 'a chicken will eat that' or the many other things it could mean. Similarly, pidgins and creoles have syntactically simple ways of encoding complex thoughts that place high demands on the hearer.

The idea that grammars come with different degrees of complexity is not new. What is new, however, is our formal treatment of the semantic domain, weaving together context, word semantics, and combinatorial structure in a principled way. Unlike previous grammatical hierarchies, such as the familiar Chomsky hierarchy (Chomsky, 1956), our goal is to create a model of increasingly complex sound-to-meaning mappings, not merely to generate sets of syntactically structured sentences. In our view, syntax is not the only generative linguistic system (Jackendoff, 2002), so we challenge the often unspoken assumption that knowledge of syntax is the most essential component in knowledge of a language.

One old problem that falls by the wayside in our approach is the question of when children start "having language." In first language acquisition, they move on from very simple to more complex language, and there is no need to make a binary decision about their having language or not at any particular point in development. Similarly, no bright line is crossed when a pidgin (often treated as "not a real language)" develops into a creole (which *is* considered a language). Our proposed hierarchy serves as a formal tool for describing a range of communication systems, allowing us to ask what natural languages and what subsystems of natural languages fall where on the hierarchy. Thus the hierarchy offers a way to classify systems such as a child's language or a pidgin, not just as either language or not language, but rather as falling somewhere along a continuum of formal possibilities.

Our hierarchy can also serve a second, and theoretically stronger, purpose. To the degree that various points on the hierarchy are found to be actually instantiated, we will be able to make an empirical claim about the human language faculty: it is not a monolithic block of knowledge, but rather a palimpsest, consisting of layers of different degrees of complexity, in which various grammatical phenomena fall into different layers (see for example Progovac, 2009, for a parallel idea).

A few comments should be made about the limits and problems in such an endeavor. First, this is a work in progress, and we draw from limited data. Any generalizations beyond the data we have analyzed are tentative and waiting to be disproven. Second, we are asking readers to throw aside commonly held beliefs about the language system and follow us on a rough ride through very diverse areas of language. In particular, for the sake of argument, we are adopting the unfamiliar and sometimes painful methodology of assuming as little syntactic structure as possible. We hope to show that there is a lot to be gained, and not much lost, in adopting this methodology, and that in doing so, a wide variety of phenomena can be tied together.

Third, we should define what we mean by "grammar", "semantics", and "syntax" here. In our terms, "grammar" is the system that relates sound to meaning. It can include syntactic tools, but it need not: less complex grammars may be direct mappings of sound to meaning, without syntactic intervention. Semantics includes all the properties that have to do with meaning, linguistic and nonlinguistic, unless otherwise specified.[2] This definition has a large scope: among other things, it includes lexical semantics, information structural properties in a sentence such as Topic and Focus, and very general semantic concepts such as Object, Action, Agent, and Patient. Syntax, by contrast, is one step removed from meaning: it is a formal system that abstracts away from semantic categories, that labels constituents in terms of abstract categories such as nouns and verbs, and that imposes formal structure on them. Often, semantic and syntactic categories intersect. For

[2] In this article, we use the terms "semantics" and "pragmatics" rather informally: semantics, unless otherwise indicated, refers to meaning as a whole; pragmatics, to the aspects of meaning contributed by non-linguistic context and world knowledge.

instance, animate objects (a semantic category) are always nouns (a syntactic category). However, although actions (a semantic category) are usually verbs (a syntactic category), there are also action nouns such as *action* and *party*. A central point in our approach is that syntactic categories and rules are not automatically and inevitably needed in every linguistic system; instead of assuming their presence, our null hypothesis is that they are absent, and we check this hypothesis by performing standard tests for syntactic categories and phrases.

Since we are not assuming that syntax necessarily matches semantics, an important part of the grammar in our model is the system of *interface rules* that mediate between semantics, syntax, and phonology. As we will see, in these less complex grammars, certain aspects of the work traditionally assigned to syntax can be stated instead in terms of the way semantics maps into linguistic expression. (For more detailed exposition of our notion of grammar, see Jackendoff, 2002, Culicover and Jackendoff, 2005.)

Finally, we should emphasize that we are not proposing a global measure of linguistic complexity, such that one could decide, say, whether Warlpiri is more or less complex than ASL. We are inclined to think that there are many dimensions along which languages can be complex or not, and we are not sure that it is important to compare across dimensions. To take about the simplest possible case: What is more complex, a language with forty phonemes and two morphological cases, or a language with twenty phonemes and six cases? How would one decide? As Daniel Dennett puts it, speaking of biological complexity, “There is no single summit in Design Space, nor a single staircase or ladder with calibrated steps, so we cannot expect to find a scale for comparing amounts of design work across distant developing branches" (Dennett, 1995, 134). Here we are exploring just one of those staircases, and in particular, steps rather near the bottom.

2. The hierarchy

This section introduces the hierarchy of grammatical complexity. On its own it is rather abstract. Its interest for us lies in what meanings can be expressed by the grammars that it defines, as will be seen in subsequent sections.

Logically, the simplest conceivable grammar is a **one-word grammar**, in which utterances are restricted to a single word, as in (1a). The next simplest is a **two-word grammar**, which allows utterances of either one or two words, as in (1b).[3] From there, it is a short way to a **concatenation grammar** (1c), which allows utterances to consist of word strings of arbitrary length. We state the grammar in the form of templates for (or constraints on) possible structures.

[3] Readers may wonder why we have singled out a two-word grammar as significant, but not, say, a three- or four-word grammar. From an empirical point of view, various phenomena we will be examining can be characterized in terms of two-word combinations. And from a theoretical point of view, moving from single words to pairs requires the speaker to manage not just the introduction of an additional word but also a semantic relation between them (see section 3.2). We speculate that this new semantic relation is the real source of complexity in two-word utterances.

(1) a. **One-word grammar** $[_{\text{Utterance}}$ Word] [*Traditional notation*: Utterance → Word]

b. **Two-word grammar** $[_{\text{Utterance}}$ Word (Word)] [Utterance → Word (Word)]

c. **Concatenation grammar** $[_{\text{Utterance}}$ Word*] [Utterance → Word*]

Notice that none of these grammars involves parts of speech or morphology. The relation of utterances to interpretations has to be conditioned by semantic distinctions, for example object vs. action, rather than by syntax, such as noun vs. verb. In other words, there is no real syntax in the usual sense, as defined by syntactic categories and syntactic markers of these categories.

The next elaboration, a **simple phrase grammar**, allows words to be grouped into phrases. An utterance consists of a concatenation of words and phrases, as in (2). But crucially, phrases cannot contain phrases, so embedding is limited to a depth of two nodes. Such a grammar requires two phrase structure rules, one for Utterances (2a) and one for Phrases (2b). There are in turn two variants of the rule for Phrases: one in which a Phrase consists only of two words (2b.i), and one in which it may consist of an unlimited number of words (2b.ii).[4]

(2) **Simple phrase grammar**

a. $[_{\text{Utterance}}$ Word/Phrase*] [Utterance → Word/Phrase*]

b. i. $[_{\text{Phrase}}$ Word Word] (2-word phrase) [Phrase → Word Word]

[4] The notation *Word/Phrase** is meant to denote an unlimited concatenation whose elements are either Words or Phrases.

ii. [$_{\text{Phrase}}$ Word*] (unlimited phrase) [Phrase → Word*]

One way to introduce phrases is by prosody, resulting in a **prosodic simple phrase grammar**. In such a grammar, prosody indicates which words belong together semantically. For example, even without syntactic categories, a child's utterance like *kiss, Julius Daddy* might mean that one should kiss Julius's daddy, while *kiss Julius, Daddy* is more likely about Daddy kissing Julius himself.

Alternatively, at this point in the hierarchy it starts to become useful to introduce parts of speech (or syntactic categories) to label words and phrases, yielding a **part-of-speech simple phrase grammar**. In such a grammar, different categories of phrases may specify different categories and orders for the words they contain, and they may be marked differently: the word *kiss* would be translated into German as a verb with umlaut and a verbal suffix (*küssen*), but as a noun without the umlaut (*Kuss*).

In the next elaboration, *phrases* can be grouped into phrases, so that there is now the potential for recursion. (3) shows the general rules for a **recursive phrase** grammar: both Utterances (3a) and Phrases (3b) can consist of any combination of Words and Phrases.[5]

---

[5] Note that one-word, two-word, and concatenation grammars are among the simplest possible finite-state grammars, and simple and recursive phrase grammars are among the simplest possible context-free grammars. In our terms, the Minimalist Program's binary Merge (Chomsky, 1995) amounts to a putatively universal recursive phrase grammar in which each nonterminal node has exactly two constituents.

(3) **Recursive phrase grammar**

a. [$_{\text{Utterance}}$ Word/Phrase*] [Utterance → Word/Phrase*]

b. [$_{\text{Phrase}}$ Word/Phrase*] [Phrase → Word/Phrase*]

A different kind of elaboration of grammars involves the possibility of making words themselves composite, that is, compositional morphology. (4) shows the prototypes for two kinds of morphological rules.

(4) **Compounding:** [$_{\text{Word}}$ Word Word]

Affixal morphology: [$_{\text{Word}}$ {Word/Stem, Affix}] (either order)

For fully complex languages, we need to add further elaborations: functional categories (e.g. Det, Aux), principles governing long-distance dependencies, local binding of anaphors, perhaps grammatical functions, and so on.

As motivation for proposing these sorts of simple grammars, we note that a number of phenomena to be discussed here are described as having no subordination, no functional categories, little or no morphology, and semantically-driven principles of word order. In the proposed hierarchy, they all appear to be versions of concatenation grammar or simple phrase grammar.

3. Interface rules

3.1 Interface rules for one-word grammars

The classes of formal grammars proposed above, of course, are useful only when they are coupled with principles that state how overt utterances are linked to meanings – *interface rules.* There are two kinds of interface rules. The first kind is, simply, *words*. A word connects a piece of phonology to a piece of meaning. If it is used in a grammar that has syntactic features, a word also carries relevant syntactic features that allow it to connect with the syntactic structure. For example, the word *cat* is a linkage in long-term memory of a piece of phonological structure, some syntactic features, and some semantic features, as in (5).

(5) Phonology: $/\text{kæt}/_1$

Syntax: $[+N, -V, \text{count}]_1$

Semantics: $[\text{FELINE, PET, …}]_1$

The subscripts indicate that these pieces of structure are connected and remain so when this unit forms part of a larger expression. Thus the word serves as part of the interface between phonology, syntax, and semantics.

The second kind of interface rule is a *combinatorial interface rule*, which specifies how the meanings of the parts of a grammatical constituent C are combined into the meaning of C. A simple case is adjectival modification of a noun: The adjective denotes a property of the entity denoted by the head noun. We will be dealing here mainly with this kind of interface rule.

An important generalization we have discovered is that the interface rules for lower levels of the hierarchy scale up to the upper levels. That is, they apply whether the constituent C is an Utterance, a Phrase, or a Word, whether its parts are Phrases, Words, or sub-word morphemes, and whether Phrases and Utterances are defined only prosodically or in terms of syntactic categories. This means that as one moves up the hierarchy, principles from lower points are not obliterated; rather they are elaborated. We will illustrate this point as we go along.

A one-word grammar (1a) presents the very simplest interface rule, (6): the meaning of the word equals the meaning of the utterance. In this rule and those to follow, the subscripts indicate links between phonology, syntax, and semantics. Here, the phonological Word, with subscript 1, is linked to the meaning with subscript 1, but this meaning is also linked to the entire Utterance.

(6) **Word = Utterance**

Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1]_2$

Semantics: $X_{1,2}$

Since utterances in this grammar contain only a single word, there is in fact no real distinction between words and utterances. So an even simpler form of this rule is the trivial (7).

(7) Phonology: Utterance_1

Semantics: X_1

(7) is the only interface rule necessary for most primate calls, which are essentially monadic utterances. What makes (7) more interesting is that it also accounts for the use of certain words of fully complex languages, words like *hello*, *ouch*, *upsey-daisy*, and *abracadabra*. Such words have no syntactic properties. Rather, they serve as full utterances on their own, and they appear in combination only in paratactic situations like *Hello, Bill* and quotative contexts like *"Hello," she said*, in which anything can be inserted, even a phrase from another language.

However, a one-word grammar need not be quite this limited. We can add a little more expressivity by allowing the utterance to paste some content around the meaning of the word, as shown in (8).

(8) **One-word grammar with pragmatics**

Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1\ ]_2$

Semantics: $[F\ (X_1)]_2$

Here, the function *F* in the semantics is present in the meaning, but it has no overt counterpart in the utterance. (8) appears to characterize the one-word stage of child language. For instance, *doggie* can be used to mean 'there's the doggie,' 'where's the doggie,' 'that looks like a doggie,' 'that belongs to the doggie,' 'I want the doggie,'

‘doggie, pay attention to me,’ and so on, where each of these extra pieces of meaning is a pragmatically determined function *F*.

Pragmatic enrichment as formalized in (8) is not confined to one-word grammars. As promised above, it scales up to syntactically more complex languages, which make use of a more general version of rule (8) that we call the **enrichment schema**. It has two cases, one for one-phrase utterances (9), and one for coercions (11). (9) represents the general form of principles for interpreting one-phrase utterances, for example those in (10).

(9) **Enrichment schema:** One-phrase utterances

Syntax: $[_{\text{Utterance}}\ \text{Phrase}_1]_2$

Semantics: $[\text{F}\ (\text{X}_1)]_2$

(10) An eagle!!

Some scotch?

[What kind of pizza do you want?] – Pepperoni.

(11) represents the general form of coercions, such as those in (12). The underlined parts of the paraphrases in (12) spell out the unexpressed function *F*.

(11) **Enrichment schema:** Coercion

Syntax: $[_{\text{Phrase}}\ \text{Phrase}_1\ ]_2$ *or* $\text{Phrase}_{1,2}$

Semantics: $[F(X_1)]_2$

(12) Plato [= ‘book by Plato’] is on the top shelf, next to Chomsky.

The ham sandwich [= ‘the person with ham sandwich’] wants more coffee.

I’m [= ‘my car’s’] parked out back.

(Sluicing) Joe ate something, but I don’t know what [= ‘what he ate’].

Notice that (8), (9) and (11) are identical except for the labels on the syntactic constituents. The basic principle for each case is that a constituent has extra added meaning beyond the meanings of its words. The literal meaning of the constituent, determined just by the meanings of the words, is semantically embedded in a larger, pragmatically determined meaning. Hence (8), the interface principle for one-word grammars, scales up to situations in more complex languages.

3.2 Basic interface rules for 2-word grammars

Next, let us amplify the grammar to two constituents that together form an utterance. Now the problem arises of how to combine their meanings to form the meaning of the whole. One thing that makes two-word utterances harder than one-word utterances is that it is necessary to establish an unspoken semantic relation between them. The simplest way to do this is just to say that the two meanings combine somehow or another. But this results in what Anderson (2004) calls “semantic soup,” in which there is far too much indeterminacy for the hearer.[6] More discipline is necessary. To ground our solution intuitively, we illustrate not with actual two-word grammars, but with a subsystem of English that approximates a two-word grammar: compounds. In English compounds, the semantics can map directly into phonology: a sequence of two words.[7] (We discuss children's two-word grammars in a moment.)

There are at least three ways the constituents of a compound can combine semantically. First, one component can be a semantic argument of the other, as in (13). Examples are shown in (14).

(13) **Function-argument schema**

Phonology: $[_{Word}\ \text{Word}_1\ \text{Word}_2\ ]_3$

[6] Gil's Association Operator (this volume) has this effect, in that the meaning of a constituent [X Y] is simply 'entity associated with X and Y.' A real example of "semantic soup" can be found in the utterances of sign-trained apes, for example Nim Chimpsky's *give orange me give eat orange me eat orange give me eat orange give me you* (Seidenberg and Petitto, 1978, Terrace, 1979). Gil's notion of Predication corresponds to our Function-argument schema ((13) below); his Attribution corresponds to our Modification schema (15).

[7] Alternatively, the semantics might map into morphosyntax. However, English compounding is not particularly restrictive about the syntactic categories of the compound's constituents, as can be seen from examples like (i); and though compound nouns are predominant, compound verbs and adjectives are also possible, as seen in (ii).

(i) long_A bow_N, $\text{under}_P\text{current}_N$, $\text{pull}_V\text{over}_P$, $\text{over}_P\text{kill}_N$, $\text{speak}_V\text{easy}_A$, $\text{hear}_V\text{say}_V$, $\text{once}_?\text{over}_P$

(ii) $[\text{skin}_N\ \text{deep}_A]_A$, $[\text{worm}_N\ \text{eaten}_V]_A$, $[\text{blue}_A\ \text{green}_A]_A$, $[\text{house}_N\ \text{sit}_V]_V$

Semantics: $[F_2 (X_1)]_3$

(exchange subscripts 1 & 2 for left-headed compounding)

(14) union member [= 'member of a union']

helicopter attack [= 'attack by helicopter']

Notice that the interface rule determines the word order, in that the semantic head, subscripted 2, corresponds to the second word (in English). Thus, syntactic structure is not needed to specify headedness; word order alone is enough. This illustrates a more general point of our approach mentioned in section 1: not all "grammatical" effects arise from the syntax. Many are a consequence of the interface rules.

A second way that two word meanings can combine is for one of the constituents to be a modifier of the other, as in (15), where the semicolon symbolizes the semantic relation of Y modifying X. Examples are shown in (16).

(15) **Modification schema**

Phonology: $[_{Word}\ Word_1\ Word_2\ ]_3$

Semantics: $[X_2; Y_1]_3$

(16) blackbird [= 'bird that is black']

plastic container [= 'container that is plastic']

Often the interpretation of a compound involves enrichment in the sense of rule (11), in addition to modification. For instance, the constituents of the compounds in (17) are enriched by the underlined material before being combined by the modification schema.

(17) snowman [= ‘simulation of man that is made of snow’]

garbage man [= ‘man who takes away garbage’]

A third way for two words to combine semantically is for them both to be arguments of an unexpressed function. (18) is the rule, and (19) gives an example.

(18) **Co-argument schema**

Phonology: $[_{\text{Word}}\ \text{Word}_1\ \text{Word}_2\ ]_3$

Semantics: $[\text{F}\ (\text{X}_1, \text{Y}_2)]_3$

(19) seahorse [= ‘something that looks like a horse and that lives in the sea’]

Notice that the form of phonology in these three schemas is the same; what differs is how it maps into semantics.

In two-word child language, we find the same interface principles at work, except that instead of mapping the semantics to a compound word consisting of a pair of words, they map the semantics to a complete utterance consisting of a pair of words, as in (20). (21) gives an example of each of the principles.

(20) Phonology: $[_{Utterance}\ Word_1\ Word_2]_3$

(21) Function-argument schema: *Mommy fix* [= 'Mommy should fix it']

Modification schema + enrichment: *Big house* [= 'that's a big house']

Co-argument schema: *Mommy pumpkin* [= 'Mommy cuts the pumpkin']

Roger Brown (1973), quoting Lois Bloom (1970), cites two instances of the same child uttering *mommy sock*. On one occasion it apparently is intended to mean 'Mommy's sock,' a case of modification. On another occasion it is intended to mean 'Mommy's putting a sock on me,' a case of co-arguments.

Martin Braine (1963) observed that many (but not all) children enter the 2-word stage with 'pivot schemas': a small number of words that occur in combination with a larger variety of other words, as illustrated in (22). This observation has been revived in recent years (Tomasello, 2003).

(22) see baby, see pretty, etc.

more car, more cereal, more cookie, more fish, etc.

no bed, no down, no fix, etc.

In our approach, each pivot can be regarded as an interface rule, for example (23), which is a special case of the function-argument schema.[8]

(23) Phonology: $[_{\text{Utterance}}\ [_{\text{Word}}\ \text{see}]\ \langle \text{Word}_i \rangle\ ]_j$

Semantics: $[\text{SEE (ME/YOU, } (\text{Y}_i)\ ]_j$

Like the enrichment schema, the interface rules in this section types apply also to phenomena in more complex languages. Compounding is not the only instantiation. The function-argument schema scales up to the familiar principles for integrating a syntactic head with its complements, and the modification schema scales up to the usual rules for integrating syntactic adjuncts with their heads. The co-argument schema scales up to, among other things, the English casual paratactic conditional (24), where there is no overt connective, and small clause constructions (25), which have an implicit copula that links the NP and the predicate (cf. Progovac, this volume).

(24) You shoot a cop, you go to jail.

(25) a. Everyone out of the car!

b. John at a baseball game?! (I can't believe it!)

c. [John at a baseball game] is hard to imagine.

d. No dogs allowed.

e. Refreshments in the kitchen.

[8] Thus we need not speak of "pivot *grammar*," as Braine did, implying that the child's language is made up entirely of pivot schemas. Rather, pivot schemas are just one possible component of the child's language.

Pivot schemas too have counterparts in complex languages. (26a) appends a free choice of name to a title to form an NP; (26b,c) combine a pivot with a free choice of phrase to form a full root utterance that cannot embed.

(26) a. *Forms of address*: Mr. X, Ms. Y, Governor Z, Rabbi W, etc.

b. *Directed epithets*: Fuck/Damn/The hell with/Hooray for NP!

c. *Inquiries*: How about XP? What about XP?

3.3 Correlating word order with thematic roles

The interface rules we have just described still leave the interpretations of many configurations indeterminate, in danger of ending up with uninterpretable "semantic soup." For instance, if we encounter the utterance *chicken eat*, we may infer that it is to be interpreted by the function-argument schema, with *chicken* as the argument of *eat*. But *which* argument? Is the chicken eating or being eaten? Similarly, the utterance *cow horse* might be interpreted by the co-argument schema to denote an action involving a cow and a horse. But which one is acting on which? The indeterminacy becomes even greater when we move to a concatenation grammar, which allows more than two words. For instance, does *cow big horse* mean that the cow is big, or the horse?

One widespread strategy to ameliorate this problem is well-documented both in the languages of the world and in the various less complex phenomena we have been looking at: the preference for Agents to precede Patients, so that the cow is interpreted as doing something to the horse, not vice versa (see Gil, 2005, for a similar discussion). This preference can be stated as a simple interface rule (27), which relates thematic roles to linear order, without any reference to further syntactic structure. In a two-word grammar, the syntactic part of the rule is (27a). By allowing the two words in question to be non-adjacent, the rule generalizes to a concatenation grammar, as in (27b). Notice that the semantic function *F* that assigns the thematic roles has no index in this particular rule, so the rule does not depend on the exact relation between Agent and Patient to be expressed. (The ellipses in the semantics allow for further unexpressed semantic information as well.)

(27) **Agent > Patient** (a special case of the Co-Argument schema)

a. *Version for two-word grammar*

Phonology/syntax: $[_{\text{Utterance}}\ \text{Word}_1\ \text{Word}_2]_3$

Semantics: $[\text{F (Agent: X}_1\text{, Patient: Y}_2\text{ , …)}]_3$

b. *Version for concatenation grammar*

Phonology/syntax: $[_{\text{Utterance}}\ \ldots\text{Word}_1\ \ldots\text{Word}_2\ldots]_3$

Semantics: $[\text{F (Agent: X}_1\text{, Patient: Y}_2\text{, …)}]_3$

Another widespread strategy is for Agents to precede Actions. The formalization of this strategy for the case of concatenation grammars is stated in (28). It is of course a prototype for subject-verb order in more complex languages.

(28) **Agent > Action** (special case of the Function-argument schema)

Phonology/syntax: $[_{Utterance} \ldots Word_1 \ldots Word_2 \ldots]_3$

Semantics: $[F_2 \text{ (Agent: } X_1, \ldots)]_3$

Actions may precede Patients, as in (29), or vice versa, as in (30). (29) is the prototype for VO order; (30) for OV order. The choice between them is arbitrary, but whichever way the order is conventionalized in a particular system, a relatively stable order makes communication of meaning more reliable and less dependent on context.

(29) **Action > Patient** (Prototype for VO order)

Phonology/syntax: $[_{Utterance} \ldots Word_1 \ldots Word_2 \ldots]_3$

Semantics: $[F_1 (\ldots, \text{Patient: } X_2, \ldots]_3$

(30) **Patient > Action** (Prototype for OV order)

Phonology/syntax: $[_{Utterance} \ldots Word_1 \ldots Word_2 \ldots]_3$

Semantics: $[F_2 (\ldots, \text{Patient: } X_1, \ldots]_3$

Another strategy correlates linear order with information structure. Most commonly, topic comes first and focus last, as in (31) and (32).

(31) **Topic first**

Phonology/syntax: $[_{Utterance}\ Word_1 \ldots]_2$

Information structure: $Topic_1$

(32) **Focus last**

Phonology/syntax: $[_{\text{Utterance}} \ldots \text{Word}_1]_2$

Information structure: Focus_1

The possible strategies in (27)-(32) lead to the possible interpretations of *chicken eat* and *eat chicken* shown in (33), depending on which principles the system in question makes use of.

(33) a. *Chicken eat* = 'chicken is eating' by Agent > Action

Chicken eat = 'someone is eating chicken' by Patient > Action (if the language has it)

Chicken eat = 'someone is eating chicken' by Topic First (if the language has it)

b. *Eat chicken* = 'someone is eating chicken' by Action > Patient (if the language has it)

Eat chicken = 'chicken is eating' by Focus Last (if the language has it)

Like the function-argument, modification, and co-argument schemas of the previous section, the interface principles in (27)-(32) map directly between a string of phonological words and a meaning. They invoke only linear order and semantic distinctions such as object vs. action, argument vs. modifier, agent vs. patient, and topic vs. focus. They show

how a fairly expressive language could be constituted without syntactic categories and even without phrase structure.[9]

At this point, one might argue that the correlation of word order with thematic roles is nothing but a "perceptual strategy" or a "habit"; it is not *real* language, which requires true syntax (Fodor, Bever, and Garrett 1974, Townsend and Bever, 2001). We find this distinction questionable. Visually perceiving an event's participants in a certain order or with a certain prominence might be a perceptual strategy (Hafri, Papafragou, and Trueswell, in press). But encoding the thematic roles in an utterance in terms of word order is a distinctly linguistic principle: it is part of the mapping between meaning and sound. In other words, even if principles like (27)-(32) might be grounded in more general cognition, they are thereby no less linguistic.

As mentioned above, Agent > Action, Agent > Patient, Action > Patient, and Patient > Action are models for SVO and SOV order in more complex languages. Moreover, Topic is often marked by initial position, while Focus is often marked by final or near-final position. So again the rudimentary interface principles for concatenation

[9] Jackendoff and Wittenberg (in preparation) discuss two further important interface principles. The first, which applies in concatenation grammars and upward, is that a semantic constituent preferably corresponds to a contiguous string in phonology. This might be considered a precursor of syntactic phrase structure. The second, which appears in prosodic simple phrase grammars and upward, is that a prosodic phrase generally corresponds to a semantic constituent; a suitably nuanced version of this principle appears in more complex languages as well – though it is usually stated in terms of a correspondence between phonological and *syntactic* constituency (e.g Selkirk, 1984, 2000, Jackendoff, 1987, 2007).

grammars scale up to widespread principles in syntactically complex languages. The simpler principles don't disappear as we move up the hierarchy.

4. Illustrations of the hierarchy

4.1 Some phenomena that involve lower levels of the hierarchy

We now briefly mention some phenomena that appear to instantiate lower levels of the hierarchy (see Jackendoff and Wittenberg, in preparation, for more detailed discussion).

Pidgins and creoles. Pidgins are often described (e.g. Bickerton, 1981, 2008, Givón, 2009; see also Gil, this volume) as having no subordination, no morphology, no functional categories, free omission of arguments, and unstable word order governed by semantic/pragmatic principles such as Agent First and Focus Last. Our outlook leads us to ask: Is there any evidence in pidgins for parts of speech? Is there any evidence for phrasal categories such as NP? Is there any phrase structure at all aside from prosody? If not, pidgins would be classified as concatenation grammars – or if prosody is doing some work, they could be prosodic simple phrase grammars. Creoles, of course, do add many features of more complex languages, such as more conventionalized word order, functional categories, and syntactic subordination. In our terms, the transition from pidgin to creole is not from non-language to language, but rather one or more steps up the grammatical hierarchy.[10]

[10] We recognize, of course, that there is considerable variation among pidgins and creoles. A description in our terms requires considerable care, differentiating one case from the next.

Late second language acquisition. Klein and Perdue (1997) describe a distinct stage in late second language acquisition that they call the Basic Variety. In their multi-language longitudinal study, they found that all learners achieved this stage; many speakers went beyond this stage, but many did not. The Basic Variety is described as lacking inflectional morphology and sentential subordination, and freely omitting arguments. It has simple, largely semantically based principles of word order including Agent First and Focus Last. (The relative position of Patient and Action seems to be determined by the target language.) Our framework leads us to ask the same questions for Basic Variety as for pidgins: Is there any evidence for parts of speech or phrasal categories? And again, is there any phrase structure aside from prosody? We conjecture that the Basic Variety too is either a concatenation grammar or, if prosody is doing any work, a simple phrase grammar.

Home sign. A third case of a language with less complex syntax is home sign, the languages invented by deaf children who have no exposure to a signed language. These have been studied extensively by Susan Goldin-Meadow and colleagues for decades (e.g. Feldman, Goldin-Meadow, and Gleitman, 1978, Goldin-Meadow, 2003). On their account, they have at most rudimentary morphology and freely omit arguments. There are morphological differences between nouns and verbs (but in our terms, the distinction may in fact be between objects and actions). There appears to be no use of prosody to delineate semantic constituency. Homesigners do produce some sentences with multiple verbs (or action words), which Goldin-Meadow describes in terms of syntactic embedding. We conjecture that these are actually rudimentary serial verb or serial action-word constructions, which need not involve embedding.[11] Our current diagnosis of home signs is that they have only a semantic distinction of object vs. action, not a syntactic distinction of noun vs. verb, and that they have a concatenation grammar with a possible admixture of a small amount of semantically-based morphology.

Al-Sayyid Bedouin Sign Language. This is a language emerging in a Bedouin tribe in Israel with three generations of hereditary deafness (Sandler, Meir, Padden, and Aronoff, 2005; Aronoff, Meir, Padden, and Sandler, 2008). Based on the very small amount of published data, the language of first generation signers looks like a one-word grammar with a slight admixture of two-word grammar. This form of the language places an extremely heavy reliance on context and pragmatics for understanding. The language of older second generation signers looks like a concatenation grammar, with little consistent use of prosody to delimit semantic constituents. The language of younger second generation signers looks like a simple phrase grammar in which prosodic constituency plays a grammatical role. There is still no morphology, and no evidence for parts of speech. In other words, as the language has developed, it has gradually climbed up the hierarchy. From what we can tell from the literature on Nicaraguan Sign Language (Kegl, Senghas, and Coppola, 1999; Senghas, 2003), it has undergone a similar development, though evidently climbing up the hierarchy a good deal faster.

[11] Hunsicker and Goldin-Meadow (2012) propose that one homesigner they have analyzed uses a syntactic NP constituent consisting of a demonstrative that functions as a determiner plus a content word that functions as head noun. Jackendoff and Wittenberg (in preparation) show how their evidence can be accounted for in terms of a concatenation grammar.

Processing strategies. The use of rules like (27-32) is not confined to emerging languages. Townsend and Bever (2001) discuss what they call semantically based "interpretive strategies" or "habits" that influence language comprehension and lead to garden path situations. In particular, hearers tend to rely on semantically based principles of word order such as Agent > Patient, which is why they have more difficulty with constructions such as reversible passives and object relatives. Similarly, Ferreira and Patson (2007) discuss what they call "shallow parsing" or "good-enough parsing" in sentences like *Bill knows the answer is wrong*: subjects in their experiments apparently rely on linear order and semantic plausibility rather than syntactic structure. As is well known, similar symptoms appear in language comprehension by Broca's aphasics (Caramazza and Zurif 1976). Dick et al. (2001) show that college students behave like Broca's aphasics when trying to understand sentences that have been low-pass filtered and sped up. As mentioned above, even though the literature tends to describe these so-called "strategies" or "heuristics" as something separate from language, they are still mappings from phonology to meaning – just simpler ones. Building in part on evidence and analysis in Kuperberg (2007), we propose that the language processor makes use both of rules of full syntactic grammar and of rules of simpler concatenation grammar. When the two types of rules produce conflicting analyses, interpretation is more difficult. And when syntactic rules break down under conditions of stress or disability, the concatenation grammar is still at work.

4.2 Two syntactically less complex languages

We now offer a brief synopsis of our work (Jackendoff and Wittenberg in preparation) on two languages whose full grammar appears not to make use of the entire hierarchy. One is **Riau Indonesian,** a vernacular with several million speakers, described in detail by David Gil (Gil 2005, 2009, this volume, and many other publications). Gil argues that this language presents no evidence for syntactic parts of speech. There is a small number (less than 20) of affixes which are completely unselective as to what they attach to. There is no inflectional morphology. Arguments can be freely omitted. There is a small number (less than 30) of closed-class items which require complements; however, they are completely unselective except on semantic grounds (some select actions, some objects). The only evidence for constituent structure comes from prosodic phrasing. The effects expressed by syntactic subordination in English are expressed in this language through syntactic parataxis plus pragmatic enrichment (e.g. the enrichment schema in (9) and (11) and a version of the co-argument schema in (18)). For instance, conditionals are expressed only by a counterpart of the English paratactic conditional, e.g. *You shoot a cop, you go to jail* (24).

The word order in this language is substantially free, but heads tend to precede modifiers, agents tend to precede actions, and actions tend to precede patients. In addition, there are information structure influences, in particular a tendency for topics to appear at the beginning of a sentence. The only rigid word order is imposed by the closed class items, which require their complements on their immediate right. The syntactic and pragmatic freedom of this language can be illustrated by Gil's example (34).

(34) *ayam makan*, 'chicken eat' can mean

a. '{a/the} chicken(s) {is/are eating/ate/will eat} {something/it}'

b. '{something/I/you/he/she/they} {is/are eating/ate/will eat} {a/the} chicken'

c. '{a/the} chicken that {is/was} eating'

d. '{a/the} chicken that {is/was} being eaten'

e. 'someone is eating with/for the chicken'

f. 'where/when the chicken is eating'

Our analysis is that the language is basically a simple phrase grammar whose constituency is determined by prosody, with a small amount of morphology. The details of our analysis are presented in Jackendoff and Wittenberg (in preparation) and are beyond the scope of this paper. But basically, Riau Indonesian is a language that is syntactically simple in our sense.

The second example of a syntactically simple language is the controversial case of **Pirahã**, studied extensively by Dan Everett (Everett, 2005, 2009). This language has exuberant suffixal morphology as well as compounding, so it is far from simple in this respect. Everett's analyses, in all their various incarnations, assume a noun-verb distinction, so the parts of speech may be at least noun, verb, and everything else. But there are no definite or indefinite articles, no markers of plurality, and no inflectional morphology. Like many of the cases we have looked at, arguments can be freely omitted, though there is also the possibility of pronominal clitics on the verb. There is fairly fixed SOV word order, with the possibility of postposing the subject and/or the object. There is an interesting restriction, namely that nouns appear to admit only one modifier, either a prenominal possessive or a postnominal adjectival or quantifier, but not both. We take this as tentative evidence for an NP constituent that is constrained to at most two words.

Everett's most notorious claim is that Pirahã lacks recursion. His evidence is that all the constructions expressed by recursive syntax in English are either paratactic or require circumlocution. In a critique of Everett (2005), Nevins, Pesetsky, and Rodrigues (2009) show that there may be multiple clauses in a sentence. However, they do *not* show that clauses can contain clauses that contain clauses, that is, true recursion. And in fact the multiple clauses they cite are arguably paratactic; so what can be shown at most is that Pirahã has a simple phrase grammar, with the only further depth of embedding being the two-word NP, essentially the same point in the hierarchy as Everett claims. In his response to the critique (Everett, 2009), Everett makes the more radical claim that these clauses are separate sentences, so the language can be described as a concatenation grammar, possibly with NPs. So whichever analysis we adopt, Pirahã looks like a syntactically simple language, though not as simple as Riau Indonesian, in that it has syntactic categories and rich morphology.

5. Implications for the evolution of language

We are reluctant to get involved in debates on the evolution of the human language capacity, because there is no evidence for who was saying what when among our hominid ancestors, not to mention among our cousins the Neanderthals. However, the grammatical hierarchy developed here offers the possibility of speculatively reverse-engineering evolution: it offers possible language-like systems that are more complex than primate calls but less complex than modern full languages. We find it plausible that such systems existed in earlier hominids, and that further evolutionary steps resulted in brain structures that permitted the more complex languages that humans speak today. (Of course, finding something plausible is the weakest of arguments, but unfortunately that is for the most part the best one can do in debates on evolution of language.)

This is essentially the idea behind Derek Bickerton's (1990, 2008) "protolanguage" hypothesis, further elaborated by Jackendoff (1999, 2002), Hurford (2011), and Progovac (this volume), among others. Bickerton's basic idea is that for a million or more years before modern language arose, hominids spoke something rather like today's pidgins. In fact, his notion of protolanguage lines up rather well with our notion of a concatenation grammar. Furthermore, the grammatical hierarchy offers a (relatively) graceful way for protolanguage itself to come into existence piecemeal, as well as for it to evolve piecemeal into fully complex language. In particular, the earlier stages don't go away – they are just elaborated upon and embedded in the more complex grammars.

Which of the levels of the hierarchy were actually instantiated in earlier hominids, and when? At the moment, we think there is no way to tell. Perhaps when more is understood about how brain structure creates specialized cognitive capacities, how genetics influences brain structure, and how the genome evolved from earliest hominids to modern humans, it will be possible to address these questions more rigorously. Meanwhile, we think the best that can be done is to keep one's eyes open for new evidence and try to develop theories of language that lend themselves to plausible evolutionary scenarios.

6. Conclusions

To sum up, we suggest that our grammatical hierarchy and the associated interface rules give us a useful way of pulling together a great number of disparate phenomena that fall under the radar of conventional generative grammar: issues in child language acquisition, late acquisition, language creation, language emergence, pidgins, language processing, language deficits, and the evolution of the language faculty.

A virtue of our analysis, we think, is that even fully complex languages display many symptoms of lower levels of the hierarchy, both in central domains such as word order and compounding and in more marginal constructions (see also Progovac, this volume). Moreover, there are full languages that appear to make little or no use of the upper reaches of the hierarchy. We think that our hypotheses lead to new and interesting conjectures about the nature of the human language faculty.

References

Anderson, Stephen (2004). *Doctor Dolittle's Delusion: Animals and the Uniqueness of Human Language.* Yale University Press.

Aronoff, Mark, Irit Meir, Carol Padden, & Wendy Sandler (2008). The roots of linguistic organization in a new language. In Holophrasis, Compositionality and Protolanguage, Special Issue of *Interaction Studies.* Derek Bickerton and Michael Arbib (Eds.) pp. 133-149.

Bickerton, Derek (1981). *Roots of Language*. Ann Arbor, Michigan: Karoma.

Bickerton, Derek (1990). *Language and Species*. Chicago: University of Chicago Press.

Bickerton, Derek (2008). *Bastard Tongues*. Hill and Wang.

Bloom, Lois (1970). *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.

Braine, Martin (1963). The Ontogeny of English Phrase Structure: The First Phase. *Language* 39, 11-14.

Brown, Roger (1973). *A First Language.* Cambridge, MA: Harvard University Press.

Caramazza, Alfonso, and Edgar Zurif (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language* 3, 572-582.

Chomsky, Noam (1956). Three models for the description of language. IRE Transactions on Information Theory (2): 113–124.

Chomsky, Noam (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.

Dennett, Daniel C. (1995). *Darwin's Dangerous Idea*. New York: Simon & Schuster.

Dick, Frederic, Elizabeth Bates, Beverly Wulfeck, Jennifer A. Utman, Nina Dronkers, and Morton Ann Gernsbacher (2001). Language deficits, localization, and grammar: Evidence for a distributive model of language breakdown in aphasic patients and neurologically intact individuals. *Psychological Review* 108, 759-788.

Everett, Daniel (2005). Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46.

Everett, Daniel (2009). Pirahã Culture and Grammar: A Response to Some Criticisms. *Language* 85, 405-442.

Feldman, Heidi, Susan Goldin-Meadow, and Lila Gleitman (1978). Beyond Herodotus: The Creation of Language by Linguistically Deprived Deaf Children. In A. Lock (ed.), *Action, Symbol, and Gesture: The Emergence of Language*, 350-414.

Ferreira, Fernanda, and N. Patson (2007). The good enough approach to language comprehension. *Language and Linguistics Compass*, 1, pp 71-83.

Fodor, Jerry, Thomas Bever, and Merrill Garrett (1974). *The Psychology of Language*. New York: McGraw-Hill.

Gil, David (2005). "Word Order Without Syntactic Categories: How Riau Indonesian Does It", in A. Carnie, H. Harley and S.A. Dooley eds., Verb First: On the Syntax of Verb-Initial Languages, John Benjamins, Amsterdam, 243-263.

Gil, David (2009). How much grammar does it take to sail a boat? In G. Sampson, D. Gil, and P. Trudgill (eds.), *Language Complexity as an Evolving Variable.* Oxford: Oxford University Press.

Gil, David (this volume). (…)

Givón, Talmy (2009). *The genesis of syntactic complexity: Diachrony, ontogeny, neuro-cognition, evolution.* Amsterdam: John Benjamins.

Goldin-Meadow, Susan (2003). *The Resilience of Language*. Psychology Press.

Hafri, A., Anna Papafragou, and Jon Trueswell (in press). Getting the gist of events: Recognition of two-participant actions from brief displays. Journal of Experimental Psychology: General.

Hunsicker, Dea, and Susan Goldin-Meadow (2012). Hierarchical structure in a self-created communication system: building nominal constituents in homesign. *Language* 88, 732-763.

Hurford, James (2011). *The Origins of Grammar: Language in the Light of Evolution.* Oxford University Press.

Jackendoff, Ray (1987). *Consciousness and the Computational Mind*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1999). Possible Stages in the Evolution of the Language Capacity, *Trends in Cognitive Sciences* 3, 272-279.

Jackendoff, Ray (2002). *Foundations of Language*. Oxford University Press.

Jackendoff, Ray (2007). A Parallel Architecture perspective on language processing. *Brain Research* 1146, 2-22.

Jackendoff, Ray (2011). What is the human language faculty? Two views. *Language* 87, 586-624.

Jackendoff, Ray, and Eva Wittenberg (in preparation). *What You Can Do Without Syntax*.

Kegl, Judy, Ann Senghas, and Marie Coppola (1999). Creation through contact: Sign language emergence and sign language change in Nicaragua. In Michel DeGraff (ed.), *Language Creation and Language Change*, 179-238. Cambridge, MA: MIT Press.

Klein, Wolfgang, and Clive Perdue (1997). The basic variety (or: couldn't natural languages be much simpler?). *Second Language Research*,13, 301-347.

Kuperberg, Gina (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research* 1146, 23-49.

Nevins, Andrew, David Pesetsky, and Cilene Rodriguese (2009). Pirahã Exceptionality: a Reassessment. *Language* 85.2, 355-404.

Progovac, Liljana (2009). "Layering of Grammar: Vestiges of Proto-Syntax in Present-Day Languages." In Language Complexity as an Evolving Variable, ed. by Geoffrey Sampson, David Gil and Peter Trudgill, 203-212. Oxford: Oxford University Press.

Progovac, Liljana (this volume). (…)

Sandler, Wendy, Irit Meir, Carol Padden, and Mark Aronoff (2005). The Emergence of Grammar in a New Sign Language. *Proceedings of the National Academy of Sciences* 102, No. 7. 2661-2665.

Seidenberg, Mark, and Laura Petitto (1978). Signing behavior in apres: A critical review. *Cognition* 17, 177-215.

Selkirk, Elizabeth (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elizabeth (2000). The interaction of constraints on prosodic phrasing. In Merle Horne (ed.), *Prosody: Theory and Experiment*, 231-261. Dordrecht: Kluwer.

Senghas, Ann (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. Cognitive Development 18 (2003) 511–531.

Terrace, Herbert (1979). *Nim*. New York: Knopf.

Tomasello, Michael (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Harvard University Press.

Townsend, David, and Thomas Bever (2001). *Sentence Comprehension*. Cambridge: MIT Press.