Guessing Rule 1: towards a 'Modest UG'

Avery D Andrews
Jan 2011, ANU

Generative grammar (especially the GB/Principles and Parameters wing) has traditionally pursued a rather aggressive program concerning linguistic universals, of making the strongest possible assumptions that don't contradict whatever is taken to constitute known typology. This program approach has led to the accumulation of a great amount of knowledge concerning the syntax of various languages, the details of word order in many European languages being a major example. But is also a downside, including:

(1) a. The concrete proposals about UG, such as for example Subjacency, have a short lifespan, and don't seem to be get replaced by clearly stated principles with demonstrably more accurate coverage (rather, a great deal of information about the phenomena is collected, covered by a penumbra of incomplete proposals).

   b. Certain kinds of data, such as for example 'multiple case' (Dench and Evans (1988), Evans (1995), Plank (1995), Nordlinger (1998), Round (2009)), tend to be systematically neglected (e.g. in Baker (2008)), probably because they present an insurmountable challenge to basic assumptions (such as the idea that all agreement morphemes can be usefully housed in functional projections).

   c. When not ignored entirely, embarrassing data is often managed in undisciplined ways, such, such as being attributed to 'the periphery' (or the 'phonology'), or masking it with a loophole that isn't narrow enough to preserve the empirical force of the theory being 'saved'.

This issues are not necessarily fatal, given that generative grammar is a young field without much in the way of deep established principles, but it follows from this is that it would be good to have roe than one philosophy for approach it.

I therefore see some merit in the pursuit by some people of a program of 'modest UG', where the goal is to explain the possibility of language learning with relatively weak assumptions about UG, such as for example those often used in work in LFG, HPSG, and other frameworks (those compatible with the 'Simpler Syntax' approach of Culicover and Jackendoff (2005)).

These typically include a variety of 'rule notations', inducing a concept of relative complexity of grammars that can be used for something like Chomsky's (1964, 1965) 'evaluation metric', but also need to be supplemented by some kind of learning theory in order to determine a grammar for a given body of data (the evaluation metric functioning as a bias for shorter grammars, other things being equal). As Chomsky used to point out, this is not trivial, but the idea of trying to accommodate what learners appear to acquire when they learn a language in terms of a manageable number

of parameters assisted by something like the subset principle does not seem to have worked out.

However more than just rule notations and an evaluation metric is clearly required; Pinker (1984) for example needed to make a fair number of substantive UG assumptions, such as some version of the problematic X-bar theory, to get his learning theory for LFG (still unequalled in coverage and detail, as far as I know) to work. I will here make a proposal which requires no substantive universals at all (as this term is usually understood), in the form of a 'Guessing Rule' (defeasible principle) for identifying 'noun phrases'. I will refine this notion later, so that there won't be any assumption that 'NP' is a substantive universal.

I call the principle 'Guessing Rule 1'. and originally developed it in order to make it easier to mark basic syntax problems on unknown languages, where students are often rather creative in how they manage to get the NP structures wrong. And most annoyingly, it is in my experience often quite difficult to really show that the 'wrong' analyses are actually wrong on the basis of just the data in the problem, or any similar problem appropriate to that stage of the course. It's only when more data of a different kind (e.g. agreement) is taken into account that their true wrongness is revealed. GR1 was therefore developed as a principle that I could teach, and then ruthlessly deduct points for failure to apply. It is described as a 'Guessing Rule' because it doesn't always work (which is not a problem if there's a good way to tell when it is failing), and it is not supposed to always return a verdict. Nevertheless, the fact that virtually all descriptive grammars contain a chapter or section on 'NP structure' suggest that something like it probably works often enough to be useful.

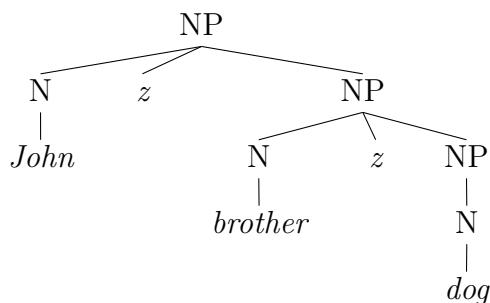## 1   The Guessing Rule

A simple statement of the proposed rule is:

(2) GR1: if a (contiguous) sequence of words, uninterrupted by a phrasing boundary, identifies a participant, guess that this sequence is an NP.

An example of the kind of misanalysis that the principle is supposed to ward off is provided by recursive possessive construction, such as *John's brother's dog('s collar)*. Typical misanalyses that students often come up with include the 'inverse analysis' (a) and the 'flat analysis' (b):

(3) a. NP   →   (Det)   AdjP*   N   (z   NP)

   b. NP   →   (Det)   (NOM   z)*   NOM
     NOM   →   AdjP*   N

The inverse analysis for example produces this tree for *John's brother's dog*

(4) inverse tree:

```
                        NP
              ┌──────────┼──────────┐
             N          z           NP
             │                ┌──────┼──────┐
            John            N       z       NP
                           │                │
                         brother            N
                                            │
                                           dog
```

Such analyses can often be marked down on various fine points, depending on the details of the assignment. For example, for English, the inverse analysis doesn't get the distribution of articles and determiners quite right, allowing things like *the boy's the dog*. But the flat analysis does, and such facts cannot be depended on (things such as *God's the beloved son* occurred in some earlier stages of English, and in many other languages). Furthermore, phrase-structure rules in general do a rather poor job of constraining the distribution of determiners, so appealing to them here would not be convincing to really good students.

The main reason we know that these analyses are almost always wrong is the way agreement works:[1] with a very few exceptions, if there is agreement of a predicate with a possessed NP argument, it will be the properties of the possessum rather than the possessor that are manifested by the agreement. But this isn't satisfying from the point of view of pedagogy or the 'projective/explanatory' goals of generative grammar.

Pedagogically, phrase structure comes naturally before agreement, because phrase-structure can help a lot with the description of agreement, but we don't want the correct treatment of things introduced in week 2 to depend on topics that aren't carefully discussed until week 5, if we can possibly avoid it. Shifting to the problem of explaining language learning, we would like the details of the more refined and subtler phenomena such as agreement to mostly follow from the more obvious and easily ascertained ones, such as basic constituent structuring, rather than having the latter depend on the former.

GR1 addresses this problem by requiring that each possessor be an NP, as well as the entire construction, thereby eliminating flat and inverse analyses. The phrase structure will then be almost always set up to as to support agreement properly, as we want in a theory that is supposed to help explain language-acquisition.

2 Practical Performance

Before discussing refinements, it might be useful to consider some examples of the performance of the guessing rule on actual material. A particular reason for doing this is its rather informal statement. This is not a fatal problem in the early life of a proposal if it seems obvious how to apply it to normal linguistic material, but would

---

[1]Case marking can also provide indications, but is trickier, due to complexities such as Suffixauf-nahme.

be fairly serious if the practical application was uncertain.

The amount of NPs it turns up seems to vary considerably with genre and language (jointly, at least). Relatively thin pickings are found for example in Evans' (2003) Bininj-Gunwok texts (narrative). In the first text, one has to go to sentence 39 to get a solid, 2-word example:

(5) Ba-yerrng-yiga-ni   ba-djoleng-m-inj ba-ru-i   na-wu   gunj
    3/3P-wood-go.for-PL 3-ready-INCH   3P-cook-PP MA-that kangaroo
    While she was going for wood it had been cooked and got ready, that kangaroo.

In context, it's clear that this kangaroo has been already introduced several sentences earlier,

There is an earlier conceivable example, in sentence 16:

(6) Njamed   na-wu, ragul
    watchamacallit na-wu, red-eyed pigeon
    that watchamacallit, the red-eyed pigeon

The problem here is that 'that watchamacallit' is not something that can be plausibly taken to constitute an identification, and *ragul* 'red-eyed pigeon', which does, is a single word, and the whole three word sequence is separated by a phrasing boundary, so we don't want to be pushed by GR1 to construe it as a single NP. This illustrates the important point that unlike many kinds of principles that one might think of, a guessing rule such as GR1 doesn't have to turn in a verdict on everything that is put in front of it: clearer application to a more limited range of cases can be taken as preferable less clear application to a wider range.

More results are found in Evans's (1995) Kayardild texts: an NP is detected in the very first line (an introductory utterance consisting of a single NP):

(7) dathin-a   dulk-a
    that-NOM place-NOM
    that place

Another 16 two-word NPs appear in the 49 lines of the first text (no longer ones). A greater tendency to use demonstratives is a major contributor to the richer pickings found in the Guninj-Binwok texts, but various possessive and attributive constructions also appear:

(8) a. dul-marra   dangka-a
      place-GEN man-NOM
      the boss of that place (line 2)

    b. bal-umban-da   jardi-y
      west-DENIZEN-NOM mob-NOM
      the western mob (line 6).

I can't say to what extent these differences might be caused by genre or predilections of individual speakers.

By contrast, the Homeric poems bristle with the complex NPs that are picked up by the principle, for example line 9 of the Iliad, where the NPs indicated by GR1 are square-bracketed:

(9) [[le:tous] kai [dios]   uios],   [ho]   gar [basile:ï] xolo:theis
    Leto(G) and Zeus(G) son(N), he(N) for king(D) having-gotten-angry
    The son of Leto and Zeus, for having gotten angry at the king (Il 1:9)

The names of Leto and Zeus are NPs, and so is the whole sequence referring to their son (Apollo). The two parents' names coordinated with *kai* 'and' on the other hand is not taken as an NP indicated by GR1, although it certainly is one by the grammar of Homeric Greek. The reason is that I consider the notion of a pair of entities (basically a kind of i-sum in the sense of Link (1998) to be a bit too abstract to bake into the first version of GR1. One might, however, later decide to elaborate the principle in such a way as to include significant social groupings, such as pairs of people that appear to be an 'item'. But I suggest that any such extension should be based on actual work on social cognition by children learning languages.

This line also indicates why we should include the reservation about phrasing breaks in GR1, since otherwise, it would seem likely that we'd have to include *ho* at the start of the next clause as part of our NP, which would make a mess.[2]

A possible reason for the rather high 'complex NP density' of the Homeric poems is that they are highly polished performances, which, even in their earliest stages of development, were probably rather well thought out before being delivered (if you were going to try to entertain drunken warlords and their retainers in the evening, you would probably think pretty hard during the day about how you were going to do it). On the other hand, it is far from clear how to set up an appropriate density metric that would work across languages and genres.

Another reason not to get too carried away with trying to do such measurements is that they only have arguable relevance to language learning when applied to the kinds of input that child (and maybe, in some cultures, older) language learners have access and pay attention to. A lot of work would need to be done to address this.

So now, on to some refinements and further qualifications.

3   Refinements

First, what do we mean by 'identify a participant'? What I intend by this is that the sequence of words must pick out some person, place, thing, plant or animal (note the deliberate imitation of Lyons (1968:317-318), extended to include the oft-ignored animals) that the speaker assumes that is known or knowable to the hearer, on the basis of the current situation. This means that the putative NPs so delineated must

---

[2]As pointed out by Nick Evans.

be definite, via familiarity, uniqueness, or deixis. As discussed above, identifying these circumstances well enough to apply the principle does not appear to be a problem.

Indefinites on the other hand are excluded from the scope of GR1. The reason for this can be seen by considering:

(10) a. A man who was talking was laughing

b. A man who was laughing was talking

c. $Past(\exists x.Man(x) \wedge Talk(x) \wedge Laugh(x))$

The problem is that when an indefinite NP establishes a new participant to talk about, both the material within and outside of the NP provide information that can later be used to identify the referent, with an NP such as perhaps *the man who was talking and laughing*. There might be something subtle that we could say about the nature of the meaning difference between (10a) and (10b), but the absence of anything really important is revealed that they would both get the essentially the same traditional logical form, (10c). Therefore, it seems unlikely that any meaning difference that existed would be clear enough to serve as data for early language-learning.[3] Therefore, the full language acquisition method would have to acquire indefinite NPs in a somewhat different way than definite ones, perhaps implicitly assuming that the same structures that are used to identify known participants are also used to introduce new ones.

Another kind of NP that we will leave out of the scope of GR1 are the quantificational ones:

(11) *Every man who was talking* was laughing

Here, there might be prospects for a workable formulation of a principle that would identify *every man who was talking* as an NP, but trying to manage this is beyond the scope of this paper.

A second point is that by 'sequence', we mean 'contiguous sequence'. Therefore, in a Warlpiri sentence such as (12), GR1 would not identify the two descriptors of the Agent of *wangka-* 'speak' as an NP:

(12) kurdu ka   wangka-mi   wita
     child PRES.IMPFV speak-NPAST small
     The small child is speaking

Neither would it suggest grouping the two apparent separated NPs into one in an English sentence with an afterthought:

(13) The child is speaking, the small one.

---

[3]A particularly striking indication of this is the difficulty that syntacticians have had in working out whether *duck in the sink* is an NP or not in *there is a duck in the sink*.

The reason for this is a metatheoretical guess that admitting discontinuous phrases into constituent structure is a Bad Idea.

A justifiable exception to requiring constituents to be uninterrupted is the treatment of 'second position clitics', which sometimes to seem able to interrupt NPs, as in this example from the Odyssey:

(14) he:n te   psuche:n kai noston hetairo:n
     his   and life/soul and return companions(G)
     both his own life and the homecoming of his companions (Od 1:5)

Here the *te* is the component of the coordinating conjunction that syntactically belongs in front of the first conjunct (functioning similarly to *both* in English. This might be taken as an example of the kind of 'loophole' that Levinson and Evans (2010) disapprove of in grammatical theory. But in this case it's all right, due to the very limited nature of the exception: the interruptors are closed-class items, which are furthermore clearly restricted to appearing after a first word or constituent. And perhaps even more strongly than that: Legate (2008) argues that when such clitics in Warlpiri interrupt a constituent, they can only appear after the first word of that constituent, and preliminary investigations suggest that this works Homeric Greek as well.

If we allow phrases to be interrupted by certain clitics, we then have to consider whether or not to have GR1 recognize these constituents; it would be simpler to have it ignore them, so that's what I'll suggest for now, although one might have to reconsider this if they proved to be sufficiently common in the child's input for learning in some languages.

A further, perhaps obvious limitation of scope is that we don't want to suggest that parts of identifications be NPs, i.e., the NPs should be maximal contiguous identifications. So in the subject of (a) below, we are only required to have one NP, in (b), three:

(15) a. The greedy, unscrupulous property developer bribed the council members with free tree-pruning.

   b. John's nice sweater from Iceland should not be washed.

It is not clear, however, what kinds of principles we want to guide the analysis of the internal structure of these NPs. In particular, determiners and possessors tend not to be repeatable, suggesting a principle that discourages analyzing NP-internal sub-descriptions as NPs (note that if the DP hypothesis of Abney (1987) and much subsequent work, is accepted, the formulation of GR1 will need some adjustment, as discussed below).

But of course partial identifications very often are NPs (as in the case of indefinites, or the two components of afterthought constructions as in (13), so we need to allow this, and will need some additional facilities, very likely some kind of distributional analysis, to find these NPs.

Another subtlety arises with closed class 'relators' appearing at either margin of an NP, such as *'s* and *from* in (15b). These don't contribute to the description of the entities described by the NPs they appear next to, so, by the principle as articulated so far, they shouldn't be included in those NPs. This seems good, since the analysis of *from* as the head of a PP (which contains an NP) is accepted by pretty well everybody, while those that don't accept Abney's DP-based analysis of possessors, where the possessive marker is a D-head taking the possessor as its Spec and the remainder of the NP as a $\overline{\text{C}}$ complement, would probably accept z as head of a 'PossP' of the form NP Poss ($=z$).

But a possible problem arises in languages such as formal-style Japanese, where, depending on the analysis, every NP either includes or is followed a 'case particle', and Tagalog (and many other Philippine languages) where every NP includes/is preceded by one. One possibility is to say that in these languages, NPs are always introduced inside 'KPs', which begin or end with a case particle (K); another is to allow an exception to GR1, whereby closed class items that signal the relationship of the individual identified by the NP to its surroundings can be included within the NP (but can also be left outside of it). In either approach, some kind of distributional analysis is required to work out whether NPs with markers fall into a different category than those without them. How hard this analysis is will depend somewhat on how sophisticated the morphological learning is, since in many languages, such as Turkish or Warlpiri, some NPs include case-markers that are morphologically part of a nominal word, while others don't, but this does not correlate to anything like an NP/PP distinction.[4] We will return to this issue later when we consider category terminology, i.e. what, exactly, do we really mean by 'NP'?

But notice that the problem of case-particles might be related to the problem of 'determiners' more generally, such as the articles *the* and *a*, which don't contribute substantive descriptive material but clearly belong in the NP anyway. An interesting contrast is the Spanish DOM-marking preposition *a*, which Alsina (1996) shows is a preposition, heading a PP (for Catalan specifically, but the arguments work identically for Spanish). This functions as a combined marker of accusative case, animacy and specificity (the last a standard function of determiners), but, but in terms of phrase-structure, it clearly goes outside the NP rather than inside it. Hindi *ko* on the other and has a very similar function, but appears to be an NP-internal case-marker rather than head of a distinct PP category. So we need to rely on some kind of distributional analysis to find the correct treatment of closed-class items on the periphery of the putative NPs mandated by GR1.

Possessors with demonstratives pose interesting issues, since they are closed-class items. Consider an example like this:

(16) that man's nose is red

As we have qualified it so far, GR1 says that there is one NP including *man's nose*, and

---

[4]But rather, to the functionally but not phrasally similar distinction between 'syntactic case' vs 'semantic case'.

another including *man* (but not *nose*), but doesn't say that either must contain *that*, even if we suppose that in most contexts, some kind of pragmatic salience principle will specify that it applies to the man rather than the nose. But given examples such as *don't laugh at that man*, etc, there is distributional evidence that it can constitute an NP with *man*, and so, by GR1, should be able to in (16). Putting it there delivers the most salient reading for (16) by the same rules as work for non-possessive NPs, but doesn't settle the question of whether the possessum NP can have its own demonstrative. That it cannot for a possessive interpretation of the construction will have to be worked out by some other means.[5]

GR1 cannot therefore replace conventional distribution-based analysis, but only jumpstart it by means of access to certain aspects of the meanings of sentences, which are available to syntax students via the translations (and sometimes even glosses) of the examples, and, presumably, to language learners from their understanding of the contexts, and meanings of individual words. We will later make a suggestion about what might be involved with the distributional analysis, especially the apparent use of negative evidence, such as the absence of of 'double determiner' sequences, such as *this that man's house*.

We restate GR1 with the qualifications as follows:

(17) GR1: if a maximal contiguous sequence of words, uninterrupted by a phrasing boundary, identifies a participant, guess that this sequence is an NP, with the qualification that closed class items with bleached/grammatical meanings on the edge might either be inside or outside the NP.

4   GR2

Prepositions and case-particles bring on a further question, which is not fully relevant to GR1, but close enough to be worth discussing here. This is whether there is any principle that can motivate PPs from a very small presentation of data, once a learner or problem-solver has decided that the P is not part of the NP. For this one can suggest a principle which I call 'Guessing Rule 2', for which a rough verbal formulation would be:

(18) If two constituents (words or phrases) tend to appear in association with each other, group them into a phrase.

But if one is developing PS rules, similar effects can be achieved in a more formally precise manner by saying:

(19) Don't put the parenthesis or parenthesis-star notation around a sequence, but only around single items.

---

[5]Note that there's also a 'characterizational genitive', illustrated by *this women's bycicle*, where the demonstrative modifies the entire NP. The genitive in this case does not appear to refer to any specific participant, and a full range of NP structures is not possible (the construction seems similar to compounding in various ways).

Given this principle, a rule such as:

(20) VP   →   V   (NP)   (NP)   (P   NP)

won't be allowed (neither, of course, its variant with S), and the PP alternative will be rigorously motivated.

The two formulations are however not fully equivalent in their effects. The 'functional' formulation (18) should probably be interpreted as having an issue with the DP hypothesis, because this analysis groups the possessive marker with the possessum (under an $\overline{\text{D}}$ constituent), rather than the possessor, as mandated by (18). The rule-form-based version (19), on the other hand, doesn't do this, since the DP hypothesis doesn't depend on putting notations around sequences.

This point requires that we at bottom reject Levinson and Evans' 2010 view of formal notations as mere toolboxes of convenience: interpretation (19) requires that we treat the rule-notation as a genuine theoretical commitment with empirical content. I say 'at bottom', because, given the depth and breadth of our ignorance, proposal such as (19) have to be taken as guesses of approximations to the truth, so I doubt that there is any practical consequence between the 'at bottom realist' view and Levinson and Evans'.

The Crow language (Siouan) of southern Montana shows that GR2 must be defeasible, with data such as below for directional and instrumental particles:[6]

(21) a. Chichúche kush dée-k
      Hardin   to   he.go-DECL
      He went to Hardin (Montana).

     b. Íilaale ii   lée-k
      Car   with he.go-DECL
      He went with a car.

     c. Íilaale Chichúche ii   kush dée-k
      Car   Hardin   with to   he.go-DECL
      He went to Hardin with a car.

GR2, in either of its formulations, suggests that *kush* 'to' and *ii* 'with' in (a,b) should be treated as postpositions forming a PP with the NP whose semantic role they mark, but this analysis falls apart when we get to examples like (c), where it is revealed that these markers can (and *kush* must) appear prefixed to the verb, leaving their associated nominals bare NPs in the clause. (c) requires an alternative analysis whereby the markers are attached to the verb, which also covers the data of (a,b); the postpositional analysis can then be dropped as unnecessary.

GR2 is clearly a much less significant principle than GR1, since the general mess caused by not recognizing PPs, PossPs, etc. is much less extensive than if NPs are omitted.

---

[6]Siouan, Southern Montana, data from fieldwork in the 1970s; the next version of this paper will hopefully be able to refer to Graczyk (2007).

And if the version of (18) is accepted, it will certainly have to be defeasible, although only positive evidence would appear to be required.

5   Categories: What is 'NP'?

Now it's time to deal with an issue that has come up several times already, namely, exactly what are these 'NP's that GR1 is telling us to use? The problem is that, in frameworks that use phrase-structure rules, there are a variety of different analyses available, such as conventional NP structure versus the DP hypothesis, plus, in some frameworks, the potential for NPs be treated as members of technically different categories in different languages. For example, $\overline{\overline{\text{N}}}$ in English, but $\overline{\text{N}}$ in Warlpiri (Simpson 1991). Similar problems arise in frameworks that don't use phrase-structure rules as such, but employ some devices with similar capacities.

One possibility is that this potential variation reflects excessive weakness in UG, which should be fixed by making some sort of strong claim, but what I'll do here is propose a more general form of the principle, where GR1 says that all contiguous sequences of words constituting an entity identification, should, by default, be analyzed as constituents of some single PS type, which we can then call NP (at least if it contrasts in structure and distribution with any other phrase-type, especially one associated with propositions). This should be seen as a variation on the technique that Lyons (1968:317-318) used to rehabilitate the traditional definition of 'noun' as 'the name of a person, place, or thing'.

This formulation is intended to be consistent with the rather dramatic claims about Riau Indonesian made by Gil (2005), to the effect that this language has only one syntactic category (S), comprising all open class lexical items as well as all phrases. A critique of this treatment has been offered by Yoder (2010), but for our purposes here, we will accept Gil's points. The reason we can do this is simple: GR1 says to put a node of some single category over all full identifications, but doesn't say not to put it over anything else (indeed, it would face big problems if it contained such a restriction). In particular, it doesn't prohibit putting that node over assertions, questions, commands, etc. In English, attempting to use the same phrase type for entity identifications and normal main-clause assertions doesn't work out at all well, but this does not mean that it won't be viable for some other language.

And in fact, without a great deal of careful preliminary work, it would be unwise to try to insist on any principle that would rule out such a permissive use of the NP node, for a number of reasons, even without reference to Gil's claims about Riau Indonesian. One is that English has a variety of utterance types used in various situations that appear to structurally consist of a single NP:

(22) a. 'Confirmatory Response Exclamation': *what a cute kitten!* (when presented with a picture of a kitten that you're obviously expected to consider cute).

b. 'Dismay Exclamation': *my novel!* (When you trip carrying the only typescript on a bridge across a river, and a wind gust comes up and scatters most of it onto the water).

    c. 'Critter Spotting Game NP': *a (big) kangaroo (with a joey in its pouch)!* (pointing to one, while taking a toddler through a zoo)

Intuitively, these appear to be 'special' utterance types used in a limited range of situations, but it would probably not be an easy enterprise to produce a coherent and typologically plausible principle that allowed them to be analyzed as NPs but not *John loves Mary*.

A further problem is afforded by the diachronic phenomenon of 'insubordination' (Evans 2007, Mithun 2008), in which various kinds of subordinate clause structures, including nominalized ones, come to be used as main clauses. Sorting out what the correct analyses for these things are would be a precondition for a principle requiring 'NP' to be different from 'S', and this would be a very difficult enterprise (consider for example the issues involved with 'mixed category' constructions). As a final point, I'll note that although having a distinction between a 'entity' and 'propositional' type is traditional in formal semantics, it is not clear that it is really necessary (Partee 2006).

Another issue is whether we should try to relate the concept of NP presented here to any concept of 'noun' defined in some way. In the first place, it is certainly not viable to require NPs to contain a noun (English examples such as *the dead*, for example, where the second word passes all tests for being an adjective rather than a noun). More tolerable would be to say that the NP should be able to contain a noun, but then we need to work out exactly what a noun is. Although Lyons' proposal is a good start, I think there are still some tricky issues to deal with (such as where classifiers fit in), so I won't pursue this any further here, beyond noting that the principle should suggest that words meaning something along the lines of 'kinds of people, places, things' (and, living things, I would add) ought to be expected to fall into a single lexical category (with some kinds of possible subcategories), without requiring this category to be different from any other categories (even phrasal ones) whose existence might be suggested or required by other principles.

A final aspect of the principle that is worth mentioning is that, unlike some versions of generative grammar that have occasionally been assumed in the past, it assumes that semantic knowledge is intimately involved in grammar acquisition. In particular, to apply the principle, the learner must be able to guess at least a reasonable approximation to the meaning of the utterance, implying a substantial grasp of what's going on in the context of utterance. As Chomsky has often pointed out, this is a far-from-innocent assumption, and in the 1970s, the idea that such access shouldn't be required for syntactic learning was formulated as the thesis of the 'absolute autonomy of syntax' (Chomsky 1975:43).[7] However, this isn't intuitively very plausible, and should be regarded as another 'conceptually best case' scenario that we don't have to invest much intellectual capital in, and probably shouldn't.[8]

---

[7]At a class at MIT in the mid seventies, I asked Chomsky if this thesis entailed that a significant amount of the syntax of a language could be learned from untranslated radio broadcasts, and he said 'yes', which would seem to settle the matter of what the thesis is supposed to mean.

[8]However, Anna Wierzbicka once devised a syntax problem wherein one could deduce a certain amount of the syntax of Basque with no meanings provided, so the thesis is not completely impossible.

6 Defeasibility of GR1

GR1 is a principle which gives strong hints about phrase structure if certain kinds of evidence are present, but just as with GR2, it doesn't always work. Here we will survey a representative (but not exhaustive) assortment of such case, and discuss how a language-learner might be able to acquire the correct analysis nevertheless.

The first we will look at occurs with inalienably possessed NPs in certain Amazonian languages, such as Jarawara and Banawa, described in Dixon (2000, 2004).[9] Jarawara (to which Banawa is closely related) has a closed class of about 200 inalienably possessed stems, which must occur with a prenominal possessor. About half of these carry gender (M vs F) marking for the gender of the possessor, and the gender of the possessor is *always* what is marked (in complex ways) on verbs of which the whole NP functions as a core argument:

(23) a. Okombi   mano   koma-ka
        Okombi(M) arm(M) sore-DECL(M)
        Okombi's arm is sore

     b. Manira   mano   koma-ke
        Manira(F) arm(F)sore-DECL(F)
        Manira's arm is sore

It is clear that if most or even a reasonably large number of languages handled agreement in their possessive constructions like this, it would be quite wrong to mark students off for coming up with the inverse analysis of possessives. Indeed, the agreement behavior of the possessum nominals here is essentially the same as that of attributive adjectives in many languages, and any theory of the syntax-semantics interface that can handle 'head-switching' modification[10] could produce appropriate semantic structures from a syntax where the possessum was expressed as a syntactic adjunct. Dixon shows that, once the considerable complexities of the agreement system (involving person and number as well as gender, and behaving in unusual ways), treating the possessor as head creates no problems and is arguably the best analysis.

However, the evidence that GR1 is not producing the expected behavior here is highly overt in the morphosyntactic forms of many of these inalienably possessed NPs, indeed, all of those in which the possessum nominal overtly marks gender. A morphology-aware learner will then have quite a lot of evidence that the structure of these NPs differs from the expectations induced by GR1, so that the somewhat less overwhelmingly obvious phenomena of agreement of predicates shouldn't be a big surprise.

A substantial remaining problem posed by these languages for UG is what the constituent structure of these examples really is. If we constrain UG to disallow head-

---

[9]Pointed out to me by Dan Everett (pc).

[10]Exemplified by the rough semantic equivalence of (a) and (b) below

    a. John is obviously unwell

    b. It is obvious that John is unwell

switching semantics for possessors, we could use a conventional PS analysis where the features of the possessor are copied onto the entire NP (perhaps due to the lack of any such features in the lexical specification of the possessum).[11] But the resulting 'loophole' might be big enough to 'let all the air out of the theory' (phrasing due to Haj Ross), so we'd need to think carefully before making grand claims.

In our next two cases, the evidence that something that GR1 suggests should be an NP is not one resides in more general properties of the overall pattern, so that a more sophisticated learning method would be required to detect it. These are 'nominal clitics' in Spanish and Indonesian.

In Spanish, direct (and indirect) objects are frequently represented by clitics which precede finite verbs and follow nonfinite ones, and serve to identify participants by providing some gender and number information:

(25) el   chico las   vió
      the boy   them(F.PL) saw.he
      the boy saw them(F.PL)

It is quite clear that these items do not constitute NPs, because the full range of NP structures can't appear in the position between an overt subject and its finite verb:

(26) a. el   chico vió   a   las mujeres
          the boy   saw.he OM the women(F.PL)
          the boy saw the women

      b. *el   chico (a) las mujeres   vió
          the boy   OM the women(F.PL) saw.he

The clitics also don't occur with the prepositional OM *a* discussed above, which is obligatory for definite/specific human direct objects.

Furthermore, the clitics don't appear in standard positions for full NPs, such after the finite verb (without following material such as adjectives and nouns, for which some of the clitic forms can be determiners). Nevertheless, they constitute entity identifications, so GR1 says to guess that they are NPs. But it is quite clear that they aren't.

So now we have an instance where the guess suggested by GR1 is wrong, and, furthermore, the evidence that this is so does not appear in the surface form of any morphologically analysed example, but requires some grasp of the wider syntactic pattern. In particular, how are we to avoid the use of 'negative evidence'[12] such as that of (26b) and (27) below:

(27) *el chico vió (a) las

---

[11]The fact that inalienable possessors count as 'pivots' looks like a potential problem for such an analysis.

[12]That is, direct evidence that a particular sentence is bad; on reflection, it should be clear that this, the conventional meaning of the phrase in linguistics, is not the only one it could have.

This is especially tricky in this instance, because Spanish word order is rather flexible, and there are quite a number of situations in which full NPs can appear directly in front of a finite verb (but not after an overt subject). So the distributional analysis would have to be fairly clever, in order not to have to presuppose notions that it was supposed to establish.

The solution I suggest is to generalize the 'mutual exclusion' principle well-known from children's acquisition of lexical meanings. The exclusion principle says that children are reluctant to accept that two different words, such as *dog* and *animal*, can be correctly applied to the same creature. The idea that we want to use here, clearly derivative on this, is that an analysis is rendered suspicious for a learner when some structure that it predicts doesn't appear in situations where their grammar says it could, but some other structure does instead. So consider the performance of a grammar in which *las* with pronominal meaning is an NP, admitting the structure (25b) and (27). In situations where (27) would make sense, and can plausibly expected, it doesn't appear. Instead, expressing the same meaning, the learner finds most often (25), but sometimes its 'clitic doubling' variant (28) below:

(28) el chico las vió a ellas

The proposal is the appearance of these structures in cases where (26) is expected (given a grammar in which the clitics are NPs), together with the failure of structures like (26b) to appear in circumstances where postnominal objects do, provides a kind of 'implicit negative evidence' that the clitics are not NPs. One might call this kind of 'negative evidence detection' 'regenerative retrodiction', since it is supposed to work by regenerating the observed utterance form its computed semantic representation.[13]

What is needed to exploit this kind of evidence is the capacity to check what ways ones current grammar provides for reformulation of an occurring utterance, together with some way of revising the grammar, presumably probabilistically, so that alternatives that aren't used get progressively deprecated. This idea is similar to the The right way to do this for syntax Gradual Learning Algorithm of Stochastic Optimality Theory (Boersma and Hayes 2001), but this assumes a very rich UG in which all constraints are universal, so we can't use it as such. I will make no specific proposals here.

Although requiring more in the way of a learning theory than standard parameter-setting, it doesn't seem unreasonably for a species that has managed to colonize as many different kinds of environments as humans have.

An immediate flowon question is whether UG sets any constraints on what kind of analysis can be used for an entity-identifying sequence when GR1 fails. One possible constraint is that the alternative is that they are instance of 'non-projecting words' in the theory of Toivonen (2001). These are single-word (non-phrasal) items that can only be adjoined to certain kinds of phrase-types (lexical and maximal phrasal categories,

---

[13]Ramscar et al. (2010) shows that actual prediction is more effective for learning than regenerative retrodiction, if the latter functions at all in the situations they investigated (learning irregular plurals and color terms), but so far at least I don't see how true prediction could be generally sufficient for language acquisition.

Toivonen suggests). This fits the Spanish clitics reasonably well, but we then face the further question of whether any further restrictions can be placed on them, such as for example limitation to any particular part of speech, 'pronouns', for example.

Here I suggest that the answer is 'no', based on the properties of the 'second passive' of Bahasa Indonesia, discussed for example in Sneddon (1996:248-251) and given an LFG analysis in Arka and Manning (2008). In this construction, the regular subject position is occupied by the NP that would normally appear in object position, and a single word expressing the agent appears right before the verb, but after any auxiliary, in a position where full NPs are not allowed:

(29) a. Rumah itu   akan saya jual
         house   that FUT I   sell
         I will sell the house

    b. *Buku itu   akan ayah   beli
        book that FUT father buy
        Father will buy the book

Semantically third person common nouns are excluded from this position, but common nouns can appear there if they refer to discourse participants, which is not uncommon in Indonesian:

(30) a. Buku itu   sudah   Tini kembalikan
        book that already Tini return
        I, Tini, have already returned the book

    b. Surat ini   harus bapak tandatangani
       letter this must father sign
       You(respectful) must sign this letter.

The post-auxiliary, preverbal Agent in these constructions seems to be quite a good candidate for being a nonprojecting word, presumably N or some subcategory thereof[14] some kind of N-related category. But given both the nature of the restrictions on it, and those of the Spanish pronominal clitics, it seems unlikely that there is any universal formal restriction can be placed on what kinds of words can be used to identify entities in nonprojecting word position.

Another kind of challenge comes from noun incorporation in polysynthetic languages. Often, incorporated nouns have generic reference, rather than referring to particular participants in the situations and events described in the discourse. But, this is not always the case, and even definitely referring sequences can appear incorporated, as illustrated by this verb from the Wichita text presented in Rood (1976:272):

(31) kiya- ki:-   cíwa:cé:- rhirʔa- s- ʔirhawi
     quot aorist.preverb big   buffalo inc lie
     there was the big buffalo lying there

---

[14]with the functional behavior of a 'logical subject'.

'quot' marks the verb as a reported event (in a story), the preverb normally signals the presence of a dative object or a possessed one; I'm not sure what it's doing here (*ki:* is a phonological reduction of *ki-uc*). The sequence glossed 'big buffalo' is referring to a previously introduced participant in the story; one sign that it is incorporated rather than a freestanding NP is that the order of noun and adjective is reversed from what it would be if it were freestanding,[15] another is the '-s' (inc) formative is suffixed to animate nouns ending in a vowel when they are incorporated. And at the end comes the verb stem.

If we assume that syntax learning is here supported by a morphology learner that can figure out that (31) is a single word, then there is no problem for GR1, since we can simply suppose that it only works on sequences of freestanding words. But this is not an innocent assumption, so polysynthetic languages may require more work for GR1 (and certainly do in many other areas!).

The previous three GR1 failures are all cases where the sequences in question are clearly not NPs at all; our next two will be one where the correct analysis is not entirely clear, but the kind of behavior that GR1 leads us to expect does not appear. But, unlike the possessors in Jarawara, it is the generalized exclusion principle rather than overt morphological form that ameliorates the potential problem with learnability. Both cases involve adnominal possessive constructions.

First, the celebrated example of possessors in Pirahã, either pre- or post- nominal, can only be proper names (Everett 2005:620), indicating that NPs in Pirahã are not recursive, in some sense. In exactly what sense is an issue that has caused considerable discussion. I will here propose two senses of recursion for constructions:

(32) a. A construction is recursive with respect to the whole grammar if it can contain substructures of unlimited depth (not necessarily of the same type, since this raises issues of how finely to discriminate construction types).

b. A construction is recursive with respect to some component, such as PS rules, if that component would allow it to contain structures of unlimited depth, even if the results are blocked by some other component, such as morphology.

Both conceptions are bound to theory to a certain extent, (b) only making sense relative to assumptions about what components the grammar consists of, (a) assuming that we can distinguish structural embedding from mere iteration, such as with unlimited coordination of NPs.

Pirahã has possessive constructions in which a possessor noun is juxtaposed with the possessum noun (either before or after):

(33) xipoógi hoáoíi   hi xaagá
     Xipoógi shotgun 3 be
     That is Xipoógi's shotgun (Everett 1985:205)

---

[15]Adjectives are however incorporated into the freestanding nominals they modify, but follow rather than precede the nominal.

But the possessor can only be a proper name: overtly recursive possessors such as this attempt are quite consistently rejected (Everett 2009:420):

(34) *kó'oí   hoagí kai   gáihií 'íga
     [name] son   daughter that   true
     That is Kó'oí's son's daughter (Everett 2005:630)

One consideration is that since the Pirahã are a small community where everybody knows each other, occasions for using such sentences don't arise very often.

But they do sometimes arise, and when NP structures such as (34) might be expected, one gets a 'flat' multi-clausal discourse instead:

(35)   'ísaabi kai   gáihií 'íga
        [name] daughter that   true

        kó'oí   hoagí 'asigí -aí
        [name] son   same be

        That is Ísaabi's daughter. Kó'oí son being the same.

"Here the juxtaposition shows that 'Ísaabi is Kó'oí's son" (Everett 2005:630).

GR1 wrongly predicts that (34) will be grammatical, because it indicates that the possessor in (33) should be an NP. But, clearly, learners either don't acquire an NP-within-NP analysis of possessors at all, or else acquire a restriction on possessor structure that prevents the whole grammar from being recursive here, even if the phrase structure itself or equivalent component is functioning recursively. In either case, we face the problem of explaining how this restriction can be learned (at least without heroic measures, such as complex processing of the entire corpus as in Bannard et al. (2009)). Regenerative retrodiction would appear to be a plausible solution, since, in those contexts where a learner who had a recursive possessive construction would expect structures like (34) to sometimes occur (i.e. those where (35) actually does), they don't.

An alternative, parameter based treatment is suggested by Roeper and Snyder's (2005) analysis of the restrictions on prenominal possessives in German. They observe (along with many others) that in German, prenominal possessors (the 'Saxon genitive' construction) appear to be rather restricted:

(36) a.   mein Auto
          my   car

     b.   Marias Auto
          Maria's car

     c. *meines/Marias/des Autos Motor
        my/Maria's/the   car's   motor

    d. *des   Mädchens Auto
       the(G) girl(G)   car

For cases like the latter two, one of the two postnominal possessor constructions must be used, either *von*+dative NP or genitive NP:

(37) a. das Motor vom   Auto/des Autos
         the motor from.the(D) car/the(G) car(G)
         The car's motor (finable on the web)

      b. das Auto vom   Mädchen/des Mädchens
         the car   from.the(D) girl/the(G)   girl(G)
         the girl's car

Snyder and Roeper propose a parameter which is supposed to allow such constructions to be recursive or not, with default value 'not', switched to 'yes' by the appearance of a multiword rather than single-word modifier.

In principle, this seems like a simpler and better account than what is consistent with GR1, which suggests treating these pronominal possessors as NPs, and so requires a nontrivial learning theory to pick up the implicit negative evidence that either they aren't NPs at all, or don't recurse in the manner expected of full NPs. But it has the problem discussed by Everett (2007:9) and Nevins et al. (2009:367) that more complex prenominal possessor have at least a limited degree of acceptability for some speakers as long as the nouns fall into a limited class of relational/kinship nouns. The following example provided to Everett by Manfred Krifka was for example accepted by 3 of Nevins et al.'s consultants, rejected by 7:

(38) Peters/meines Vaters Auto
      Peters/my   father's car
      my father's car, Peter's father's car

7 rejections as opposed to 3 acceptances is normally enough to justify treating a construction as substantially ungrammatical, but this isn't sufficient to save the parameter-setting account, because we still have the problem of explaining the even more general (universal, as far as I know) badness of (36c-d).

More dubious is the status of some rather complex examples that can be found on the web, such as:

(39) [der halbe FreundesKreis] meines Freundes Schwesters Exfreunds Bruder
      [half of   the   circle of   friends   of]   my

      friend's sister's ex-boyfriend's brother

      `http://www.apfeltalk.de/forum/illegale-windows-xp-t38878.html`
      (I detect a tone of jocularity in such examples, and no intent at all to actually designate any actual individual)

German speakers I have presented this too tend not to like it very much (the wrong morphology of *Schwesters* is one obvious problem), but not all completely rejected it. But any degree of acceptability at all by people who reject that examples marked bad in (36) is inconsistent with the idea of a blanket ban on recursion in the PS rules, and the details of different varieties might illuminate the nature of the actual restrictions.

Another point that might be worth mentioning is that the position taken here is agnostic about whether the absence of recursive NP structures in Pirahã is related to anything else about the language, in particular Everett's proposed 'Immediacy of Experience' principle. It is certainly consistent with the IEP, but it seems to me that this principle is not quite enough, since, after all, occasions to use (35) do occur, which would also be occasions to use (34), which is ungrammatical. So we seem to need some grammatical specification that goes beyond the actual content of the IEP, even if the principle might plausibly create good conditions for this grammatical specification to propagate itself.

7 Absence of Evidence vs. Evidence of Absence

A final topic to consider is languages in which NPs might exist, but are not motivated by any presently known evidence from those languages. Examples are provided by certain Australian languages, such as Kalkatungu (Blake 1983), and Kanyara and Mantharta (Austin 1995), where there are sequences of words that might be NPs, since they do identify participants, but no solid evidence that these sequences must be treated as constituents. This situation arises because these languages have extensive 'NP splitting', whereby the components of a description can be scattered across the sentence (interpretive reassembly supported by case-marking), but also happen to not provide any known evidence that any of these constituents are NPs, as is provided for example by the clitic-positioning system of Warlpiri (Simpson 1991).

Such languages afford cases where there are some sequences that GR1 suggests analyzing as NPs, and no real problem if they are so-analysed, but neither is there any real evidence that they should be, at least given the rather limited data available about them. But a significant point is that in many languages where more data is available, NP splitting and position turns out to be closely involved with aspects of information structure, which require either a great deal of text, or good access to native speakers to work out. See for example Legate (2002) for Warlpiri, Devine and Stephens (2006) for Latin, and Devine and Stephens (2000) for Greek. Close investigation is therefore required to tell whether GR1 is behaving well or badly when it suggests NPs in places where there isn't obvious evidence for them. If further analysis of such cases revealed a clear need for NPs, GR1 would get significant corroboration, and if it revealed that they were creating problems, it would be substantially embarrassed.

8 Conclusion

GR1 has a considerable amount of valid predictive power. Indeed, anyone with a stock-picking technique that worked as well would not have to continue working very long at any job whose properties they found to be annoying. But it does not produce *any*

Type 1 Universals, of the sort that have proved to have a rather high mortality rate, nor any tangible implicational universals. The reason is, basically, that it doesn't do anything until confronted with specific strings of words that it says ought to be NPs, and then says in effect that all of the structures that appear in any of these positions can be defeasibly expected to appear in all of them, up to details involving closed-class items, such as case-marking (since they're all NPs, and expansion of phrases is mostly context free, so if you allow one of them, you're inviting all of them). 'Defeasibly' means that some exceptions are possible, although if the rule didn't work often enough to be useful, it should be taken as being wrong, and there ought to be a reasonably straightforward way of identifying the situations where GR1 fails, and the expected NPs absent. But we don't have to specify in advance the details of how to do this, if all we're proclaiming is a guessing rule.

GR1 and GR2 can be taken as part of a 'modest UG' approach to 'theoretical linguistic typology', whereby one tries to make relatively weak assumptions about what UG consists of, that are sufficiently strong to produce definite results, and are also consistent with known typology, under relatively generous assumptions about what can be taken as 'known'.[16] As noted at the beginning of the paper, GR1 (and GR2) seem to require some kind of classic generative architecture, with sentences being produced by combinatorial processes working over a discrete finite collection of basic entities, describable in terms of rule notations (including, in principle, universal schemas with slots for parameters) It is worth pointing out that this is not an innocent assumption at all: nothing that anybody knows about neurons suggests that they should act as if they were following generative rules, but this is what they appear to do.[17]

The proposal here is modest in another sense, that it doesn't claim to account for all of language-acquisition, but only help with one small part of it. There is certainly a place for comprehensive models such as that of Pinker (1984) and the 'degree-2 learnability theory' of Wexler and Culicover (1983), but these are unavoidably committed to considerably more detail than can be well substantiated, which leads to various difficulties. I think we can at least hope that smaller scale proposals might be easier to update and combine into a more comprehensive picture. Two obvious next things to tackle after GR1 and GR 2 are:

(40) a. How is basic clause-structure acquired (especially the VP that seems to usually or even always exist in SVOX languages).

b. How are NPs that GR1 doesn't work apply to acquired.

(a) might well be dependent on (b), while an approach to (b) might come from investigating techniques for learning grammars from partially annotated corpora, where a 'lazy annotator' has picked out some but nowhere close to all NPs (but not too many

---

[16]Not, for example, requiring that a language be investigated by at least five people working for 10 years minimum on it before anything is taken as 'known' about it.

[17]Smolensky and Legendre (2006) show how this might be possible, although I'm not aware of any evidence that their proposed mechanism is actually the correct one.

false positives). This is something which computational linguists might well have already looked into, although I'm not aware of anything relevant at the moment.

References

Abney, S. P. 1987. *The English Noun Phrase in its Sentential Aspect.* PhD thesis, MIT, Cambridge MA.

Alsina, A. 1996. *The Role of Argument Structure in Grammar.* Stanford, CA: CSLI Publications.

Arka, I., and C. Manning. 2008. Voice and grammatical relations in Indonesian: a new perspective. In S. Musgrave and P. Austin (Eds.), *Voice and Grammatical Relations in Austronesian Languages*, 45–69. Stanford CA: CSLI Publications.

Austin, P. 1995. Double case marking in Kanyara and Mantharta languages, Western Australia. In F. Plank (Ed.), *Double case: Agreement by Suffixaufnahme*, 363–79. Oxford: Oxford University Press.

Baker, M. C. 2008. *The Syntax of Agreement and Concord.* Cambridge University Press.

Bannard, C., E. Lieven, and M. Tomasello. 2009. Modeling childrens early grammatical knowledge. *Proceedings of the National Academy of Sciences* 106(41):17284–17289.

Blake, B. 1983. Structure and word order in Kalkatungu: The anatomy of a flat language. *The Australian Journal of Linguistics* 3:143–175.

Boersma, P., and B. Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86.

Chomsky, N. 1964. *Current Issues in Linguistic Theory.* Mouton.

Chomsky, N. 1965. *Aspects of the Theory of Syntax.* Cambridge MA: MIT Press.

Chomsky, N. 1975. *The Logical Structure of Linguistic Theory.* Springer Verlag. published version, with some omissions, of original PhD thesis, which was circulated in various versions from 1955-56.

Culicover, P. W., and R. S. Jackendoff. 2005. *Simpler Syntax.* Oxford University Press.

Dench, A., and N. Evans. 1988. Multiple case-marking in Australian languages. *Australian Journal of Linguistics* 8:1–47.

Devine, A., and L. D. Stephens. 2000. *Discontinuous Syntax: Hyperbaton in Ancient Greek.* Oxford: Oxford University Press.

Devine, A., and L. D. Stephens. 2006. *Latin Word Order: Structured Meaning and Information.* Oxford University Press.

Dixon, R. 2000. Categories of the Noun Phrase in Jarawara. *Journal of Linguistics* 36:487–510.

Dixon, R., and A. R. Vogel. 2004. *The Jarawara language of Southern Amazonia.* Oxford and New York: Oxford University Press.

Evans, N. 1995. *Kayardild: The Language of the Bentick Islanders, with Elements of a Grammar of Proto-Tangic.* Mouton.

Evans, N. 2003. *Bininj Gun-Wok: a pan-dialectal grammar of Mayali, Kunwinjku and Kune.* Canberra: Pacific Linguistics.

Evans, N. 2007. Insubordination and its uses. In I. Nikolaeva (Ed.), *Finiteness*, 366–431. Oxford: Oxford University Press.

Everett, D. 1985. Pirahã. In D. Derbyshire and G. Pullum (Eds.), *Handbook of Amazonian languages*, Vol. 1, 200–326. Berlin: Mouton de Gruyter.

Everett, D. 2005. Cultural constraints on grammar and cognition in pirahã: Another look at the design features of human language. *Current Anthropology* 621–646.

Everett, D. 2007. Cultural constraints on grammar in pirah˜: A reply to Nevins, Pesetsky, and Rodrigues (2007). URL: `http://ling.auf.net/lingBuzz/000427` (viewed August 2010).

Everett, D. 2009. Pirahã culture and grammar: A response to some criticisms. *Language* 85.

Gil, D. 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*, 243–263. John Benjamins Publishing Company.

Graczyk, R. 2007. *A Grammar of Crow.* Lincoln NE: University of Nebraska Press.

Legate, J. A. 2008. Morphological and abstract case. *Linguistic Inquiry* 39:55–101.

Legate, J. A. 2002. *Warlpiri: Theoretical Implications.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA. URL: `http://www.ling.upenn.edu/~jlegate/` (viewed Apr 2010).

Levinson, S., and N. Evans. 2010. Time for a sea-change in linguistics: Response to comments on 'the myth of language universals'. *Lingua* 120.

Link, G. 1998. *Algebraic Semantics in Language and Philosophy.* Stanford CA: CSLI Publications.

Lyons, J. 1968. *Introduction to Theoretical Linguistics.* Cambridge University Press.

Mithun, M. 2008. The extension of dependency beyond the sentence. *Language* 81:69–119.

Nevins, A., D. Pesetsky, and C. Rodrigues. 2009. Piraha exceptionality: a reassessment. *Language* 85:355–405.

Nordlinger, R. 1998. *Constructive Case*. Stanford CA: CSLI Publications.

Partee, B. H. 2006. Do we need two basic types. URL: `http://people.umass.edu/partee/docs/`.

Pinker, S. 1984. *Language Learnability and Langage Development*. Harvard University Press.

Plank, F. 1995. *Double Case: Agreement by Suffixaufnahme*. Oxford University Press.

Ramscar, M., D. Yarrett, M. Dye, K. Denny, and K. Thorpe. 2010. The effects of feature-label-order and their implications for symbolic learning. *Cognitive Science* 34:909–957.

Roeper, T., and W. Snyder. 2005. Language learnability and the forms of recursion. In A. M. D. Sciullo (Ed.), *UG and external systems: Language, brain and computation*, 155–169. Amsterdam: Benjamins.

Rood, D. S. 1976. *Wichita Grammar*. New York and London: Garland Publishing Company.

Ross, J. R. 1967. *Constraints on Variables in Syntax*. PhD thesis, MIT, Cambridge MA. distributed by the Indiana University Linguistics Club; revised version publised as Ross (1986).

Ross, J. R. 1986. *Infinite Syntax!* NJ: Norwood. revision of Ross (1967).

Round, E. 2009. *Kayardild morphology, phonology and morphosyntax*. PhD thesis, Yale. URL: `http://www.ling.yale.edu/~erich/Home_files/Round2009PhDDissertation.pdf`.

Simpson, J. H. 1991. *Warlpiri Morpho-Syntax*. Kluwer Academic.

Smolensky, P., and G. Legendre. 2006. *The Harmonic Mind (Volume I: Cognitive Architecture)*. Cambridge MA: MIT Press.

Sneddon, J. N. 1996. *Indonesian: a Comprehensive Reference Grammar*. London and New York: Routledge.

Toivonen, I. 2001. *The Syntax of Non-Projecting Words*. PhD thesis, Stanford University, Stanford CA.

Wexler, K., and P. W. Culicover. 1983. *Formal Principles of Language Acquisition*. Cambridge MA: MIT Press.

Yoder, B. 2010. Syntactic underspecification in Riau Indonesian.