Ling 507
Dan Jinguji
Final Paper

Empty Categories as Subjects of Finite Clauses
A case of *pro*-drop — or not

Principles and Parameters

The modern study of syntax is highly influenced by the concept of universal grammar, that there exists a common, innate structural framework underlying the grammar of the languages of the world. This is not a new idea. It can be traced back to Wilhelm von Humboldt (18th century), René Decartes (17th century), Roger Bacon (13th century), and Platonic, Socratic, and Pythagorean philosophy of ancient Greece. One of the primary arguments for this innate grammatical framework comes from observations of first language acquisition (L1A) in children. Noam Chomsky coined the term "poverty of the stimulus" to describe the situation of children acquiring native proficiency in their mother tongue in spite of the fragmented information given them about the grammar.

One of the hallmarks of current linguistic thought regarding universal grammar is the Principles and Parameters approach published by Chomsky in *Lectures on Government and Binding*. In this approach, the characteristics of universal grammar can be segmented into two groups, those that are invariant and those that may vary from language to language. The invariant characteristics are termed "principles". These can be found in all languages. The aspects of structural variety in language are not thought to vary freely, but rather, the variation is based in choices made from a limited set of alternates provided within the innate grammar postulated by this approach. These choices, then, allow for observed structural diversity found in the world's languages. These choices are termed "parameters" since they are settable values within the grammar of a given language. Some quick examples may help to clarify.

There is a principle that language structure is hierarchical, that is, that the elements in an utterance can be grouped into logical structures that contribute to the meaning of the utterance. This can be viewed at many levels, for instance that the intrinsically meaningless sounds within the utterance can be grouped into meaningful morphemes, that these morphemes can be grouped into separate words, that words can be grouped into phrases, and that phrases can be grouped into larger phrases. Moreover, in general, the meaning of the utterance can be viewed compositionally from the meanings of the constituents.

As an example of a parameter, the basic order of these phrasal constituents can be shown to vary radically between languages. Taking a simple declarative sentence with a transitive verb, we can identify three constituents: an action (the primary semantic content, called "verb"), an actor (the performer of the action, called "subject"), and the theme (the one who undergoes the action, called "object"). These three elements can be ordered in 6 different configurations: VSO, VOS, SOV, SVO, OSV, and OVS. Languages can be categorized by the predominant ordering for these elements in the utterances of that language into one of these 6 types, or a seventh type which allows free variation in the ordering. This does not imply that all declarative sentences will demonstrate this order, since it may be altered by other factors. However, it provides a structural bulwark for language acquisition and language understanding. Given the idea that this is parametric, once the child sees this predominant pattern in the ambient language, the

parameter value can be set and the child's grammar grows in its conformity to that of the rest of the speech community. Additionally, it has been observed that acquisition of this parametric knowledge aids in the understanding, and therefore, acquisition, of other aspects of the grammar as well, for example, in this case, the lexicon. The thematic relationship of the constituents of the utterance clarifying the referents for unfamiliar nouns and verbs, based on their location relative to the more familiar ones.

Another example of a parameter is the necessity for an overt subject within a sentence, the *pro*-drop parameter. There are a number of languages where the subject is obligatory, even if recoverable from context, for example English, "I have seen it already." Alternately, in Spanish, the subject pronoun can be dropped, "Ya lo vi." In fact, in Japanese, both of the pronouns may be omitted, "Mō mita." These are examples of *pro*-drop languages. In fact, in both of these languages, the norm is to drop the pronoun(s). It is important to note that the *pro*-drop parameter addresses more phenomena than just the omission of subject pronouns. It also licenses the "displacement" of the overt subject to after the verb in SOV languages like Italian "È arrivato Paolo," and Spanish "Llego Pablo," but not English "Paul arrived." So, the parameter represents a cluster of grammatical features / structures, rather than a rule that licenses a single type of construction.

The Principles and Parameter approach addresses the issue of poverty of the stimulus. As one can see from the multiple actions of a single parameter setting, the triggering of a parameter setting during language acquisition can account for a number of changes in the surface grammar of the child. Not only does this help minimize the amount to be apprehended during L1A, it also parallels the observed development of grammar in young children as they learn language.

Van Valin's paper

It is commonly attested that English does not permit the empty category as the subject of a tensed (finite) clause. (1)

(1) * *e* left.
* That *e* was arrested shocked everyone.

In *An Empty Category as the Subject of a Tensed S in English*, Robert Van Valin argues that the empty category can appear as the subject of a tensed clause in English. In support of this, he offers several examples. Here are two of them. (2)

(2) a. John talked to Mary today and will ask her for a date tomorrow.
b. The robber just ran out of the bank and will be arrested by the police any moment now.

He continues by observing that these are sentences with coordination. The question he raises is the level at which the coordination occurs. Two obvious analyses are coordinated VPs or coordinated IPs. For the purposes of this initial discussion, (2a) will be used as the example.

Van Valin initially rejects the analysis of coordinated IPs since it results in an empty position for the Spec of the second IP. (3)

(3) [$_{IP}$ John talked to Mary today] and [$_{IP}$ *e* will ask her for a date tomorrow.]

The EPP requires a subject for both of the finite clauses. This would lead to positing an empty category for the subject of the second clause, contrary to the attested stance.

Van Valin continues by exploring the possible analysis of this sentence as coordinated VPs. The conjoined VP structure could then share the NP "John" in the [Spec, IP] position. (4)

(4) [$_{IP}$ John [$_{I'}$??? [$_{VP}$ [$_{VP}$ talk to Mary today] and [$_{VP}$ ask her for a date tomorrow.]]]]

This leads to a "collision" at the Infl position. It must carry both [+past] for the first clause and [–past] for the second. (It could be argued that "will" serves as an auxiliary here, but that is outside the scope of this discussion.) This leads Van Valin to conclude that the first analysis is correct, that the structure is coordinated IPs, with an empty category as the subject for the second IP.

He continues to bolster this analysis by considering θ-role assignment. The Theta Criterion provides for each argument to be supplied exactly one theta role. In the coordinated VP analysis, it appears that [Spec, IP] receives the AGENT role twice, once from each tensed verb. In (2b), the problem is heightened. The subject "the robber" receives an AGENT role from the first verb and a THEME role from the second verb since it is passive. (5)

(5) [$_{IP}$ [$_{NP}$ The robber]$_{AGENT}$ [$_{I'}$ just ran out of the bank]] and
[$_{IP}$ e_{THEME} [$_{I'}$ will be arrested by the police at any moment.]]

Once again, it seems that the coordinated IP structure provides a more satisfactory analysis. The empty category provides an independent recipient for the second θ-role.

Van Valin then asks about the nature of the empty category. He posits that it cannot be PRO since it is the subject of a finite clause. He rejects it being a trace or variable since there is a lack of Move α in this sentence. Therefore, he concludes is must be *pro*. This is consistent with binding Principle B, since it must be free in its governing category, which it is. It seems then, at the surface, that we have an example of *pro*-drop in English.

However, Van Valin then notes that this *pro* is odd in that it must be controlled, unlike an overt pronoun one might place in that location. For this discussion, he alters (2a) slightly.

(6) a. John$_i$ talked to Bill and e_i will ask Mary for a date.
b. John$_i$ talked to Bill$_j$ and he$_{i/j/k}$ will ask Mary for a date.

In sentence (6a), the empty category must be coindexed with "John", whereas in (6b), the overt pronoun "he" can refer to "John" or "Bill" or potentially some third masculine individual not otherwise mentioned in the sentence. Thus, the empty category, in this case, is obligatorily controlled.

Is there an alternate analysis for this data? Yes. What are the options? This empty category may indeed represent ellipsis, the process by which portions of an utterance can be omitted when recoverable from context. One common example of this context is discourse.

(7) Q. Who can play the guitar?
A. (Not) John.

Within this context (7), it is clear the answer is "John can play the guitar," in the case of the positive answer. It is interesting to note that the negative answer shows negation of the overt NP, even though, in the underlying IP, it is the verb that is negated "John cannot play the guitar."

Another context (8) for ellipsis is coordination. Between the coordinands, common verbiage is frequently omitted.

(8) a. John can play the guitar and Mary can, too.
b. John can play the guitar, and Mary, the violin.
c. John wrote with a pencil, and Mary, a pen.
d. Few dogs eat Whiskas or cats Alpo.

Here are the same constructions as in (8), but without ellipsis. (9)

(9) a. John can play the guitar and Mary can play the guitar, too.
b. John can play the guitar, and Mary can play the violin.
c. John wrote with a pencil, and Mary wrote with a pen.
d. Few dogs eat Whiskas (and) few cats eat Alpo.

These examples go from the more straightforward to more complex. In (8a), the VP "play the guitar" is omitted, a single constituent. In (8b), the modal "can" and the verb "play" are omitted. Notice that this no longer is a simple constituent, though in "traditional" English grammar, this would constitute the complete verbal portion of the predicate. Example (8c) shows even more disparate pieces of sentential structure elided between the coordinands; in this case, both the verb and the preposition which introduces a PP. In (8d), the elided material is disjoint; both the determiner for the subject NP and the verb are elided. It is interesting to note that modifier scoping rule prompt a change in the coordinating conjunction to retain the same sense. (This, too, is without the scope of this paper.)

Within the context of ellipsis, sentences (2a) and (2b) can be handled relatively easily. (10)

(10) a. John$_i$ talked to Mary today and (John)$_i$ will ask her for a date tomorrow.
b. The robber$_i$ just ran out of the bank and (the robber)$_i$ will be arrested by the police any moment now.

We have two complete IPs which are coordinated. Their coordination licenses ellipsis of the common verbal (word) material, in these cases the subject in the second clauses. Thus the analysis is Move and Merge[1]. Move the common material to the first coordinand and Merge the common material. (11)

(11) a. John$_i$ talked to Mary today and t_i will ask her for a date tomorrow.
b. The robber$_i$ just ran out of the bank and t_i will be arrested by the police any moment now.

The empty category is then *trace*. This is mediated by the presence of the conjunction. This is curiously similar to the suggestion by Daniel McCloy that the conjunction is serving the role of filling the empty category.

The in-class presentation also examined briefly a quote from Rosa Parks: (12)

(12) Memories of our lives, of our works, and our deeds will continue in others.

Ellipsis in this construction will handle the apparent non-parallelism in this coordinated structure. It is interesting to note that semantically this seems to be coordinated PPs even where the surface analysis might suggest coordination at the NP or N level. (13)

[1] This is a guess on my part regarding analysis of ellipsis across conjunct structures within the P&P approach.

(13) a. Memories of our lives, our works, and our deeds will continue in others.
b. Memories of our lives, works, and deeds will continue in others.
c. Memories of our lives, our works, and deeds will continue in others.

In fact, one might argue that semantically the coordination is occurring at the level of the higher NP. (14)

(14) Memories of our lives, memories of our works, and memories of our deeds will continue in others.

This is more clearly illustrated in (15).

(15) She likes books about history and gardening.

The coordination here, at least in the most common interpretation, is at the level of the argument for "like". (16)

(16) She likes books about history and books about gardening.

To get the other read one must specify the scope for coordination. (17)

(17) She likes books that are about both history and gardening.

Van Valin vs. Parameter

As noted in the earlier discussion, at the surface Van Valin's examples seem to be cases of *pro*-drop within English. This would be a revolutionary claim, changing one of the most commonly cited parameter setting for English[2].

While Van Valin does associate the empty category with *pro*, he also notes that this is an odd *pro* since it must be controlled within its sentential context. This leads his discussion into a blending of *pro* and PRO as a means to explain this phenomenon. He closes his article with a rearticulation of the finding of an empty category as the subject of a finite clause, conceding that its exact nature is unclear.

Within Van Valin's conclusion, there is strong evidence that he is not proposing this as an example of *pro*-drop within English. As he notes, a too-hasty reanalysis of *pro* and PRO would license patently non-grammatical sentences such as those in (1). That is, it would lead to the conclusion that English is a *pro*-drop language.

Van Valin, instead, proffers coordinated IPs as a conditioning environment in which we see the empty category in [Spec, IP] for [+finite] Infl. One important distinction to make is the limited scope of this phenomenon. We see similar examples of the empty category as the subject of finite clauses in some particular registers. In *Introduction to Government and Binding Theory*, as Exercise 3 in the Introduction chapter, Liliane Haegeman offers several examples of empty categories as the subject for finite clauses. She notes that it is restricted in register to what she terms "abbreviated writing". She continues to show that the subject pronoun can be dropped in limited circumstances, even within this register. (18)

(18) a. Brilliant (—) could have stayed all day.
b. Could see everything from wheelchair.

[2] This statement should probably be limited to Modern English. Middle English is questionable, but Old English seems quite likely to be a *pro*-drop language.

c. I must work as *(I) told Sally G…
d. What shall *(I) write …
e. The next book *(I) think of calling *Answers to Correspondents* …
f. Never have *(I) worked so hard at any book …

The "fragileness" of this phenomenon is indicative of its nature as a stylistic element rather than as a parametric setting with far-reaching effects. In fact, not only is the phenomenon observable in quite limited environments, one cannot see any of the other manifestations of the *pro*-drop parameter setting in English.

This points to one of the cornerstones of research into parametric settings in language. In identifying the parametric settings for various languages of the world, one has the opportunity to discern those recurring patterns. This can enhance our understanding of universal grammar, allowing us to combine those areas where the current understanding does not see the interrelatedness of existing parameters, or to tease apart into separate parameters those which are currently incorrectly coalesced. However, every living language is always reforming itself. It is not inconceivable that this reformation could involve changes in parametric settings for the language in question. Does linguistics have the power to properly identify those intermediate states as such or will serious consideration of those intermediate language states misdirect the search for the fundamental nature of universal grammar. Ultimately, the quest for a definitive description of universal grammar may prove quixotic.

References

Cartesian linguistics. (2006, September 7). In *Wikipedia, The Free Encyclopedia*. Retrieved December 12, 2006, from http://en.wikipedia.org/wiki/Cartesian_linguistics.

Citko, B. (2006). *Missing Labels*. manuscript.

Citko, B. (2006). personal communication.

Elliptical construction. (2006, October 23). In *Wikipedia, The Free Encyclopedia*. Retrieved December 12, 2006, from http://en.wikipedia.org/wiki/Elliptical_construction.

Haegeman, L. (1994). *Introduction to Government and Binding Theory, second edition.* Oxford, England; Malden, MA: Blackwell Publishers, Ltd.

Kimball, S., (2001). *Basics of Reed-Kellogg diagrams*. Retrieved December 12, 2006, from University of Texas, E360K English Grammar: Dr. Kimball Spring 2001: http://www.utexas.edu/courses/langling/e360k/handouts/diagrams/diagram_basics/basics.html.

McCloy, D. (2006). personal communication.

Mitchell, R., & Myles, F. (2004). *Second Language Learning Theories, second edition.* London, England: Hodder Arnold.

Modistae. (2006, December 10). In *Wikipedia, The Free Encyclopedia*. Retrieved December 12, 2006, from http://en.wikipedia.org/wiki/Modistae.

Poverty of the stimulus. (2006, December 12). In *Wikipedia, The Free Encyclopedia*. Retrieved December 12, 2006, from http://en.wikipedia.org/wiki/ Poverty_of_the_stimulus.

Universal grammar. (2006, December 6). In *Wikipedia, The Free Encyclopedia*. Retrieved December 12, 2006, from http://en.wikipedia.org/wiki/Universal_grammar.

Van Valin, R.D., Jr. (1986). An Empty Category as the Subject of a Tensed S in English. *Linguistic Inquiry, 17,* 581–586.