

INTD0111A/ARBC0111A

The Unity and Diversity of Human Language

Lecture #5
Sept 26th, 2006

Announcements

- Assignment #1 is posted online. It's due on Tuesday October 3 at 4:15pm in class (recommended), or by e-mail no later than 5pm.
- Policy on delay in turning in assignments:
 - 5% off if turned in after the deadline on the day it's due (that means prior to midnight).
 - 10% off if turned in on the next day after the deadline.
 - 20% off if turned in later than that.
 - Not accepted after I post the solutions (I know this is self-evident, but just in case)

Announcements

- For your Language Adoption Project (LAP), you may want to have a look at "*The world atlas of language structures*", which is available in the main library at the Main Atlas Case [G1046.E3 W6 2005](#).
- Should you fall in love with one particular language there, you'll have to start looking for a descriptive grammar of that language, and if it's not available in our library, then order it through the inter-library loan service.
- I'll be posting links to lists of some descriptive grammars for different languages pretty soon on the course website.

Summary of where we are

- The formalist approach to unity and diversity in human language is based on the main assumption that humans are born with an innate abstract system of linguistic knowledge which allows them to acquire their native languages with such ease and in such a short period of time. This abstract system is what we called **UG** (for *Universal Grammar*, in case you've forgotten).

UG: principles and parameters

- UG has two components: **principles** and **parameters**. The principles are invariant; they hold in all languages. Parameters are also universal, but unlike principles, they come in the form of (typically binary) options, and this is where the locus of cross-linguistic variation exists.

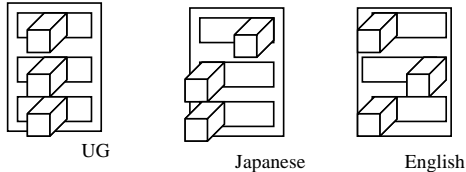
UG: principles and parameters

- As Chomsky notes:

"We can think of the initial state of the faculty of language as a fixed network connected to a switch box; the network is constituted of the principles of language, while the switches are the options to be determined by experience. When the switches are set one way, we have Swahili; when they are set another way, we have Japanese. Each possible human language is identified as a particular setting of the switches—a setting of parameters, in technical terminology."

UG: principles and parameters

- Or, as Paul Haegstrom represents this graphically:



UG: principles and parameters

- Under this approach, a child's job is to "set" the value of each parameter on the basis of the primary linguistic data (PLD) around him/her.
- This should answer Mrs. Advocate's question from last time: If you're born in Tokyo, then your PLD are different from the PLD of someone born in Montreal, hence the acquired system will be different.
- Language acquisition is thus the result of interaction between nature (principles and parameters) and nurture (PLD).

One UG principle: structure-dependency

- One UG principle that we discussed quickly last time is structure-dependency, which states that rules have to make reference to sentence structure (and not to linear order for example).
- Consider how a child can learn the rule for yes-no question formation in English on the basis of the PLD. Here's a couple of sentences in the input:
 - John must leave.
 - Must John leave?

One UG principle: structure-dependency

- Hypothesis #1 (structure-independent): Invert the first word and the second word to form a yes-no question.
- Does it work?
 - Well, let's expand the PLD space:
 - This boy must leave.
 - *Boy this must leave?
- Something went wrong here, and it wasn't the child's error. Children never make these mistakes.

One UG principle: structure-dependency

- Hypothesis #2 (structure-independent): Move the auxiliary verb to the front to form a yes-no question.
- Does it work?
 - The boy should have left.
 - ✓Should the boy have left?
- But:
 - *Have the boy should left?
- So, the modified rule may generate ill-formed questions. Can we do better?

One UG principle: structure-dependency

- Hypothesis #3 (structure-independent): Move the first auxiliary verb to the front to form a yes-no question.
- Does it work? How about this:
 - The boy who must leave has been sick.
 - *Must the boy who leave has been sick?
- This is not English, obviously. So, something went wrong again.

One UG principle: structure-dependency

- Hypothesis #4 (structure-dependent): Invert the auxiliary verb of the whole sentence and its subject to form a yes-no question.
- Does it work?
 - The boy who must leave has been sick.
 - Has the boy who must leave been sick?
- That worked.

One UG principle: structure-dependency

- As it turns out, children never produce any of the bad forms above. Why?
- Because hypotheses like 1, 2 and 3, are not even considered. Why?
- Because they are not structure-dependent. Structure-dependency is a universal principle of grammar, and as such constrains language acquisition by children.

Ok, what's a parameter then?

Can you give us an example?

One UG parameter: The null subject parameter

- Consider these data from English, French, and Italian, all of which allow SV orders:
 - (1) John will leave.
 - (2) Jean arrivera. French
 Jean will-arrive
 - (3) Gianni verrá. Italian
 Gianni will-come.

One UG parameter: The null subject parameter

- Italian, however, allows the subject of a tensed sentence to be omitted, an option that is not available in English or French:
 - (5) *Will leave.
 - (6) *Arrivera. French
 will-arrive
 - (7) Verrá. Italian
 will-come.

One UG parameter: The null subject parameter

- This is an example of *parametric variation*, which Baker formulates as follows: "In some languages (e.g., French, English, Edo) every tensed clause must have an overt subject. In other languages (e.g., Italian, Spanish, Romanian, Navajo, Arabic) tensed clauses need not have an overt subject."
- This case of cross-linguistic variation is typically referred to as the **null subject (NS) parameter**.

One UG parameter: The null subject parameter

- The children's task in acquiring their language is to "set" the parameter value on the basis of the PLD in their linguistic environment.
- The interesting thing about the null subject parameter is that it also explains to us a "cluster" of differences between these two "types" of languages.

One UG parameter: The null subject parameter

- For one thing, unlike English and French, Italian also allows VS orders:
 - (8) *Will leave John.
 - (9) *Arrivera Jean. French
will-arrive Jean
 - (10) Verrá Gianni. Italian
will-come Gianni.

One UG parameter: The null subject parameter

- Similarly, an overt "dummy" subject with "weather verbs" is required in both English and French, as opposed to its absence in Italian:
 - (11) It is raining.
 - (12) Il pleut. French
it rains
 - (13) Piove. Italian
Is-raining.

One UG parameter: The null subject parameter

- Also, subject wh-questions that, as we've seen before, are disallowed in English (and French) when the embedded clause has "that", are fine in Italian:
 - (11) *Who did you say that — saw Chris in the park?
 - (12) *Qui veux-tu que — épouse Jean?
who want-you that marries Jean?
 - (13) Chi crede che — verrá?
who you-think that will leave

One UG parameter: The null subject parameter

- As Baker notes, there are three reasons why linguists think this "cluster" of linguistic properties are interrelated:
- First, Spanish and Romanian behave exactly like Italian with regard to the three properties above, so it cannot be an accident that such properties occur or do not occur together in natural languages.

One UG parameter: The null subject parameter

- Second, Old French was exactly like Italian, i.e., it had all the properties discussed above. At one point in its historical development, though, French became just like English, i.e., it lost all these three properties. This makes sense only if there is a correlation in the occurrence or lack of occurrence of such syntactic properties.

One UG parameter: The null subject parameter

- Third, the properties are indeed interrelated, since they all involve the notion “subject of the sentence” in one way or another. Without getting into technical detail, somehow the on-setting of the NS parameter allows subjects to behave in a way that is not possible when the parameter is assigned an “off” value.

One UG parameter: The null subject parameter

- The parametric approach thus seems promising: Not only does it tell us why languages differ with regard to a particular property, but it also ties together what seem to be (at least on the surface) a set of unrelated linguistic phenomena.

Remember word order correlates in English and Japanese?

- Maybe it's time to revisit the word order correlates from last time, and see what the principles and parameters framework can say about them.
- Here's the data again, followed by the table for word order correlates that we observed earlier:

Remember word order correlates in English and Japanese?

- English:
The child might think that she will show Mary's picture of John to Chris.
- Japanese:
Taroo-ga Hiro-ga Hanako-ni zibun-no
Taroo-SU Hiro-SU Hanako-to self-POSS
syasin-o miseta to omette iru
picture-OB showed that thinking be
“Taro thinks (literally, is thinking) that Hiro showed a picture of himself to Hanako.”

Remember word order correlates in English and Japanese?

Element A	Element B	English	Japanese
Verb	Direct Object	A precedes B	A follows B
Verb	Pre-/post-position phrase	A precedes B	A follows B
Verb	Embedded Clause	A precedes B	A follows B
Pre-/post-position	Related Noun Phrase	A precedes B	A follows B
Noun	Pre-/post-position phrase	A precedes B	A follows B
Complementizer	Embedded Clause	A precedes B	A follows B
Auxiliary	Main verb	A precedes B	A follows B

Introducing some syntax

- There's a couple of things one can notice here about the A and B elements in the table. Can you tell what these are?
- For one thing, all the A elements are single words (e.g., a verb, a pre- or post-position, a complementizer, etc.), whereas the B elements could be as simple as a single word or as complex as a phrase or a clause.

Introducing some syntax

- The second difference has to do with meaning. There's a sense in which the B elements are there by virtue of the A elements. For instance, a verb like "rent" requires an object, whereas a verb like "sleep" does not:
 - *John rented.
 - John rented a car.
 - John slept.
 - *John slept the bed.

Introducing "heads", "complements", and "phrases"

- To capture the asymmetry between the A and B elements, we call each A element a "head" and each B element a "complement".
- A head plus its complement form a *phrase* of the head's kind. For example, a Verb (V) plus a noun phrase (NP) complement form a verb phrase (VP).

Introducing PSRs

- We express phrasal structure in the form of rewriting rules, called *phrase structure rules* (PSRs), for an obvious reason.
- For example, the PSR for VPs in English is expressed as follows:
 $VP \rightarrow V NP$

More PSRs in English

- And we can do the same for the rest of the A and B elements in English:

VP \rightarrow V PP
VP \rightarrow V CP
PP \rightarrow P NP
NP \rightarrow N PP
CP \rightarrow C S
AuxP \rightarrow Aux VP

where PP = Prepositional Phrase; CP = Complementizer Phrase; S = Sentence; C = Complementizer; Aux = Auxiliary.

PSRs for Japanese

- How about the PSRs for Japanese then? These should be similar to the English PSRs above, except for the position of the head within the phrase.
- So, one rule for VP rewriting in Japanese is as follows:
 $VP \rightarrow NP V$

PSRs for Japanese

- And the same applies to the rest of the rules:

VP \rightarrow PP V
VP \rightarrow CP V
PP \rightarrow NP P
NP \rightarrow PP N
CP \rightarrow S C
AuxP \rightarrow VP Aux

The head directionality parameter

- The difference between English and Japanese thus comes down to the “directionality” of the head within the phrase: left in English, right in Japanese.
- Abstracting away from the type of categories involved in the rules, we can express the difference between English and Japanese phrase structure in two simple rules:
 - XP → X complement (English)
 - XP → complement X (Japanese)

The head directionality parameter

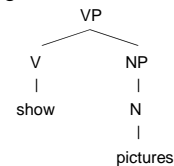
- This is another instance of parametric variation in human languages, which we can state as follows:
 - Heads occur initially (i.e., before their complements) or finally (i.e., after their complements) within phrase structure.
- And this is the so-called *head directionality (HD) parameter*.

The head directionality parameter

- The head-initial setting of the HD parameter holds in English, Edo, Thai, Khmer, Indonesian, Zapotec and Salish, while the head-final setting holds in Japanese, Lakhota, Turkish, Basque, Navajo, the languages of the Eskimos, and Quechua.

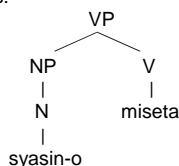
Some tree-drawing fun

- Graphically, a visually convenient way of representing phrase structure is by means of a tree diagram, as shown below for the VP “*show pictures*” in English:



Some tree-drawing fun

- The structure of the same VP in Japanese, by contrast, will look like this:



- To push the family tree analogy further, let’s call VP the “mother” node, and NP and V its “daughters”. And yes, you can talk about NP and V as being “sisters”.

So, where’s the subject in the tree?

- Notice that the HD parameter does not say anything about the position of subjects, since these are not complements. Is this good or bad?
- It’s actually good, since English and Japanese are both subject-initial. We don’t want to parameterize that.
- Let’s assume then that each sentence is of category “S”, whose daughters are a subject NP and an AuxP, and add this to our PSRs:
 - S → NP AuxP

A mini-grammar for English phrase structure

- Adding all our rules so far together, we have the following mini-grammar for English:

$S \rightarrow NP \text{ AuxP}$
 $\text{AuxP} \rightarrow \text{Aux VP}$
 $\text{VP} \rightarrow V \text{ NP}$
 $\text{VP} \rightarrow V \text{ PP}$
 $\text{VP} \rightarrow V \text{ CP}$
 $\text{NP} \rightarrow N \text{ PP}$
 $\text{PP} \rightarrow P \text{ NP}$
 $\text{CP} \rightarrow C \text{ S}$

A mini-grammar for English phrase structure

- In addition, we have to assume a set of rules that insert words under "terminal" nodes in the tree, e.g.,
 - $N \rightarrow \{\text{man, dog, justice, ...}\}$
 - $V \rightarrow \{\text{love, hit, leave, ...}\}$
 - $\text{Aux} \rightarrow \{\text{will, must, Past, Future, ...}\}$
 - etc.
- As you should expect, these are called *lexical insertion* rules.

A mini-grammar for Japanese phrase structure

- Similarly, the mini-grammar for Japanese phrase structure will have the following rules as well as a set of lexical rules for Japanese words:

$S \rightarrow NP \text{ AuxP}$
 $\text{AuxP} \rightarrow \text{VP Aux}$
 $\text{VP} \rightarrow NP \text{ V}$
 $\text{VP} \rightarrow \text{PP V}$
 $\text{VP} \rightarrow \text{CP V}$
 $\text{NP} \rightarrow \text{PP N}$
 $\text{PP} \rightarrow \text{NP P}$
 $\text{CP} \rightarrow \text{S C}$

So, why do English and Japanese look dramatically different then?

- Now, let's try to make things more interesting and see how and why English and Japanese do really look dramatically different on the surface.

So, why do English and Japanese look dramatically different then?

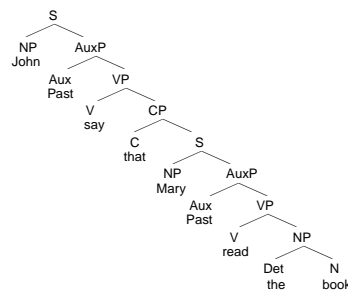
- Compare English and Japanese again:

John said that Mary read the book.

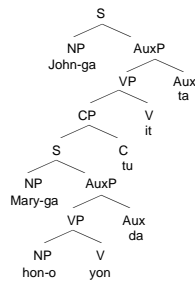
John-ga Mary-ga hon-o yon-da-tu it-ta
 John-SU Mary-OB book-OB read-past-comp say-past

- Given the PSRs for both English and Japanese, the structural trees will look as follows:

First: English



Second: Japanese



So, ...

- The principles and parameters approach thus accounts for word order correlates in SVO and SOV languages in a straightforward manner.
- Notice also how a simple difference in head directionality leads to a dramatic variation on the surface, due to its cumulative effect on all heads and complements in a language.
- In addition, since the HD parameter has two settings only, it predicts two types of languages, SOV and SVO, which is exactly what we find in language samples: these two orders represent about 90% of human languages.

Japenenglish!

- But equally important, the HD parameter also correctly predicts the non-existence or at least the rarity of Japenenglish-type languages, i.e., languages in which the verb precedes the object but that are also postpositional, or languages in which the verb follows the object but that are also prepositional, (in contradiction with Greenberg's *Universal 4*).

Japenenglish!

- As Baker notes, in such languages we expect to find structures like this:
 - Chris put the book the table on.
 - Chris the book on the table put.
- But Japenenglish-type languages are very rare, if existent (but see assignment #1). This is good news for the parametric approach since Japenenglish is predicted to be an unattested human language under this approach.

Some work left to be done, though:

- How do we explain the existence of other basic word orders in human languages? Particularly, how does the principles and parameters approach account for the existence of VSO languages (about 9% of human languages per Tomlin's 1986 sample).
- And how about the "rare" types: VOS, OSV, and OVS. Why do these even exist?
- Even more seriously, how about these languages that do not seem to impose restrictions on word order, e.g., Mohawk or Warlpiri? How are these possible human languages?
- We discuss all these questions and others next time. So, see you then!