

## Word order and parameters in multiple languages

### DO ALL LANGUAGES HAVE CONSTITUENT STRUCTURE?

Constituent structure = a tree structure that can be (mostly) derived from a context-free grammar

Warlpiri (Australia):

These-S big-O children-S chased those-O small-S dogs-O.

or any variation thereof.

(S = subject, O = object)

Other languages that may be free word order languages: Latin, Hungarian.

### DIFFERENT WORD ORDERS IN CONSTITUENT-STRUCTURE LANGUAGES

Navajo:

Ashkii at'eed yiyiiltsa.  
Boy girl saw

"The boy saw the girl."

Japanese:

John-ga Mary-o butta.  
John-S Mary-O hit

"John hit Mary."

Both Navajo and Japanese are SOV languages.

English is an SVO language.

These are 2 of the 6 cases in Greenberg's typology.

SVO	42%	English, Edo (Nigeria), Indonesian
SOV	45%	Japanese, Turkish, Quechua
VSO	9%	Zapotec, Welsh, Niuean (Pacific)
VOS	3%	Tzotzil, Malagasy
OVS	1%	Hixkaryana (S. America)
OSV	0%	Warao (S. America)

Why is Greenberg's typology interesting?

- Parsing
- Relation to other language phenomena

## MORE WORD-ORDER PROPERTIES

Prepositions vs. postpositions:

Navajo:

ee            biih    naasdza  
clothing into I-got-back

“I got back into (my) clothes.”

Japanese:

John-ga Mary to kuruma da Kobe ni iita.  
John-S Mary with car by Kobe to went

“John went to Kobe by car with Mary.”

Both languages have PP --> NP prep.

English has PP --> prep NP.

(Therefore, in Japanese we should refer to “postpositions” instead of prepositions.)

Genitive order:

Navajo:

Chidi        bi-jaad.  
Car         its-leg

“the wheel of a car”

Japanese:

John-no imooto-ga sinda.  
John-'s sister-S died

“John's sister died.”

In both languages, the possessor always precedes the possessed object.

Edo (Nigeria), a language not historically related in any way to English, also has SVO word order:

John mien Mary.  
John found Mary

“John found Mary.”

Edo also has prepositions:

John rhie nene ebe ne Mary.  
John gave the book to Mary

“John gave the book to Mary”

In Edo, the possessor follows the possessed object:

Omo John rre.  
child John come

“A child of John’s came.”  
“John’s child came.”

More of Greenberg’s universals:

Languages with dominant VSO order have prepositions.

Almost all SOV languages have postpositions.

In languages with prepositions, the possessor almost always follows the possessed object.

In languages with postpositions, the possessor almost always precedes the possessed object.

So it looks like you can’t just choose random grammar rules and get a human language.

#### ALTERNATE APPROACHES TO CASE MARKING

How do you know which noun is the subject of the sentence and which is the object?

Alternatively, how do you know which object is the agent (actor) of the proposition and which is the “patient” (Latin: ‘the one who suffers’)?

(Side note: Thus, in the current movie, “The Passion of ...”, ‘passion’ has the older English meaning of ‘suffering’, not the current meaning of ‘feelings’.)

#### AN ASIDE ON SENTENCES VS. PROPOSITIONS

Note 2 different meanings of ‘object’ here: one as grammatical term and one as a synonym for ‘thing’.

Also note the difference between ‘sentence’ and ‘proposition’. A sentence is a string of words, and a proposition is a logical expression.

A sentence can represent a proposition. An ambiguous sentence represents more than one proposition.

A proposition can be true or false.

Is it possible to have a sentence that doesn’t represent a proposition at all? What about Chomsky’s famous sentence, “Colorless green ideas sleep furiously”?

What kind of proposition does a question (“Who ate my cheese?”) represent?

What kind of proposition does an imperative sentence (command) represent?

Sometimes we need context to know what proposition a sentence represents:

Example 1: Yes.

Example 2:

A: I would like you to do my grocery shopping for me today.

B: OK...

A: I would like you to do my grocery shopping for me today.

B: OK!

So are we talking about case marking, i.e. a property of sentences, or identifying themes, i.e. identifying parts of propositions?

This would make a difference if the subject of the sentence weren't always the actor. In fact, that can happen.

But let's ignore that problem for the moment.

#### BACK TO CASE MARKING

There are at least 4 ways to mark case in the languages of the world:

1) Word order.

E.g. English:

John kissed Mary.

vs. Mary kissed John.

Likewise the Romance languages (languages historically derived from Latin): French, Spanish, Italian, etc.

2) Explicit case marking (noun declensions):

E.g. Latin (note the difference between Latin and its daughter languages):

Puer basivit puellam.

boy-S kiss-PAST girl-O.

"The boy kissed the girl."

Also: Puer puellam basivit.

Puellam puer basivit.

Puellam basivit puer.

etc.

### 3) Explicit case marking (constants):

E.g. Japanese:

John-ga Mary-o butta.  
John-S Mary-O hit.

### 4) Agreement markers on the verb.

E.g. Mohawk:

John Mary shako-núhwe's.  
John Mary he/her-likes.  
"John likes Mary."

John Mary ruwa-núhwe's.  
John Mary she/him-likes.  
"Mary likes John".

shako- is used only when the subject is a masc. sing. NP and the object is a fem. sing. NP.

ruwa- is used when the subject is fem. sing. and the object is masc. sing.

Mohawk has 58 such prefixes.

## HEAD MARKING VS. DEPENDENT MARKING

Latin and Japanese (2) and 3) above) are examples of *dependent-marking* languages.

Mohawk is a *head-marking* language.

The verb is the head of the VP: it defines the main action. Choosing the verb determines how many NP's there will be and what their cases will be (subject, direct object, indirect object [recipient], etc.) These NP's are dependents of the head.

Now let's look at what happens inside an NP:

Japanese:

John-no imooto  
John-'s sister  
"John's sister"

Latin:

soror pueri  
sister of-boy  
"the boy's sister"

In both of these languages, the marker is on the possessed noun.

Mohawk:

John rao-wise'  
John his-glass

“John’s glass”

The noun is the head of the NP: it defines what is being talked about. So the possessed noun is the head of the NP. The possessor is a dependent that modifies the head.

In Mohawk, the possession marker is on the head noun.

So for both NPs and VPs, Mohawk marks the head noun, while Japanese and Latin mark the dependent nouns.

Languages tend to have head marking or dependent marking, but not a combination of both.

Again, these means that not all possible sets of context-free rules are possible in a human language.

#### PRO-DROP AND OTHER PROPERTIES

Languages with a similar historical background (e.g. both are Romance languages) and many of the same word-order properties (e.g. both are SV0) can still have other differences.

(Even within the Romance languages, French and Italian are more closely related than either is to other languages.)

1) Possible sentence word orders:

Italian:

John verrà.  
John arrive-FUT  
“John will arrive”

French:

John arrivera.  
John arrive-FUT  
“John will arrive”

But not when using VS0:

Italian:

Verrà John.  
arrive-FUT John  
"John will arrive"

French:

\* arrivera John.  
arrive-FUT John

English patterns like French in this regard.

2) Things are not the same when using a pronoun as the subject of the sentence (i.e. we already know who's being referred to):

Italian:

Verrà.  
arrive-FUT  
"He/she/it will arrive"

French:

\* arrivera.  
arrive-FUT

Again, English patterns like French in this regard.

Italian is called a *pro-drop* language because it can drop subject pronouns.

French:	no VSO	no pro-drop
Italian:	VSO	pro-drop
Spanish:	VSO	pro-drop
Rumanian:	VSO	pro-drop

Historically, French and Italian are more closely related to each other than to the other Romance languages, and Rumanian is the farthest from all the others (for geographical and political reasons). So history doesn't explain this.

Again, we don't get a random choice of context-free rules.