LING 1010

Language and Mind

Prof. Jon Sprouse

03.10.21: The acquisition of syntax

Nativism vs Empiricism

Just to recap, here are the properties of Nativism and Empiricism, including their stances on the type of innate knowledge that is possible.

Modern Nativism	**Modern Empiricism**
 Substantial innate knowledge	 Minimal innate knowledge
 Input/experience still plays a role, but less than the role it plays in empiricism.	 Input/experience plays the largest role in learning
 The innate knowledge can be domain-specific.	 If there is innate knowledge, it is domain-general

Last time: Morphological Rule Learning

Last time we looked at the knowledge that children must have in order to solve word learning problems, and asked which theory it fits with:

Modern Nativism

Modern Empiricism

Memorizing the (past tense) form of verbs

Generalizing to a rule

Let's look at syntax and see if anything suggests domain-specific knowledge

The question is whether there are any aspects of rule learning that require domain-specific knowledge.

Transformations

Transformations

While it is possible to create phrase structure rules that capture all of the sentences in a given language, linguists have noticed that sometimes sentences appear to be related to other sentences:

Declarative: John is running.

Interrogative (Yes-No Question): Is John running?

To capture this relationship, linguists have postulated a second type of syntactic rule in addition to phrase structure rules. These additional rules are called **transformations**.

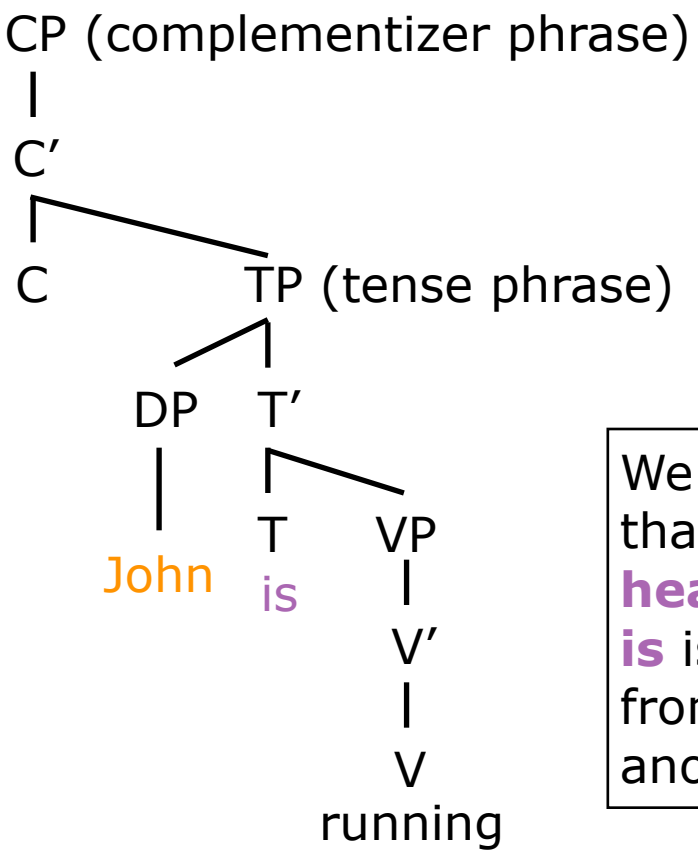
A transformation is exactly what it sounds like: it is a syntactic rule that takes the output of the phrase structure rules, and **rearranges** (or **transforms**) that output into a new output.

There are two steps: Phrase Structure rules followed by transformations

Here is the two-step process in tree form: (1) we apply the normal PS rules, then (2) we apply a transformation that moves is to a new location in the tree:

Step 1: Apply PS rules

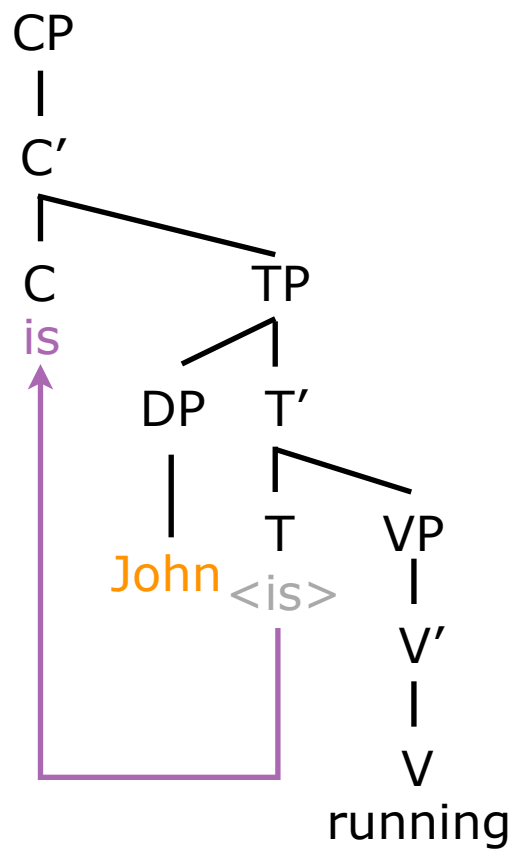
Output: John is running.



We call the transformation that does this **head movement** because **is** is a **head**, and it moves from one position to another.

Step 2: Apply transformation

Output: Is John running?



Head movement seems to occur in all yes-no questions

The head-movement transformation appears to be part of the process for forming all yes-no questions in English:

Declarative	**Question**
John is running.	Is John running?
John can juggle.	Can John juggle?
John will cook.	Will John cook?
John has left.	Has John left?

So it is going to be a pretty important grammatical rule for children to learn!

Is there domain-specific, innate knowledge in learning transformations?

Looking at head-movement more deeply

So far we've defined head-movement as "move is". But what happens when there is more than one "is" in the sentence?

John is thinking that Mary is smart.

Which "is" are we allowed to move in English (to form a question)?

In this sentence you can move the first one:

Is John <is> thinking that Mary is smart?

But you can't move the second one:

*Is John is thinking that Mary <is> smart?

This suggests that the definition of the head-movement transformation is more complicated than just "move is". It has to be something like "move the first is".

Testing a new definition (a new theory)

So let's see if we can test the theory that the definition of head-movement is "move the first is".

Theory to be tested: In English, you move the first is.

Here is a new sentence to test it on:

The woman that is happy is smart.

You CAN'T move the first one: *Is the woman that <is> happy is smart.

But you CAN move the second: Is the woman that is happy <is> smart.

Uh-oh. This sentence disproves our theory (in science, we say it falsifies the theory). This sentence works exactly opposite to our theory: our theory says you move the first one, but in this sentence you move the second one.

So here are the facts

For some sentences, you move the first instance of is.

John is thinking that Mary is smart.

Is John is thinking that Mary is smart?

For other sentences, you move the second instance of is.

The woman that is happy is smart.

Is the woman that is happy is smart.

So we need a theory that **does not rely on linear order**. Because the linear order doesn't seem to be the deciding factor.

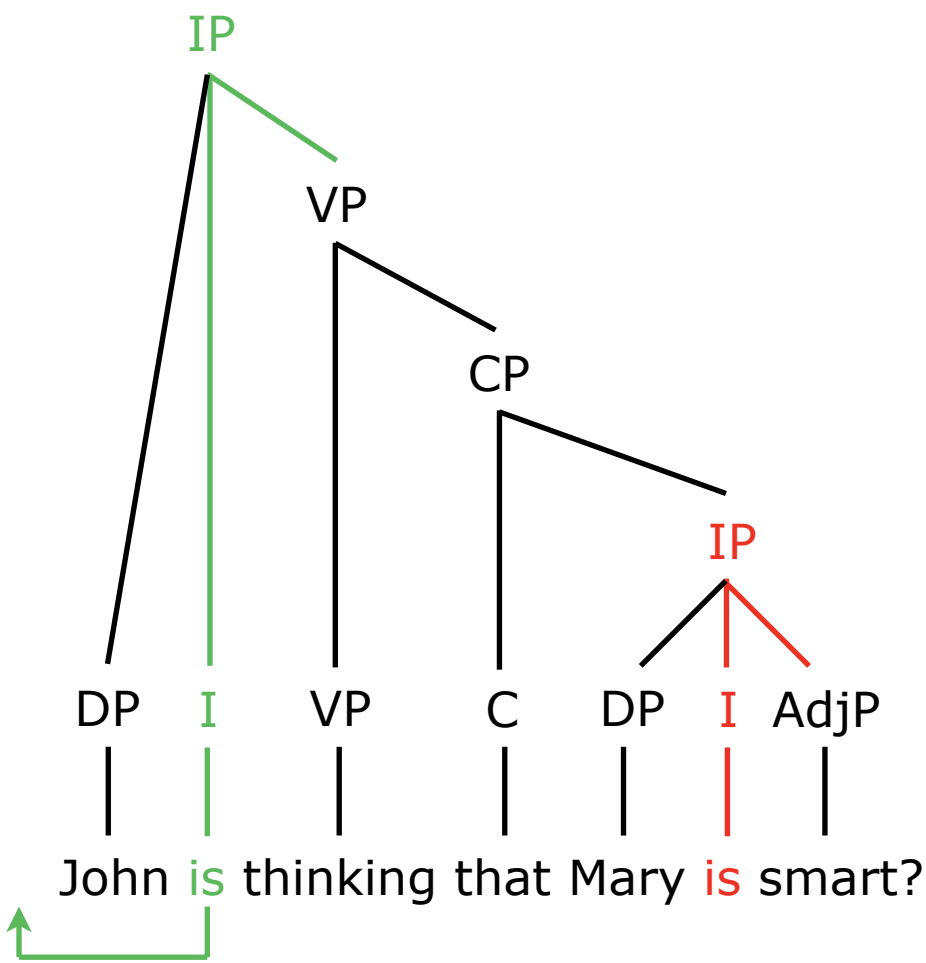
Creating a better theory: sentence 1

To create a better theory, we need to distinguish the two instances of is in the sentence using something other than linear order.

Let's use **structure** to distinguish them. In the first sentence, the two instances of is show up in different Inflection Phrases (IPs).

The green IP is called the matrix IP because it is the primary IP of the sentence. It forms the structure of the sentence. All other phrases fit into it, like a matrix. It is the **highest IP**. In grammar school, you would call this the main clause.

The red IP is called the embedded IP because it is embedded in the matrix of the sentence. It will be **lower than the matrix IP**. In grammar school, you would call this the embedded clause.

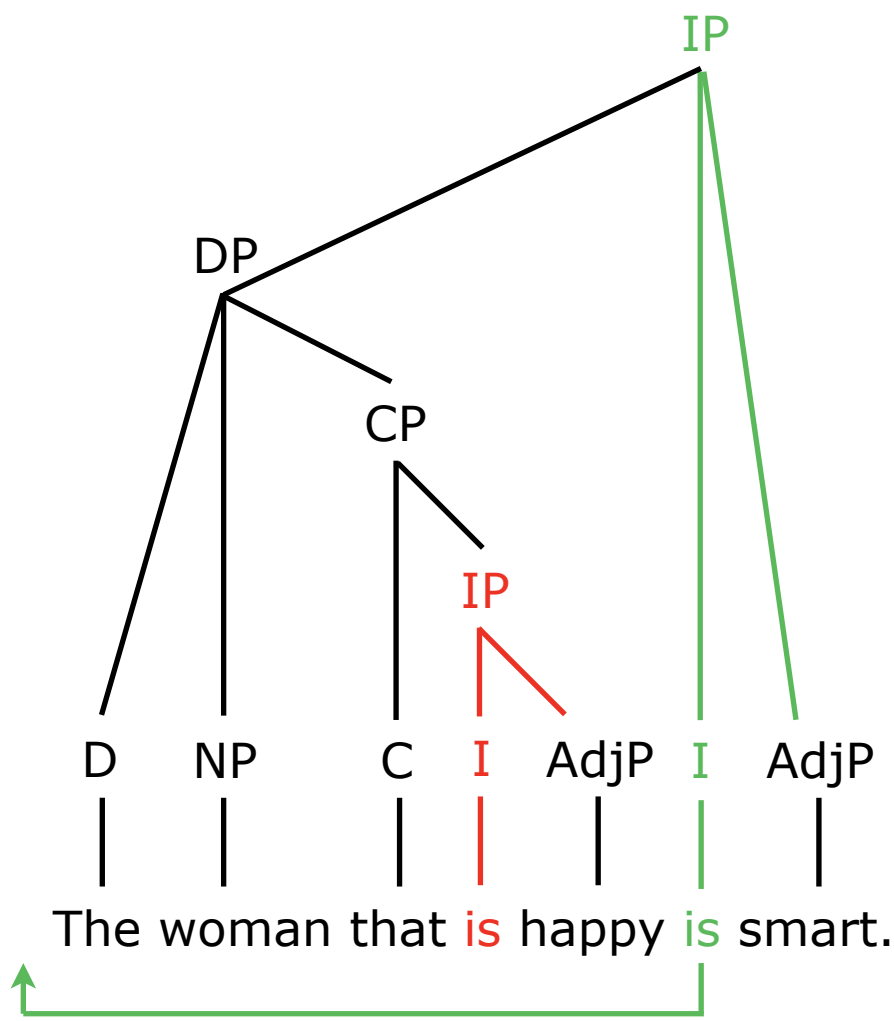


Creating a better theory: sentence 2

And if we draw the structure for the second sentence, a pattern emerges: the is that can be moved is always from the matrix IP!

The phrase "The woman that is happy" is called a **relative clause**. It is one long clause that forms the subject of the sentence. So the IP that is inside of it is an embedded IP.

The first is is inside an embedded IP. We can't move that one. The second is that is inside of the matrix IP. And that is the one that we move!



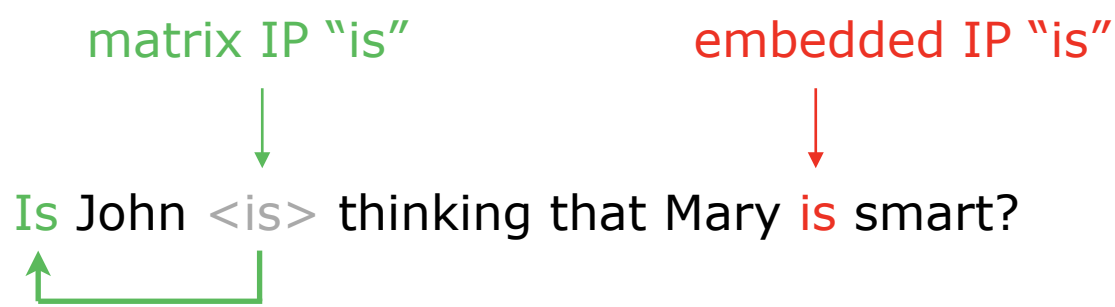
The correct theory

So now we see that the correct theory is something like "move the is that is in the matrix clause"

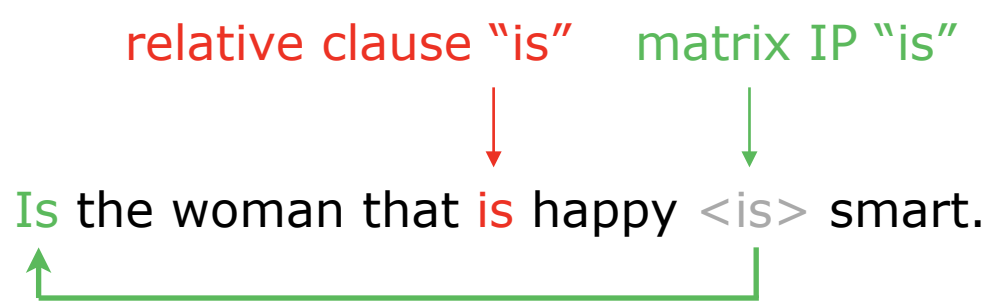
Correct theory: Move the is that is in the matrix IP

When we apply this to our two test sentences, we can see that it works:

Here you move the matrix clause is, but not the embedded clause is:



Here you move the matrix clause is, but not the relative clause is:



Structure Dependence

So in the end it looks like the "move first" theory doesn't work, but the "move the matrix IP is" does work:

Move first theory: Move the first is.

Move matrix IP theory: Move the is that is in the matrix IP

One interesting difference between these two theories is that the first one only makes reference to the **linear order** of the words in the sentence. It doesn't make reference to the hierarchical structure of the sentence at all.

But the second one makes reference to the **hierarchical structure** of the sentence. It makes a distinction between the is that is in the matrix IP, and any is's that are in other clauses.

For this reason, linguists call the second (correct) definition of head-movement a **structure dependent** rule. The transformation is defined in terms of the hierarchical structure of the sentence (e.g., matrix vs embedded IP). So it is a structure dependent rule.

Learning Structure Dependence

OK, so what is the big deal? Well, children need to learn the **correct definition of head-movement** in order to be able to create English questions.

We have seen two theories that they could try. How do they decide which one to try?

Move the first is:

Pros:	Simpler (not structure dependent)
Cons:	Ultimately incorrect

Move matrix IP is:

Pros:	Ultimately correct
Cons:	More complicated (structure dependent)

A possible (but incorrect!) learning theory

Here is one possible theory of how head-movement could be learned:

Step 1: Children notice that questions in English are formed by moving is.

Is John <is> happy?

Step 2: Children postulate the hypothesis that "move first" is the correct theory. They choose this one first because it is simpler, and because it works for a lot of questions in English:

Is John <is> thinking that Mary is smart?

Step 3: At some point, children notice a sentence that is incompatible with "move first". So they switch to the hypothesis that "move matrix" is the correct theory:

Is the woman that is happy <is> smart.

Why do we think that this theory incorrect?

Step 1: Children notice that questions in English are formed by moving is.

Step 2: Children postulate the hypothesis that "move first" is the correct theory. They choose this one first because it is simpler, and because it fits with a lot of questions in Englis:

Step 3: At some point, children notice an example that is incompatible with "move first". So they switch to the hypothesis that "move matrix" is the correct theory.

The problem with this theory is that successful learning requires hearing sentences like the following in order to notice that there is a problem with the move first theory:

Is the woman that is happy <is> smart.

Legate and Yang (2002) looked at over 20,000 questions that were spoken to a child in the CHILDES database… and they found precisely zero questions of the critical type! So if children relied on hearing this sentence to learn questions, **they would never learn how to form this question correctly!**

So how do children learn this?

Crain and Nakayama (1987) performed experiments to try to get children to produce yes-no questions in an attempt to see if they ever entertained the "move first" theory.

The prediction is that if children do entertain the hypothesis that "move first" is correct early in acquisition, then at some point early in acquisition they should produce sentences that follow the "move first" theory. These sentences will look like errors to us:

This is ungrammatical in adult English, but it is predicted to be spoken by children if they believe that "move first" is correct early in acquistion.

"Is the girl who <is> skating is tall?"

To test this prediction, Crain and Nakayama recruited a group of children ages 3;2 - 4;7, and played with puppets to try to get them to create yes-no questions. Then they looked to see if any of the yes-no questions showed the "move first" pattern. They tried really hard: they elicited 81 yes-no questions from the children.

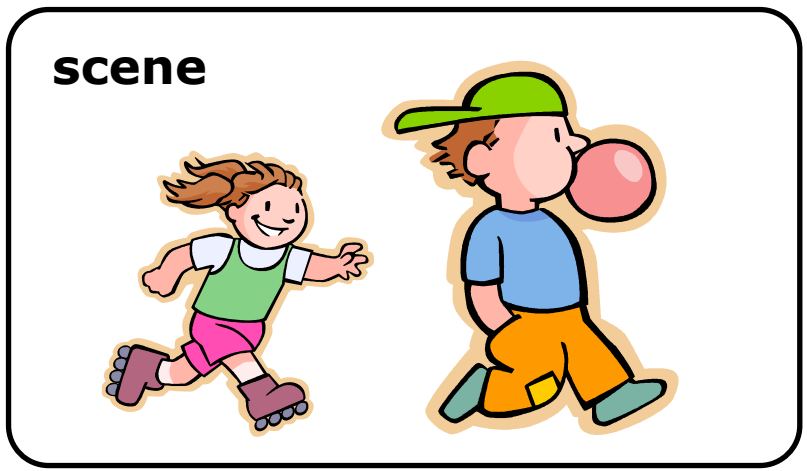
Here is an example

Experimenter: "Hey [child's name], look at that girl who is skating. Do you think she is tall?"

Child: "No! She isn't tall!"

Experimenter: "I wonder if Jabba thinks she is tall. Ask Jabba if he thinks the girl who is skating is tall."

Child: …… [creates the question] ……



Jabba Puppet

This was 1987!

The results

It should go without saying that children this young do make mistakes. In fact, they make more mistakes than correct responses. But the critical question is what type of mistakes do they make?

Do they make mistakes that suggest the "move first" theory? Or do they make other types of mistakes?

	Child Response	**Results**
adult-like	"Is the girl who is skating tall?"	38%
move first error	"Is the girl who skating is tall?"	0%
mistakes	"Is the girl who is skating is tall?"	37%
	"Is the girl who is skating, is she tall?"	12%
	other types	13%

What does this mean for learning?

These facts (the corpus facts and the experimental facts) seem to suggest that our learning theory is wrong:

Step 1: Children notice that questions in English are formed by moving is.

Step 2: Children postulate the hypothesis that "move first" is the correct theory.

No. Children do not seem to ever entertain the "move first" hypothesis. (Crain and Nakayama 1987)

Step 3: Children switch to the hypothesis that "move main clause" is the correct theory when they notice an example that is incompatible with "move first"

No. Children do not seem to ever hear sentences that would show that "move first" is wrong. (Yang and Legate 2002)

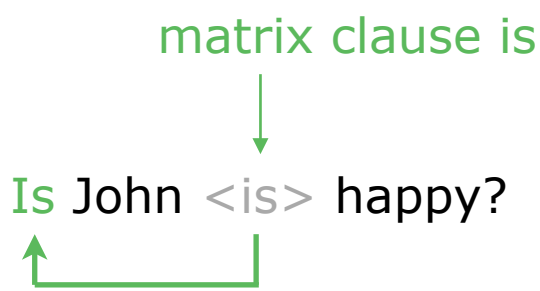
So how could children possibly learn the correct definition ("move matrix IP") given all of this?

A nativist theory might work

One possibility that might work is to postulate that children know (innately) that all transformations must be structure dependent.

Step 1: Children notice that questions in English are formed by moving is.

Step 2: Because children know innately that all transformations must be structure dependent, even a simple sentence is evidence that head-movement targets the matrix clause is:



Step 3: Therefore, as soon as children notice that a transformation is necessary, they will know the correct definition (the structure dependent definition).

Transformation Learning

So now we can look at the knowledge that children must have in order to solve the challenge of learning rules, and ask which theory it fits with:

Modern Nativism

This seems like it could be both domain-specific and innate. I am going to skew it toward Nativism because it is a good candidate. But I am open to future research showing it can be handled with domain-general knowledge

Only attempt **structure-dependent** rules.

Modern Empiricism

Hypothesizing a rule

Testing the rule

Some Conclusions

A transformation is a syntactic rule that takes the output of the phrase structure rules, and **rearranges** (or **transforms**) that output into a new output.

One ubiquitous example of a transformation in English is head movement which moves a phrasal head, like is, from one position to another in the tree.

Transformations in human syntax, such as head-movement, are **structure dependent**. This means that the transformation is defined in terms of the hierarchical structure of the sentence (e.g., matrix vs embedded IP).

Structure-dependent transformations pose a learning problem for children, because the input that children receive is often compatible with non-structure dependent transformations (like "move first"). So the question is how children could possible learn that the correct transformation is structure dependent.

One possibility that might work is to postulate that children know (innately) that all transformations must be structure dependent. There is some evidence that children never entertain non-structure-dependent hypotheses.