جامعة نيويورك أبوظبي

NYU | ABU DHABI

PSYCH-UH 2218: Language Science

Class 14: Syntax - an introduction

Prof. Jon Sprouse
Psychology

Fact 1: There are possible and impossible strings of words

Possible and Impossible strings of words

You already know this, but let's make it explicit. There are strings of words in your language that are possible, and strings that are impossible.

✓ Revolutionary new ideas occur infrequently.

* Infrequently occur ideas new revolutionary.

Reversing the order order typically results in ungrammaticality.

Create a short sentence in your native language, and write it down.

The sentence I made above has 5 words. That means there are 120 possible orders of the words (=5!).

Now write down as many of the orders of the words as you can, and ask yourself which ones are grammatical?

You will find that very few will be grammatical. Maybe only 1!

✓ Revolutionary new ideas occur infrequently.

* Infrequently occur ideas new revolutionary.

* Revolutionary ideas infrequently new occur.

* Occur new revolutionary infrequently ideas.

Two theories to explain this

Theory 1: You have memorized which sentences are possible in your language. We have great memories, so this is possible in principle.

Theory 2: You have learned a system of rules that sentences in your language must follow.

Both of these theories can potentially explain this fact. We need more evidence!

✓ Revolutionary new ideas occur infrequently.

* Infrequently occur ideas new revolutionary.

* Revolutionary ideas infrequently new occur.

* Occur new revolutionary infrequently ideas.

How can we tease apart these two theories?

Fact 2: The number of sentences in your language is probably infinite

How many sentences are there in English?

Imagine that somebody asked you this question, like in a trivia game or something.

This feels unanswerable, right? It feels like a trick question.

That is because it is very easy to demonstrate that we can probably create a very large number of sentences, probably infinite, like so:

I have one car.
I have two cars.
I have three cars.
I have four cars.

This also feels like a trick, but notice what it is doing — it is showing us why we didn't want to answer that first question!

This is not, strictly speaking, a proof that sentences are infinite in number. But it makes many linguists believe that they are.

Quick activity: Can you do this in your language?

What is the longest sentence in English?

Imagine that somebody asked you this question, like in a trivia game or something.

This feels unanswerable, right? It feels like a trick question.

This feels feels like a trick answer again, but it shows that there is no apparent upper bound to the length of sentences in English.

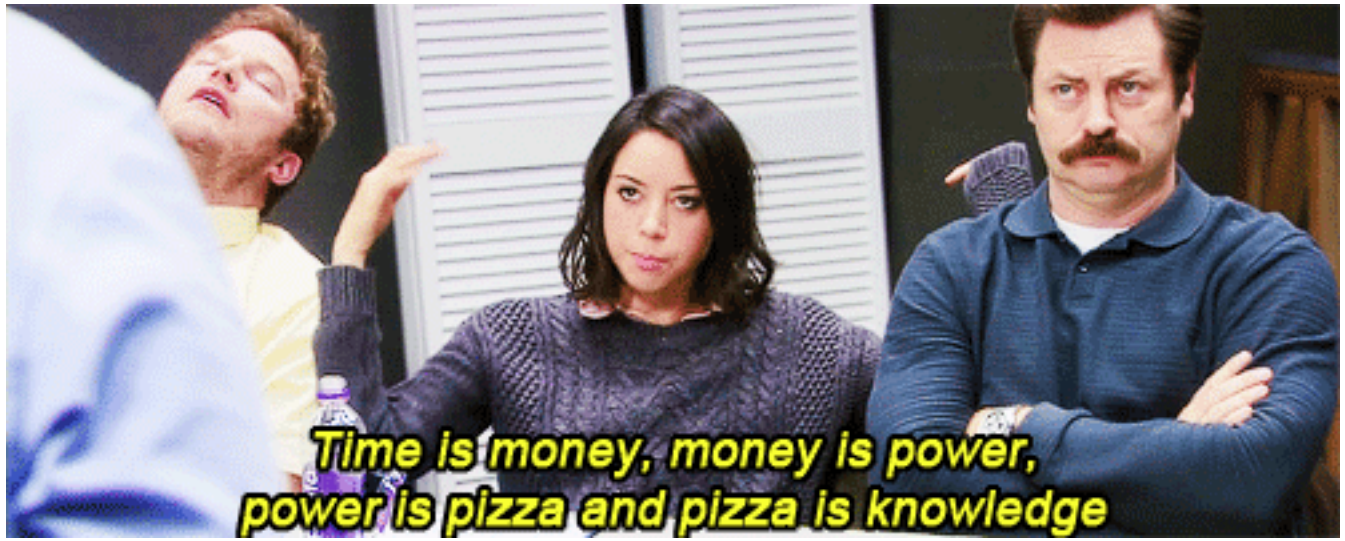
I like cookies.
Lisa said that I like cookies.
Clare thinks that Lisa said that I like cookies.
Ben claimed that Clare thinks that Lisa said that…

Again, this is not, strictly speaking, a proof that sentences are infinite in number. But it makes many linguists believe that they are.

Quick activity: Can you do this in your language?

A related idea: novel sentences

Here is a sentence that you have probably never heard before (or before you heard it on Parks and Rec). In other words, it is a novel sentence:



Almost all jokes that you find funny are novel sentences — because jokes tend to be less funny the second time you hear them.

Novel sentences are not quite about infinity. But they do show that the number of sentences in a language is very large — so large, that there are sentences you have not yet heard before. In fact, linguists argue that most of the sentences you hear are novel — including most of those on this slide!

Quick activity: Can you make a novel sentence in your language?

Two theories to explain word order restrictions

Theory 1: You have memorized which sentences are possible in your language. We have great memories, so this is possible in principle.

> **No:** If the number of sentences in a language are infinite, it is not possible to have memorized them all individually, because we have finite memory.
>
> **No:** If the length of sentences in a language are potentially infinite, it is not possible to have memorized those sentences, because we have finite memory.
>
> **No:** If we can understand novel sentences, then understanding sentences is not about memorization, because you could not have memorized something that you did not hear before!

Theory 2: You have learned a system of rules that sentences in your language must follow. That system leads to an infinite number of sentences, leads to infinitely long sentences, and allows you to understand sentences you have never heard before!

The shape of the theory

Before we jump into this, I just want to reassure you that the theory of syntax will ultimately have the same general shape as the other theories that we have seen:

	phonology	**morphology**	**syntax**
explananda:	sequences of sounds	sequences of morphemes	sequences of words
primitives:	phonemes	morphemes	syntactic categories
rules:	phonological rules, syllable structure	morphophonological, structure building	phrase structure, transformations
constraints:	phonotactic constraints	headedness deriv./inflect. ordering	headedness X'-theory …

What types of phenomena will help us uncover the mental representations of sentences?

We will look for unexpected ungrammaticality

Large changes in word order, like reversing the order of the sentence, are fairly straightforward to notice, and fairly straightforward to capture with rules.

The more subtle effects arise when you would expect certain word order patterns to be possible given logical reasoning. This is the syntactic version of gaps in the paradigm like we saw in phonology and morphology!

✓ Mary seems to be happy.

✓ Mary tried to be happy.

✓ It seems that Mary is happy.

* It tried that Mary is happy.

✓ It is likely that Mary will leave.

✓ It is probable that Mary will leave.

✓ Mary is likely to leave.

* Mary is probable to leave.

✓ I believe that Mary is honest.

✓ I think that Mary is honest.

✓ I believe Mary to be honest.

* I think Mary to be honest.

We will look for words acting as units together

Take a look at these two sentences. They differ by one word - the 7th word in the sentence. The other 6 words are identical. But I am sure you notice that there is an interesting psychological shift in these two sentences!

1	2	3	4	5	6	7
Jack	and	Jill	ran	up	the	hill
Jack	and	Jill	ran	up	the	bill

What seems to be happening is that up the hill acts as a unit in the top sentence, while it is only the bill (no up) that acts as a unit in the second sentence. When words act as a unit we say they form a **constituent**.

✓ Up the hill, ran Jack and Jill.

* Up the bill, ran Jack and Jill.

Up the hill can be moved to the front of the sentence, but up the bill cannot.

* Jack and Jill ran the hill up.

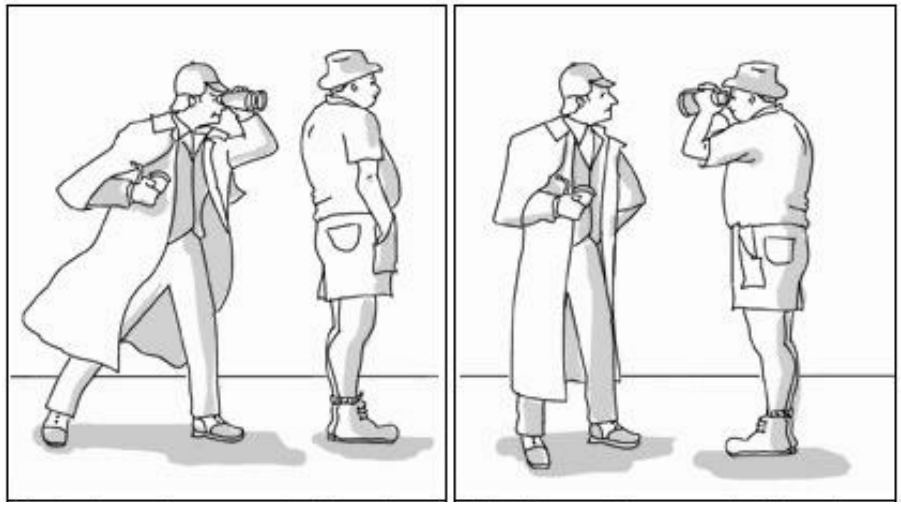
✓ Jack and Jill ran the bill up.

The bill can be moved between ran and up. But the hill cannot.

Sentences can also be ambiguous

Sentences can be ambiguous just like we saw with words. For example, this string of words can have two different meanings. Assuming that the meaning of sentences is compositional (based on the words in the sentence), how can we explain the same set of words leading to two distinct meanings?

We will use the same logic that we used for morphology, and conclude that sentences must have **hierarchical structure**.



Sherlock saw the man with binoculars.

Quick activity: Can you make an ambiguous sentence in your own language?

Cross-linguistic variation

Finally, and perhaps most importantly, this should be a theory of language as a cognitive ability for all humans. Therefore the rules we make should explain all languages, not just one specific language.

Here is a sentence in Japanese, with the word-by-word translation in the second line (called a gloss in linguistics) and an equivalent English sentence below it in quotes to show you what the sentence means (called the translation in linguistics):

	Taro-ga	Hiro-ga	Hanako-ni	neko-no	syasino	miseta	to	omotte	iru
gl:	Taro	Hiro	Hanako-to	cats-of	pictures	showed	that	thinking	is

tr: 'Taro is thinking that Hiro showed pictures of cats to Hanako.'

Most English speakers feel as though Japanese word order is very different from English. In fact, when given glosses, many English speakers have no idea what the sentence means. It comes across as gibberish. But our theory of syntax needs to explain both Japanese and English (and all other languages).

A quick excursion — it is not just about a lack of meaning

There is clearly a relationship between syntax and semantics

This is something you already know. The word order in a sentence conveys important information about the meaning

The dog bit the man. vs. The man bit the dog.

Given this, one tempting idea is that the impossible strings that we have seen are not ungrammatical because of a problem with the syntax, but rather because of a problem with the semantics.

In other words, when we see the asterisk indicating ungrammaticality, it is tempting to say "that is just because the sentence doesn't make sense".

✓ Revolutionary new ideas occur infrequently.

* Infrequently occur ideas new revolutionary.

* Revolutionary ideas infrequently new occur.

* Occur new revolutionary infrequently ideas.

Syntacticians would say that this gets the situation backwards. These sentences "don't make sense" because their syntax is incorrect.

Syntactic rules can be respected or violated even when there is no meaning

Argument 1: syntactically grammatical sentences that have no meaning

✓ Colorless green ideas sleep furiously.

* Furiously sleep ideas green colorless.

* Sleep green furiously colorless ideas.

* Ideas colorless green furiously sleep.

The meanings of these words simply cannot go together. Something can't be both green and colorless. Ideas don't have color. Ideas can't sleep. And you can't sleep furiously.

Speakers recognize the first sentence as syntactically grammatical despite having no meaning.but its reverse is not. You can even see this in how fluently you can speak these strings - the first is fluent, the second becomes a list of words.

Crucially, other orders are not. You can even see this in how fluently you can speak these strings - the first is fluent, the second becomes a list of words.

Syntactic rules can be respected or violated even when there is no meaning

Argument 1: syntactically grammatical sentences that have no meaning

Here is another example that is slightly different:

More people have been to France than I have.

Ask yourself: What does this sentence mean?

It does not have a meaning. But on first hearing it, English speakers tend to find the sentence syntactically grammatical. (There are several experimental studies on this!).

Syntactic rules can be violated while maintaining meaning!

Argument 2: semantically interpretable sentences that are syntactically ungrammatical

* Dog the bit man the.
* Three dog bit the man.
* The boy quickly in the house the ball found.

Each of these sentences is syntactically ungrammatical (for different reasons). But English speakers typically have no trouble determining the meaning of the sentences.

More broadly, we know that English speakers are able to understand other people who do not speak English exactly the same as them — such as children learning the language, adults learning the language, or even speakers of slightly different varieties of English (UK vs US, etc).

So what is the relationship between syntax and semantics?

Ultimately, we will build a theory in which syntax feeds semantics. What this means in practice is that when we see dramatic changes in semantics, there is probably a difference in the syntax:

Mary appeared to John to be polite.

Mary appealed to John to be polite.

Who is "being polite" in each of these sentences?

Mary is eager to please.

Mary is easy to please.

Who is doing the pleasing and who is being pleased in each of these sentences?

The cat is out of the bag.

The cat seems to be out of the bag.

The cat tried to be out of the bag.

What is "the cat" in each sentence? An actual cat, or a secret?

This suggests that the syntactic representations of these sentences are different from each either even though the linear order is the same. This is similar to ambiguous sentences — it suggests hierarchical structure!

The first component of a our theory of syntax: syntactic category

We don’t want a different rule for each word

Estimates for the number of words in English range from 100,000 to 1,000,000. It turns out it is really complicated to count words (remember the Inuit). But it doesn’t really matter what the exact number is. It is very large!

Syntactic rules tell us where to put each word in the sentence. If we had a rule for each word, that would be a lot of rules for children to learn, and a lot of rules to store in our minds. We’d need one for each word, and perhaps several to account for the other words it can appear with.

Fortunately, it looks like the human mind takes advantage of the idea of **categories** to reduce the number of rules. The idea is that two objects in the same category share some relevant property. For example, two movies in the “horror” category will share the property of being scary.

The same idea seems to apply to words. Two words that share the same **syntactic category** can **appear in the same position in a sentence.**

A test for syntactic categories

Two words that share the same **syntactic category** can **appear in the same position in a sentence.**

This definition suggests its own cognitive test. We can determine whether two words are in the same syntactic category by asking whether they can appear in the same position within a specific sentence.

So all we need to do is choose various sentences, delete a word from it, and ask which words can go in the blank. These are sometimes called **frames**.

The ____ existed.

dog

homework

idea

All of the words that can fit in this position are the same **syntactic category,** which in this case we call **nouns**.

* eat

* of

* quickly

The words that can't fit in this position are not nouns. We need more tests to see which category each of these words are.

More frames to test syntactic several categories in English

Nouns:	The ____ existed.	dog, book, idea
Verbs:	The cat will ____.	sleep, meow, jump
Prepositions:	It hid right _____ here.	over, under, near
Adjectives:	They are very ____.	happy, sad, tired
Adverbs (manner):	She coughed ____.	loudly, quickly
Adverbs (sentential):	____, you are a liar.	honestly
Determiners:	He wrote ____ other work(s).	the, an, five
Complementizers:	I know ____ John is a liar.	that, if, whether

Mad Libs!

The children's game Mad Libs takes advantage of the fact that words of the same syntactic category can replace each other in a sentence.

In this game, there is a humorous story with words missing. The missing blanks are labeled with a category of word. You come up with words of each type without reading the story, then plug them in to see if the resulting story is funny.

The meaning will sometimes be strange (or funny), because syntactic category **is not about meaning**. Syntactic category is just about where in the sentence the word can go according to syntax. It is not about semantics!

MAD LIBS

TALK LIKE A PIRATE

Ye can always pretend to be a bloodthirsty _____ (NOUN), threatening everyone by waving yer _____ (ADJECTIVE) sword in the air, but until ye learn to _____ (VERB) like a pirate, ye'll never be _____ (ADVERB) accepted as an authentic _____ (NOUN). So here's what ye do: Cleverly work into yer daily conversations _____ (ADJECTIVE) pirate phrases such as "Ahoy there, _____ (PLURAL NOUN)," "Avast, ye _____ (PLURAL NOUN)," and "Shiver me _____ (PLURAL NOUN)." Remember to drop all yer *g*s when ye say such words as *sailin'*, *spittin'*, and *fightin'*. This will give ye a/an _____ (PART OF THE BODY) start to being recognized as a swashbucklin' _____ (NOUN). Once ye have the lingo down pat, it helps to wear a three-cornered _____ (NOUN) on yer head, stash a/an _____ (NOUN) in yer pants, and keep a/an _____ (NOUN) perched atop yer _____ (PART OF THE BODY). Aye, now ye be a real pirate!

Syntactic category is not the same as the "parts of speech" that you learned in school

Though we are using labels for syntactic categories that you have see before (noun, verb, etc), they are not the same. (Linguists just like to take names from traditional grammar that people are familiar with.)

Think about the definitions that your teachers gave you for parts of speech:

"A noun is a person, place, thing, or idea."
"A verb is an action or state of being."

These are definitions based on **meaning** (semantics). They are not based on syntax. **Syntactic category** is only about syntax - the places in the sentence the word can appear.

Also, remember that science abhors disjunctions - if you have to say "something is either x or y", you don't have a very good theory. Notice that we don't have that issue with syntactic category. You either are, or are not, each category. There is no "person, place, or thing".

Syntactic Categories must be part of lexical entries!

Up until now, items stored in the lexicon (morphemes) have consisted of a pair of representations: a sound (phonemic representation) and a meaning (semantic representation):

Phonemic representation: [k æ t]

Semantic (meaning) representation:

Syntactic Category: noun

morpheme

Now we need to add a third piece of information to the entry: **syntactic category.**

Activity: Find these in your language!

Nouns:	The ____ existed.	dog, book, idea
Verbs:	The cat will ____.	sleep, meow, jump
Prepositions:	It hid right _____ here.	over, under, near
Adjectives:	They are very ____.	happy, sad, tired
Adverbs (manner):	She coughed ____.	loudly, quickly
Adverbs (sentential):	____, you are a liar.	honestly
Determiners:	He wrote ____ other work(s).	the, an, five
Complementizers:	I know ____ John is a liar.	that, if, whether