LING 1010



Language and Mind Prof. Jon Sprouse

02.08.21:

Morphology

Morphology: shape and meaning

Words are (at least) a pairing of **sound** and **meaning**:

sound

meaning

[kæt]



Much like phonology is driven by puzzles (e.g., regularities in the distribution of sounds), **morphology** is driven by puzzles about the relationship between the meaning and the shape of words. By shape, we simply mean the sequence of phonemes that make up the word.

Morphology is the study of the shape of words.

The pairings are (mostly) arbitrary

For simple words (we will get better at defining this later), the pairing between sound and meaning is arbitrary. There is no reason why the meaning cat is paired with the sound cat in English. We can see this by looking at all of the different sounds that are paired with this meaning in different languages:

language

arabic ethiopian farsi gaelic hawaiian icelandic korean mayan swahili tamil thai tsalagi

word besseh domadh gorbeh piscin popoki köttur koyangi miz paka poonai maa-oh we'sa





But some pairings are systematic

Let's take a look at a classic puzzle in morphology. We can use this puzzle to motivate quite a bit of the theory of morphology:

What does this word mean?

Meaning 1: Can't be locked

unlockable



Meaning 2: Can be unlocked



Achievement Unlocked Obsessed With Gamerscore

Why are there two meanings for this word? Why isn't there one, or more than two? And why are they the meanings that they are?

But some pairings are systematic

Every word that has the form un-X-able, where X is a reversible verb, seems to have two meanings (no more, no less). And those meanings always seem to be the same two meanings:



This seems like something more than a coincidence. Morphology wants to find a way to explain this.

Insight 1: These words seem to be made of smaller parts

If you look at all of these words together, a pattern emerges. They all seem to be built from the same two parts (un, able) and a different verb in the middle.

This should spark an idea. If word meanings are built from their parts, then the similarities between these words can be explained: they are similar because they use the same parts (2/3!). un-lock-able

un-do-able

un-fold-able

un-learn-able

Compositional:	In linguistics, we say that the meaning of a string is
	compositional if the meaning can be derived from the
	independent meanings of the parts (if the meanings are
	composed of the meanings of the smaller parts).

Compositionality is another great example of structure in the mind. We interpret meanings from smaller pieces, but don't ever realize it!

Defining the parts: Morphemes

Some words are clearly one piece:

fierceThese words are complete units. There are no sub-parts thatdeskhave meaning outside of the word. So these are each clearly abootseparate lexeme. For example, the **erc** in **fierce** does not haveatan independent meaning.

However, if we start looking around we see that lots of words are made up of smaller pieces; and that those pieces seem to have regular meanings:

desks pencils boots	The s in these words seems to mean something like multiple : we can call it plural .
jumped failed labeled	Similarly, the ed in these words seems to mean something like in the past: we can call it past tense
preset	
prepay	The pre in these words seems to mean something like before
preboard	

Defining the parts: Morphemes

Morpheme: The smallest unit of language that carries a distinct meaning.

Some words are only a single morpheme:

fierce desk boot at Each of these words is a single morpheme - there is one unit in the word that carries meaning: the entire word itself

Some words contain two morphemes:

desks pencils boots	The ${f s}$ in these words is a morpheme.
	The rest of the word is a morpheme too!
preset prepay preboard	The pre in these words is a morpheme.
	The rest of the word is a morpheme too!

Bound vs Free morphemes

It is possible to investigate all of the types of morphemes in a language, and develop a theory of the types of morphemes. That theory is called a theory of **morphology** (the shape of words).

-	
fierce desk boot at	Morphemes that can be a stand-alone word are called free morphemes
desks pencils boots	Morphemes that only occur attached to a free morpheme are called bound morphemes
jump <mark>ed</mark> fail <mark>ed</mark> label <mark>ed</mark>	
preset prepay preboard	

Affixation: combining bound and free morphemes

Bound morphemes are sometimes called **affixes** because they must be affixed to a free morpheme.

In general, there are three types of affixes:

prefix: a bound morpheme that appears before the free morpheme

preset repay unqualified **suffix:** a bound morpheme that appears after the free morpheme

jump<mark>ed</mark> fail<mark>ing</mark> labels

infix: a bound morpheme that appears inside of a free morpheme

English doesn't really have infixes, but we have two slang examples that are close:

hizouse shiznit	fanfuckingtastic
Why doesn't iz	n count as an infix?
Why doesn't f u	ucking count as an infix?

More structure: the rules of "infixation" in English

Here is another great example of structure in the mind: there are rules to the way that the word "fucking" can be inserted into words in English.

Every native speaker of English knows this rule, but I bet you were never taught it.

Let's take a look at some examples:

fantastic	\rightarrow	fan <mark>fucking</mark> tastic
absolutely	\rightarrow	abso <mark>fucking</mark> lutely

Notice that you can't put it in a different spot:

*fantasfuckingtic

*absolutefuckingly

The asterisk means imposible.



http://xkcd.com/1290/

More structure: the rules of "infixation" in English

The rule is based on **word stress**. Word stress is the extra acoustic prominence that we give to certain syllables inside of words.

The "fucking"The word "fucking" can only be inserted in the positioninsertion rule:immediately before the primary stressed syllable.

You can find the primary stress in a word by putting your hand horizontally under your chin, and saying the word. The syllable that makes your hand go the furthest down is the primary stress.

We can mark stress with an acute accent diacritic. And you can see that "fucking" is always inserted before the primary stress:



Compounding: combining two free morphemes:

Compound words are words that are composed of two (or more) free morphemes.

Some languages, including English, form noun-noun compounds by combining two (or more) nouns

Novel compounds are compounds that you make up on the fly. They tend to have a fully compositional meaning. They also tend to have two (or more) possible meanings:

cookie chair A chair made of cookies / shaped like a cookie. A chair for (eating?) cookies.

Lexicalized compounds are compounds that have become stored as complete units. They tend to be written as a single unit. They tend to have one meaning, which may not be completely compositional any longer:

teacup

cupcake

flagship

Insight 2: These words are ambiguous

Ambiguity:

In linguistics, we say that the meaning of a string is ambiguous if there is more than one possible meaning.



The puzzle of ambiguity

Ambiguity raises a real puzzle for compositionality: How is it that two meanings can come from the same pieces?



If meaning comes from the parts (compositionality), then when we have the same parts, we should get the same meaning, right???

Ambiguity through hierarchical structure

Instead of throwing out compositionally, we save it by saying that complex words are compositional, and that the word has **hierarchical structure**. It is a difference in the structure that leads to a difference in the meaning!

Meaning 1: lock+able, then un + lockable



Meaning 2: un+lock, then unlock + able



Hierarchical structure:

smaller units are combined to form larger units.

We can use **trees** to demonstrate the hierarchical structure.

Two items that combine are linked with two lines that converge into a node. We label that node in order to show that they formed a new unit with certain properties.

Ambiguity through hierarchical structure

Instead of throwing out compositionally, we save it by saying that complex words are compositional, and that the word has **hierarchical structure**. It is a difference in the structure that leads to a difference in the meaning!

Meaning 1: lock+able, then un + lockable





Meaning 2: un+lock, then unlock + able

unlockable unlock able un lock



Achievement Unlocked Obsessed With Gamerscore

WARNING!

We are now using the word "**structure**" in two different contexts. It is important not to be confused by it!

The first use is when we talk about the **mind having structure**. What we mean there is that the mind is not some undifferentiated blob that can work any which way that we need. It works a certain way. It is structured just like any other organ in the body (the heart functions a certain way, the liver functions a certain way, etc).

The second use is when we talk about **words having structure** (and later we will see that sentences also have structure). What we mean here is that there is a specific shape (or structure) to words. This shape/structure might not always be apparent, but it is there. It explains certain properties of words (and sentences), like the way that certain words are ambiguous.

Both of these senses are the same in a way: structure means having a specific form, just like buildings have structure. But it is important to know exactly what we mean when we say that "the mind has structure" and "words have structure", because the details are a bit different.

Structure-building rules

We can capture the hierarchical structures of words with structure-building rules. These rules combine two objects together to yield a third (larger) object:

Meaning 1: lock+able, then un + lockable



Meaning 2: un+lock, then unlock + able

unlockable ightharpoonup definition of the second state of the

Making the rules more general

These rules are specific to the word unlockable. But we saw earlier that this ambiguity is true of other words (undoable, unlearnable, etc). We can make the rules more general by replacing the non-affixes with parts of speech:

Meaning 1: lock+able, then un + lockable



Meaning 2: un+lock, then unlock + able

adjective \bigvee VERB + able \rightarrow ADJ \bigotimes Achievement Unlocked \bigvee Obsessed With Gamerscore \bigvee un + VERB \rightarrow VERB \bigcirc un | lock

Applying the rules to other words

To see that this general form works, let's try other words. First, let's try the word **undoable**.

Meaning 1: do+able, then un + doable



Meaning 2: un+do, then undo + able





Applying the rules to other words

To see that this general form works, let's try other words. Next, let's try the word **unlearnable**:

Meaning 1: learn+able, then un + learnable



Meaning 2: un+do, then undo + able



The standard form of the rule

The format of the rules on the previous slides is very easy to understand. However, there is a more standard format for the rules that reverses the order of the left/right sides, and removes the plus sign.

Meaning 1: can't be locked



Hierarchical Structure and Ambiguity

And here is the big payoff from structure-building rules. The two meanings come from two different hierarchical structures, which we get through the application of different rules. In this case, from three rules, we get two distinct meanings. Both meanings use the "able" rule. But they each use a different "un" rule, and use it in a different order.



Rethinking arbitrariness

The pairing of sound and meaning for individual morphemes is arbitrary. Our example "cat" is a single morpheme, so its pairing is arbitrary.

language	I
arabic	
ethiopian	
farsi	
gaelic	
hawaiian	

word besseh domadh gorbeh piscin popoki



But once that relationship is established, the presence of that morpheme in a **multi-morphemic** word will have **systematic effects due to compositionality**.

$$cat + s =$$

A real-world application of a theory of morphology

How many words do the "Eskimo" have for snow?

We should actually call them speakers of Inuit-Yupik/Aluet languages, but here I use the term "Eskimo" because that is what this saying/meme uses.



First, let's confront our biases

Why do we think that Inuit-Yupik speakers would have lots of words for snow?

Do we think that they are excellent snow researchers, and therefore need a precise vocabulary for the different types of snow formations?

No, we don't. So this is not a comment on their scientific interest in snow. Could it be something negative?



First, let's confront our biases

Why do we think that Inuit-Yupik speakers would have lots of words for snow?

Or could it be a subtle form of language-oriented prejudice?

Language prejudice is something that we will discuss in more detail later in the semester. But for now, I want you to be aware that it exists, and show you a little bit about how we can apply our scientific theories of language to claims about languages that may have a root in prejudice.



Now let's try to answer the question for English...

How many words for snow are there in General American English?



Now let's try to answer the question for English...

How many words for snow are there in General American English?



So how many "words" for snow are there in Inuit ("Eskimo") languages?



There are around 12 distinct morphemes for snow-like phenomena, including both canonical snow (snow, blizzard, flurries), and related things (slush, sleet, etc). **This is not much more than English!**

The Inuit languages have about 280 different grammatical forms of each word (e.g., snows, snowed, snowing) that can be formed through affixation. This is MUCH more than English, and is due to the grammatical properties of the languages.

So there are two answers. If we only count distinct morphemes then the number is very close to English (\sim 12). If we count the different forms that come from affixation, then then number is ridiculously large (>1000). But the large answer is simply a grammatical fact of the language, not an indicator of precision in discussing snow!

Some conclusions

Morphology is the study of the shape of words (the sequences of phonemes in a word). It is driven by puzzles surrounding the relationship between the shape and the meaning of words.

We say that the meaning of a string is **compositional** if the meaning can be derived from the independent meanings of the parts (if the meanings are composed of the meanings of the smaller parts).

The smallest unit of language that carries a distinct meaning is called a **morpheme**.

We say that the meaning of a string is **ambiguous** if there is more than one possible meaning.

Ambiguity leads us to believe that multi-morphemic words have hierarchical structure, which allows multiple meanings to be captured compositionally.

We can use **structure-building rules** to capture hierarchical structure.

The relationship between the sound of a morpheme and the meaning of morpheme is arbitrary. But once that relationship is established, the presence of that morpheme in a multi-morphemic word will influence the meaning of the word due to compositionality and structure-building rules!