Linguistics 001 Lecture 3 The structure of words

Words and morphemes

home

schedule

homework

Today we turn to **morphology**, which deals with how words are put together out of smaller pieces that linguists call morphemes, the minimal units of linguistic form and meaning.

So what does "minimal unit of form and meaning" mean? Consider some English words.

dog
dogs
bulldog

walk
walks
walked
walking
moonwalk

red
reddish
redden
reddens
redder

Most of these words can be divided up into identifiable parts, each of which has some kind of independent status, as evidenced by the fact that it occurs in other words (usually with a similar meaning or function).

dog + s

cf. hand+s, cat+s, book+s

walk + ing

cf. talk+ing, runn+ing, sing+ing

redd + en

cf. black+en, whit+en, short+en

Each of these independent elements is a **morpheme**. The definition includes "minimal" because reddens breaks down into not just redden + s, but into redd + en + s -- and no further.

We've started talking blithely about words and morphemes as if it were obvious that these categories exist and that we know them when we see them. This assumption comes naturally to literate speakers of English, because we've learned through reading and writing where white space goes, which defines word boundaries for us; and we soon see many cases where English words have internal parts with separate meanings or grammatical functions, which must be morphemes.

In some languages, the application of these terms is even clearer. In languages like Latin, for example, words can usually be "scrambled" into nearly any order in a phrase. As Allen and Greenough's New Latin Grammar says, "In connected discourse the word most prominent in the speaker's mind comes first, and so on in order of prominence."

Thus the simple two-word sentence *facis amice* "you act kindly" also occurs as *amice facis* with essentially the same meaning, but some difference in emphasis. However, the morphemes that make up each of these two words must occur in a fixed order and without anything inserted between them. The word *amice* combines the stem /amic-/ "loving, friendly, kind" and the adverbial ending /-e/; we can't change the order of these, or put another word in between them. Likewise the verb stem /fac-/ "do, make, act" and the inflectional ending /-is/ (second person singular present tense active) are fixed in their relationship in the word *facis*, and can't be reordered or separated.

In a language like English, where word order is much less free, we can still find evidence of a similar kind for the distinction between morphemes and words. For example, between two words we can usually insert some other words (without changing the basic meaning and relationship of the originals), while between two morphemes we usually can't.

Thus in the phrase "she has arrived", we treat *she* and *has* as separate words, while the /-ed/ ending of *arrived* is treated as part of a larger word. In accordance with this, we can introduce other material into the white space between the words: "she apparently has already arrived." But there is no way to put anything at all in between /arrive/ and /-ed/. And there are other forms of the sentence in which the word order is different -- "has she arrived?"; "arrived, has she?" -- but no form in which the morphemes in *arrived* are re-ordered.

Tests of this kind don't entirely agree with the conventions of English writing. For example, we can't really stick other words in the middle of compound words like *swim team* and *picture frame*, at least not while maintaining the meanings and relationships of the words we started with. In this sense they are not very different from the morphemes in complex words like *re+calibrate* or *consumer+ism*, which we write "solid", i.e. without spaces. The question of whether a morpheme sequence is written "solid" is largely a matter of orthographic convention, and in any case may be variable even in a particular writing system.

Indeed, even using more reliable tests based on real data from spoken data rather than the arbitrary patterns of writing, it can sometimes be difficult to determine how to draw the line between words and morphemes. Nonetheless, **word** and **morpheme** are very useful and perhaps even indispensable concepts for our discussion of morphology.

Combining morphemes: the constituent structure of words

Now, we can say that the relationship between words and morphemes is that words are made out of one or more morphemes put together. (An example of a one-morpheme word would be *under*.)

We must ask, then, how this works. Are words just strings of morphemes, or do they have more structure, like sentences do? It turns out that words are like sentences, i.e. they have internal constituent structure.

This can be demonstrated with English examples. Notice two uses of the prefix *un-*. **UN-** added to a verb gives another verb.

tie, un-tie
cover, un-cover
button, un-button
cage, un-cage

The meaning it conveys is called "reversative." Prototypically it is used only with a rather restricted type of verb.

UN- can also be added to an adjective to give another adjective with a simple "not" meaning. This use is very productive.

happy, un-happy
true, un-true
important, un-important
ashamed, un-ashamed

We'll now see that this ambiguity in the use of the prefix can lead to an ambiguity in words that contain it, which can only be understood in terms of a hierarchical structure.

Consider the example unusable. It contains three morphemes:

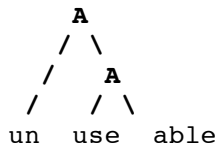
1. prefix un-
2. verb stem use
3. suffix -able

What is the structure?

Is it first use + able to make **usable**, then combined with un- to make unusable?

Or is it first un + use to make **unuse**, then combined with -able to make unusable?

Since unuse doesn't exist in English, while usable does, the first structure is correct.



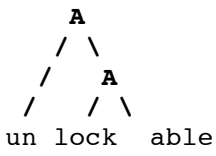
This analysis is supported by the general behavior of these affixes. As we saw, there is a prefix un- that attaches to adjectives to make adjectives with a negative meaning (unhurt, untrue, etc.). And there is a suffix -able that attaches to verbs and forms adjectives (believable, fixable, readable). This gives us the analysis pictured above. There is no way to combine a prefix un- directly with the verb use, so the other logically possible structure won't work.

Now let's consider the word unlockable. This also consists of three morphemes:

1. prefix un-
2. verb stem lock
3. suffix -able

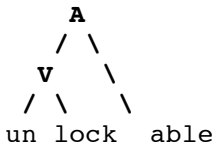
This time, though, a little thought shows us that there are two different meanings for this word. One is "not lockable," as with a box that simply has no latch on it.

> *Don't store your money in that box, it's unlockable.*



The second meaning is "able to be unlocked," in contrast with something that can't be unlocked because it's rusted shut or the key is missing.

> *Now that we have the right key, the box is finally unlockable.*



These two structures permit us to account for the two senses of unlockable.

> We can combine the suffix -able with the verb lock to form an adjective lockable, and then combine the prefix un- with lockable to make a new adjective unlockable, meaning "not able to be locked".
>
> Or we can combine the prefix un- with the verb lock to form a new verb unlock, and the combine the suffix -able with unlock to form an adjective unlockable, meaning "able to be unlocked".

By making explicit the different possible hierarchies for a single word, we can better understand why its meaning might be ambiguous. Because use is **not** a verb that effects a change, it cannot form the derived word *unuse. So unusable must be based on usable, and therefore is not ambiguous.

Types of morphemes and how they are combined

Now that we've seen the general mechanism for morpheme combination, we can look at what types of morphemes are and the specific ways in which they come together to form words.

Morphemes are usually discussed in terms of binary oppositions. I.e., a morpheme is either of type x or type y. To a certain extent, the distinctions overlap, but never completely, and each distinction demonstrates a different property of natural language morphology, so we will go through them in turn here.

Bound versus free

There are two basic types of morphemes according to their freedom of occurrence.

- **bound morphemes: cannot occur on their own as full words**
  - **-s in dogs**
  - **de- in detoxify**
  - **-ness in happiness**
  - **cran- in cranberry**
- **free morphemes: can occur as separate words**
  - **dog**
  - **walk**
  - **berry**
  - **yes**

In a morphologically complex word -- a word composed of more than one morpheme -- one constituent may be considered as the basic one, the core of the form, with the others treated as being added on. The basic or core morpheme in such cases is referred to as the stem or root, while the add-ons are affixes. Affixes that precede

the stem are called prefixes, while those that follow the stem are suffixes.

Thus in rearranged,

re- is a prefix,
arrange is a stem, and
-d is a suffix.

Often a distinction between root and stem is made, though the details depend on the language at hand. Pinker uses them in the following way.

ROOT: The most basic morpheme in a word or family of related words, consisting of an irreducible, arbitrary sound-meaning pairing: *electricity*, *electrical*, *electric*, *electrify*, *electron*.

This is essentially any bound morpheme, excluding affixes.

STEM: The main portion of a word, the one that prefixes and suffixes are stuck onto. So associated with the root *electr-* we have stems like *electrify* and *electron*, to which we can add further endings to get *electrifies* and *electrons*

In English, stems can also appear as independent words without additional endings, but in some languages, stems are always followed by a suffix in order to make the word complete. For example in Latin, there is a root *agr-* having to do with fields and agriculture, from which are formed stems like *agro-*,

"field". This stem cannot occur on its own, but must have a suffix indicating case and number (which we'll discuss below), e.g. an /s/ for the accusative plural as in *Properavi ad agros* "I hastened to the fields."

Whereas a root is normally a single morpheme, a stem might contain two or more. For example, a compound noun might function as a stem for the addition of the plural suffix.

Morphemes can also (more rarely) be infixes, which are inserted within another form, rather than before or after.

The ancestor of most of the languages of Europe, which we will talk about in the lecture on historical linguistics, had an infix /n/ that marked certain verb stems as present. This can still be seen in a few relics in Latin. For example, 'I conquer' is *vinco*, with an /n/, but I conquered is *vici*, without the /n/, as in Julius Caesar's famous quote " *Veni, vidi, vici*", 'I came, I saw, I conquered.'

English doesn't really have any infixes, except for certain expletives in colloquial expressions like these:

fan-fucking-tastic
Missi-fucking-ssippi
***fantas-fucking-tic**

This is "infixation" because the expletive goes inside a morpheme, not between morphemes. It's not random,

however: the expletive precedes the main stress of the word, which is why it sounds ridiculous to say *fantas-fucking-tic. As we've seen before, even nonstandard expressions follow rules.

Prefixes and suffixes are almost always bound. Stems are most often free in English, but sometimes are bound. Here are some words containing bound stems (or "roots").

ruth-less
grue-some
un-kempt
cran-berry

Sometimes these are called "morphans" (i.e. morphological orphans), and they're the basis of a whole class of bad jokes.

Internal changes

A range of morphological processes involve not the addition of some element (such as a suffix) but rather some change in the stem.

In English, some irregular inflections involve internal changes of this type -- for example, the past tense and past participle.

swim	swam	swum
drink	drank	drunk
begin	began	begun
sit	sat	
win	won	
come	came	
run	ran	
shine	shone	
find	found	

Many verbs, such as wear / wore / worn, show a combination of pure internal change (for the past tense) and irregular suffixation (for the past participle).

A small number of noun plurals also have internal changes.

foot	feet
mouse	mice
man	men

In Modern English these are all irregularities. There are no morphological categories that are regularly marked by internal change. But the pattern shown by the verbs is what's leftover from an older system that was once quite regular. If we go back far enough, we find that the languages from which English descends quite regularly marked tense differences by internal changes.

The most dramatic examples of internal change are found in the Semitic family of languages. For example:

> **In Arabic, noun plurals are most often formed by changing the vowels in a root.**
>
> > **kitaab** "book"
> > **kutub** "books"
>
> **In Modern Hebrew, verbs are derived from nouns in a similar way.**
>
> > **faks** "a fax"
> > **fikses** "to fax"

This type of morphology is often called templatic, where *template* refers to the patterns of vowels used in various contexts.

Content versus function - or - open class versus closed class

Morphemes can also be divided on a roughly semantic basis into categories of content and function morphemes, a distinction that is conceptually distinct from the free/bound distinction but partially overlaps with it in practice.

The idea behind this distinction is as follows:

> **some morphemes express some general sort of**

referential or informational content, a meaning that is essentially independent of the grammatical system of a particular language

other morphemes are heavily tied to a grammatical function, expressing syntactic relationships between units in a sentence, or obligatorily marked categories such as number or tense.

Thus (the stems of) nouns, verbs and adjectives are typically content morphemes: throw, green, Chris, sand are all English content morphemes.

Content morphemes are also often called open-class morphemes, because they belong to categories that are open to the addition of arbitrary new items. People are always making up or borrowing new morphemes in these categories: smurf, nuke, byte, grok, chalupa, baathist.

By contrast, the following are typically function morphemes:

- **prepositions: to, by, from, with**
- **articles: the, a**
- **pronouns: she, his, my**
- **conjunctions: and, but, although**
- **affixes: re-, -ness, -ly**

Such morphemes either serve to tie elements together grammatically:

hit by a truck

Pat and Chris
they saw their dog

or to express morphological features such as definiteness that may be required in a particular language:

she found a table
she found the table
***she found table**

Function morphemes are also called "closed-class" morphemes, because they belong to categories that are essentially closed to invention or borrowing -- it is very difficult to add a new preposition, article or pronoun.

For example, for years, some people have tried to introduce non-gendered pronouns into English, for instance sie (meaning either "he" or "she", but not "it"), but have had essentially zero success. This is much harder to do than to get people to adopt a new noun or verb, due to the basic distinction between open and closed class morphemes: the pronouns are part of a limited system, whereas normal nouns are a long list to which items can easily be added.

Inflectional versus derivational

Words are often "related" to one another in various ways. When we know these words, we understand that relationship, which often can be generalized to create new words. This is the essence of morphology.

A basic distinction in type of relationship among words

is reflected in the following terms.

> **Inflectional morphology creates new forms of the same word (in a relevant sense): the core meaning is the same, but the word reflects new grammatical properties.**
>
> > **For example, *walk* and *walked* describe the same action, but at different times.**
>
> **Derivational morphology, on the other hand, creates new words from old ones: the core meaning might change significantly, and the resulting word will still require additional inflectional morphology appropriate to the context in which it is used.**
>
> > **For example, *walk* and *walker* have fundamentally distinct (though, of course, related) meanings: one is an action, the other is a person (or a device to aid a person).**

In this section we'll explore this difference, and related issues.

Inflectional morphology

Part of knowing a word in English (or any language) is knowing how to inflect it for various grammatical categories that the language includes, such as singular / plural or past / present tense. One basic distinguishing properties of inflectional morphology is that it creates different forms of the "same" word.

For every verb, for example, a speaker can create inflected forms that express these grammatical categories. Together, this set of related forms is called a paradigm.

walk turn download gimble

walks	turns	downloads	gimbles
walked	turned	downloaded	gimbled
walking	turning	downloading	gimbling

Generally, inflectional morphology in English is entirely productive, i.e. there are not arbitrary restrictions on how the affixes are combined with stems.

> **Productivity is reflected in the fact that new words such as download participate in these inflections, as well as completely made-up words, such as gimble from the poem *Jabberwocky*.**

Even when there are irregularities in how the inflections are formed, each slot is normally filled. (A row is added here to distinguish the past tense in I walked from the participle in I have walked, since many irregular verbs distinguish these categories.)

walk	see	go	am
walks	sees	goes	is
walked	saw	went	was
walked	seen	gone	been
walking	seeing	going	being

Forms like saw and gone are irregular, since they aren't formed by simply combining a stem and the usual (or any) affix, though there's still some relation to the basic sounds of the stem.

> **A special kind of irregularity is suppletion, where there's no relation between the stem and the irregular form. Examples are went (cf. go) and am (cf. be).**

In English verbs, irregularity is most common in the past tense and past participle. Regular verbs such as walk don't even have distinct suffixes for these categories, just -ed for both.

> **Irregularity is very rare in the present tense -- it's only really found for be -- and is completely absent for the present participle, which is always formed by adding -ing.**
>
> **The 3rd singular present tense forms does and says have an irregular pronunciation of the stem vowel (i.e. "duz" and "sez"), but morphologically the suffix is regular -s added to the basic stem. So this is a different sort of irregularity than what we find in is, am, are which have no relation to the basic stem be.**

In some languages, verbs are inflected for many more categories than we find in English. Here are some verb forms in Swahili.

anapenda	"s/he likes"	alinipenda	"s/he liked me"
alipenda	"s/he liked"	aliwapenda	"s/he liked them"
atapenda	"s/he will like"	nitakupenda	"I will like you"
amependa	"s/he has liked"	nitawapenda	"I will like them"

Many of these distinctions are marked in English by other words (such as pronouns) rather than by morphology (within the same word).

Nouns in English enter into much smaller paradigms, essentially just singular and plural. (The possessive 's is actually a property of phrases, not individual nouns.) These are most often regular, taking the plural -(e)s, whose pronunciation is predictable based on the preceding sound. This is the suffix added to new words, including made-up examples such as wug.

dog	horse	byte	wug
dogs	horses	bytes	wugs

Like verbs, however, noun inflection can be irregular as well, and also suppletive.

foot	child	deer	person
feet	children	deer	people

Nouns like deer, which have no overt marking of plural (one deer, ten deer), are similar to verbs such as hit, which have no overt marking of past tense (I do hit, I hit yesterday, I have hit).

In some languages, a major inflectional category for nouns is case, which marks the relationship of the noun to a verb or preposition, or otherwise indicates its function in the sentence. Here are some noun forms in Icelandic.

hestur	"the horse [is...]"	hestar	"the horses [are...]"
hest	"[...see] the horse"	hesta	"[...see] the horses"
hesti	"to the horse"	hestum	"to the horses"
hests	"of the horse"	hesta	"of the horses"

In Modern English, only pronouns are inflected for case; this is the difference between he / him, we / us, etc. For the most part, English expresses these differences syntactically: either with word order (for subject vs. object) or with prepositions (such as of or from).

But case-marking on nouns is by no means exotic. Modern German,

Finnish, Hungarian, Japanese and Russian have it, as did most of the familiar ancient languages like Latin and Greek. In fact, Old English (spoken ca. 600-1000 CE) had it too, which should not surprise us since English is related to Icelandic and German. Indeed, the similarities to Icelandic are easy to observe:

stân	"the stone [is...]"	stânas	"the stones [are...]"
stân	"[...see] the stone"	stânas	"[...see] the stones"
stâne	"to the stone"	stânum	"to the stones"
stânes	"of the stone"	stâna	"of the stones"

Finally, some adjectives in English can be inflected for comparative ("more") and superlative ("most").

thick	big	fast	stupid
thicker	bigger	faster	stupider
thickest	biggest	fastest	stupidest

As before, we find some irregularity and suppletion.

far	good	bad
farther	better	worse
farthest	best	worst

But many adjectives -- in particular, those longer than a syllable or two -- have to express these categories by using separate words.

beautiful
more beautiful
most beautiful

intelligent
more intelligent
most intelligent

Again, this is another example of the similar functions of inflectional morphology and syntax.

Also, as with nouns and verbs, other languages have additional types of inflection on adjectives. The most common type is called agreement or concord, which is where an adjective takes endings which indicate information about the noun they modify, like whether it is singular or plural, what gender it is or what case it is in.

Consider, e.g., the difference in French between *vin rouge* 'red wine' and *vins rouges* 'red wines'.

General properties of inflectional morphemes:

- **They do not change basic syntactic category**
  - **big, bigg-er, bigg-est are all adjectives.**
- **They express grammatically-required features or indicate relations between different words in the sentence**

  - In Pat love-s Chris, -s marks the 3rd person singular present form of the verb, and also relates it to the 3rd singular subject Pat.
- They occur "outside" any derivational morphemes (closer to the edge of the word)
  - In ration-al-iz-ation-s the final -s is inflectional, and appears at the very end of the word, outside the derivational morphemes -al, -iz, -ation.
- In English, they are all suffixes. Here are the regular forms (there are also numerous irregulars).

-s	Plural	dog + s
-ed	Past	walk + ed
-s	3rd sing Present	sing + s
-ing	Progressive	say + ing
-er	Comparative	tall + er
-est	Superlative	tall + est

The idea, then, is that walk, walks, walked, walking are all specific instances (inflections) of the same basic word, rather than "new" words.

Derivational morphology

We can contrast these properties with derivational morphemes, which make new words from old ones. Thus creation is formed from create by adding a morpheme that makes nouns out of (some) verbs. Basic properties:

- change the part of speech (noun, verb, etc.) or the basic meaning of a word.
  - -ment added to a verb forms a noun (judg-ment)
  - re-activate means "activate again"
- are not required by syntactic relations outside the word.
  - un-kind combines un- and kind into a single new word, but has no particular syntactic connections outside the word
  - thus we can say he is unkind or he is kind or they are unkind or they are kind, depending on what we mean
- are often not productive or regular in form or meaning -- derivational morphemes can be selective about what they'll combine with, and may also have erratic effects on meaning.
  - the suffix -hood occurs with just a few nouns such as brother, neighbor, and knight, but not with most others.

        - **e.g., *friendhood, *daughterhood, or *candlehood.**
        - **brotherhood can mean "the state or relationship of being brothers"**
        - **but neighborhood cannot mean "the state or relationship of being neighbors"**
    - **some derivational affixes, though, are quite regular in form and meaning, e.g. -ism.**
- **typically occur "inside" any inflectional affixes (i.e. closer to the root)**
    - **in governments, -ment, a derivational suffix, precedes -s, an inflectional suffix.**
- **in English, they may appear either as prefixes or suffixes**
    - **pre-arrange, arrange-ment.**

Here are some derivational affixes in English:

-ation	is added to a **verb**	to give a **noun**
	finalize	finalization
	confirm	confirmation
un-	is added to a **verb**	to give a **verb**
	tie	untie
	wind	unwind
un-	is added to an **adjective**	to give an **adjective**
	happy	unhappy
	wise	unwise
-al	is added to a **noun**	to give an **adjective**
	institution	institutional
	universe	universal
-ize	is added to an **adjective**	to give a **verb**
	final	finalize
	sterile	sterilize

Keep in mind that most morphemes are neither derivational nor inflectional! For instance, the English morphemes Joe, twist, tele-, and ouch.

Also, most linguists feel that the inflectional/derivational distinction is not a fundamental or foundational question at all, but just a sometimes-useful piece of terminology whose definitions involve a somewhat complex combination of more basic properties. Therefore we will not be surprised to find cases for which the application of the distinction is unclear.

Category-changing derivation

As mentioned, inflectional affixes, since they create a form of the same word, don't change the syntactic category or "part of speech" of that word.

walk is a present tense **verb**
walks is a present tense **verb** (3rd person singular)
walked is a past tense **verb**
(is) walking is a progressive **verb**

dog is a singular **noun**
dogs is a plural **noun**

Some derivational affixes that create new words also happen to preserve the syntactic category.

-DOM added to a **noun** creates a **noun**.

king, king-dom
star, star-dom
martyr, martyr-dom

-STER added to a **noun** similarly creates a **noun**.

gang, gang-ster
road, road-ster
prank, prank-ster

But -dom can also be added to other parts of speech, as in freedom and boredom; and -ster can be added to verbs, as in spinster; in all cases the result is a noun, in which case the part of speech may change.

Other derivational affixes always change the syntactic category of a word, as part of their basic function.

-AL added to a **noun** creates an **adjective**.

person, person-al
cause, caus-al
tribe, trib-al

-AL added to a **verb** creates a **noun**.

betray, betray-al
dispose, dispos-al
approve, approv-al

This ability to change category is one of the best diagnostics for derivational morphology, since inflectional affixes simply create a new form of the same word (retaining its original category).

Derivational gaps

While an inflectional paradigm is characterized by all slots filled (whether by general rule or by irregular process), it is often the case in derivational morphology that there are

gaps in a chart like this.

Noun	Verb	Adjective
water	to water	water-y
knight	to knight	knight-ly
hospital	to hospital-ize	???
cat	???	cat-ty ? feline ?

Such a chart is not a paradigm of forms of the "same" word, but rather just an array of related but "different" words. What justifies this statement?

It's often not even clear what should count as filling the slot. There may be several possible candidates for a particular slot, with various meanings.

> Should the adjective for cat be catty or feline?
>
> One is closer in **form** (i.e. it shares the same root), but the other is closer in **meaning** (it refers to cats rather than the behavior of humans).

There's also not a consistent relationship between the related forms.

> The relation of talk :: talked is exactly that of run :: ran, i.e. past tense, because it is inflectional morphology that is involved.
>
> But despite the meaning of to water, the verb to knight doesn't mean "to sprinkle with knights", but rather something like "to make into a knight". This type of inconsistency is the norm for derivational morphology.

The **way** the slots are filled is often quite erratic as well. Notice some of the many ways **verbs** can be changed into **nouns** in English. (Sometimes there are changes in the form of the stem, which we won't get into today.)

suffix **-al**	refuse	refusal
	arrive	arrival
	confer	conferral
suffix **-ion**	confuse	confusion
	commune	communion
	extend	extension
suffix **-ation**	derive	derivation
	confirm	confirmation
	perturb	perturbation
suffix **-ance / -ence**	disturb	disturbance
	refer	reference
	dally	dalliance
suffix **-ment**	confine	confinement
	treat	treatment
	develop	development
"conversion": no suffix, but often with a stress change	to convert	a convert
	to permit	a permit
	to invite	an invite

While irregularity is certainly found in inflectional paradigms, it's only in derivational morphology that we find such an erratic situation. For example, notice refusal but confusion (*confusal, *refusion), arrival but derivation (*derival, *arrivation).

A more appropriate way to think about derivation is as a **network** of related words, which may vary considerably for different roots, rather than a fixed paradigm for all words of the same class.

An extreme example is a word with no derivational relatives. This is most common for borrowed words that haven't been extended to new, derived words, such as this Aleut loanword.

parka

This word has at least one other inflected form, i.e. the plural parkas, but no established derived forms.

Here's a relatively simple example with a handful of derivational relatives.

happy	happiness	
	happily	
	unhappy	unhappiness
		unhappily

And a more complex example.

grace	(to) grace		
	graceful	gracefulness	
		gracefully	
		ungraceful	ungracefulness
			ungracefully
	graceless	gracelessness	
		gracelessly	
	gracious	graciousness	
		graciously	
		ungracious	ungraciousness
			ungraciously
	disgrace	disgraceful	disgracefulness
			disgracefully
		(to) disgrace	

Notice that each word that results from a derivational process can then participate in a further derivation. For example *disgracefully* is derived from *disgraceful*, which is derived from *disgrace* which is derived from *grace*. This is quite unlike inflection, where the set of relationships is fixed by the overall grammar of the language.

Lexicalization

Because the words that result from a derivational process are new words, different from the original word, they can take on a life of their own. This process is often termed **lexicalization**, which essentially refers to becoming an independent word.

For example, RE- is added to a verb to create a new verb with the extra meaning "again."

think, re-think
fill, re-fill
create, re-create

But not all uses of RE- are of this semantically transparent (or "compositional") type -- that is, where you can take the meanings of the parts and determine the meaning of the whole.

move, remove
turn, return
form, reform

These words, in their most common uses, do not mean simply "move again," "turn again," or "form again." (If they're pronounced more deliberately, it's possible to interpret them as compositional words; cf. recreate "to relax" and re-create "to create again.")

Words like this were formed hundreds or even thousands of years ago (typically in French or Latin). Originally they had a compositional meaning, but over the centuries the meanings of the two related words became disconnected.

Other examples:

dis+comfort means essentially "lack of comfort"

but **dis+ease** no longer means simply "lack of ease"

quick+ly means "in a quick manner"

but **fair+ly** most often means "to a moderate degree" (originally from an older meaning of *fair*)

e.g. fairly expensive doesn't mean "expensive in a way that I consider fair"!

gang+ster means "member of a gang"

but **team+ster** no longer necessarily means "one who drives a team of horses"

It's because derivation creates new words that this lexicalization is possible. Clearly, speakers of a language must memorize them as independent words with potentially independent meanings.

You don't find this sort of lexicalization with inflectional morphology: walked can't refer to a different kind of movement than walks. That's not surprising if these are both forms of the same word WALK, with a single basic meaning.

The only time this happens with inflectional morphology is when older, irregular forms can take on a special meaning after they've been replaced by a new, typically regular form. Here are some examples in English.

basic word	regular inflection	old, irregular inflection
brother	brothers	brethren
old	older	elder
late	latest	last
bereave	bereaved	bereft

Since, for example, brothers has taken over for brethren, the latter word achieves a new independence (if it's not forgotten). It is no longer an inflectional form of the word *brother*, and thus is free it to shift in meaning. Not suprisingly, such words usually take on a more restricted meaning with traditional or archaic connotations.

Regularization

A type of morphological change that affects both inflectional and derivational morphemes is regularization.

Children learning a language often regularize forms that don't follow the general pattern, e.g. goed. Over time irregularities tend to be eliminated, even in adult speech, or if retained they shift to a more limited function as brethren did.

For example, some irregular verb forms have been abandoned today, but were used by Shakespeare, in what is termed Early Modern English.

> **crew** as the past tense of crow, now regular crowed, which is also found (in one use, as a past participle)
>
> It was about to speak, when the cock crew.
>
> Bernardo, Hamlet, Prince of Denmark, Act 1, Scene 1
>
> ---
>
> My lord, I did;
> But answer made it none: yet once methought
> It lifted up its head and did address
> Itself to motion, like as it would speak;
> But even then the morning cock crew loud,
> And at the sound it shrunk in haste away,
> And vanish'd from our sight.
>
> Horatio, Hamlet, Prince of Denmark, Act 1, Scene 2
>
> ---
>
> Come, stir, stir, stir! the second cock hath crow'd,
> The curfew-bell hath rung, 'tis three o'clock:
> Look to the baked meats, good Angelica:
> Spare not for the cost.
>
> Capulet, Romeo and Juliet, Act 4, Scene 4

holp as the past tense and participle of help, now exclusively regular helped, which is found in both functions as well

> Let him thank me, that holp to send him thither;
> For he was fitter for that place than earth.
>
> Gloucester, King Richard III, Act 1, Scene 2

> Sir, how comes't that you
> Have holp to make this rescue?
>
> Sicinius, Coriolanus, Act 3, Scene 1

> The last was I that helped thee to the crown;
> The last was I that felt thy tyranny:
> O, in the battle think on Buckingham,
> And die in terror of thy guiltiness!
>
> Ghost of Buckingham, King Richard III, Act 5, Scene 3

> Would I had been by, to have helped the old man!
>
> Shepherd, The Winter's Tale, Act 3, Scene 3

A similar trend can be found in modern colloquial usage such as the following.

Standard	Nonstandard
I saw it	I seen it
I had gone	I had went
I walked	(same)
I had walked	(same)

The general distinction in form between past tense and past participle, absent from regular verbs such as walk, is being eliminated from many irregular verb paradigms as well.

It should be noted as well, however, that sometimes regular verbs can become irregular by analogy with existing irregularities.

Earlier	Innovation	Source of analogy
dive :: dived	dive :: dove	drive :: drove
dig :: digged	dig :: dug	stick :: stuck

Note that dove is an American innovation; dived is still the standard in Britain (as it remains the norm in America for the past participle, i.e. have dived).

The same attention to subregularities is responsible for the nonstandard bring :: brang, which is a stronger pattern with more examples, especially ending "ng" (sing :: sang, ring :: rang) than standard bring :: brought.

Compounds

One special type of morphology stands somewhat outside these distinctions because none of the morphemes involved are necessarily bound, thus none can be considered as true derivational or inflectional morphemes.

This is the process of compounding, which is the combination of two or more stems, rather than a single stem with an affix. Although in English we often write spaces between the elements of a compound, they function as single words.

In fact, as note above, the spelling of compounds in English is rather erratic. Basically, the more familiar and standardized a compound is, the more likely it will be spelled with a hyphen or with no space at all.

> lawnmower
> classroom
> shoptalk
> pickpocket
> pushup
>
> half-life
> spoil-sport
> sit-up

Newly formed compounds, and many established ones, are written with spaces.

> high school
> credit card
> shoe polish
> drug dealer
> weed whacker
> border control officer

In other languages, such as German, the elements are written consistently without spaces, making them easier to identify.

> Kreditkarte "credit card"
> Grenzkontrolloffizier "border control officer"
> Donauschiffahrtskapitänsmützenquaste "Danube shipping captain's cap tassel"

In English, the most common kind of compound is a sequence of two or more nouns forming a single complex noun, such as olive oil, credit card, or employee training manual.

> These are "single" nouns in the sense that they can substitute in a sentence for a one-word noun, from the point of view of the syntax.
>
> > I put olive oil on the bread.
> > I put butter on the bread.
> >
> > I lost my credit card.
> > I lost my wallet.

Similarly, to pluralize a compound noun, a single -s is added (not one for each element).

credit cards (*credits cards)

Another test for compounds is that in syntactic phrases in English (such as *adjective* + *noun*), stress normally falls on the rightmost word; whereas in a compound (such as *noun* + *noun*), stress falls further to the left.

a funny cárd, an expensive cárd
a bírthday card, a crédit card

a really boring mánual
an employee tráining manual

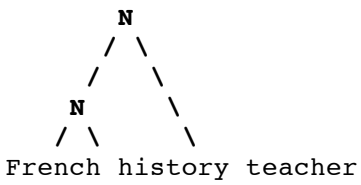
When a compound contains more than two words, understanding its meaning usually requires figuring out how to put the words together, in the same way that we need to figure out the structure of a syntactic phrase. That is, **structural ambiguities** become possible.

French history teacher
Enron document shredder

Here are two possible constituencies for the first compound, leading to ambiguity.

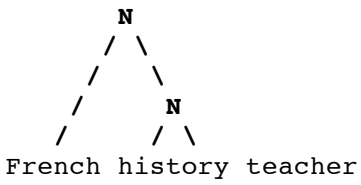
1. ((French history) teacher) = "a teacher of French history"

including someone of Spanish origin, not French



2. (French (history teacher)) = "a French teacher of history"

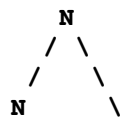
including someone who teaches about the history of Spain, not France

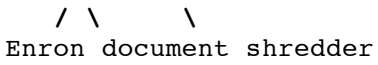


And similarly for the other example.

1. ((Enron document) shredder) = "a shredder for Enron documents"

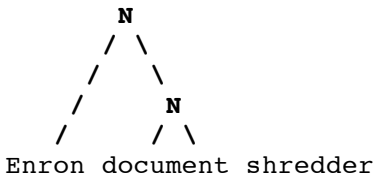
such as a shredder that happens to be used for Enron documents, though it may belong to Arthur Andersen; the term could also refer to a **person** who shreds such documents<`/p>





2. (Enron (document shredder)) = "a document shredder belonging to Enron"

> such as a device owned by Enron, whatever its use may be; possibly "borrowed" by an employee for personal use, not for Enron documents at all



This is the same idea as for unlockable, where the affixes and stem can have different constituency relationships.

1. ((unlock) able) = "able to be unlocked"

2. (un (lockable)) = "not able to be locked"

Thus the same notions of constituency apply for the structure of phrases and for the internal structure of words.

The inflection of compounds: "flied out" and "flatfoots"

Words are like syntactic phrases in that they have a main element with which subordinate elements are combined. This main element is called the **head**

So a syntactic Noun Phrase has a noun as its head, which has combined with things like adjectives and determiners. Notice that the properties of the phrase are determined by the property of the head, so a noun phrase is noun-like in its distribution, and furthermore if the head noun is singular, the NP will be singular.

This turns out to be the property of all sorts of heads, not just syntactic ones. So compound and derived words, for example, although treated by the syntax as though they were an unanalyzable unit, actually have a constituent structure, as we've seen, and they have heads.

So the compound *dog food* has *food* as its head, because *dog food* is a type of food, not a type of dog, and the head of *blackboard* is *board*, because it is a kind of board, not a shade of black, and it is a noun, not an adjective.

Inflectional properties of compounds are also determined in this way, so a compound will inflect like its head. The head of *oversee* is *see*, so the past tense is *oversaw*, and the past participle is *overseen*.

So why is that in baseball we say that the batter *flied out*, not that he *flew out*? The head of this compound verb is *fly*, so it should have the same inflection as the simple verb, shouldn't it? And why is it hard to figure out what the plural of *walkman* should be? It would seem to be headed by *man*, so why do we hesitate to say *walkmen*, when we know that the plural of *policeman* is *policemen* without question?

The answer to the first question is that, while the head of *fly out* is indeed *fly*, it is not the verb, but rather the noun, as in *a fly to shallow center field*. Now the verb *fly* does indeed

have an irregular past tense associated with it, but the noun has no past tense at all. When we make it into a verb we have to start from scratch, and all that's available is the regular past tense in *-ed*.

The problem with *walkmen* is slightly different. It turns out that *walkman* is just a made up word with no compositional meaning. It is not a kind of man at all, but a portable stereo, so *man* cannot be the head. In fact, it has no head, and in a sense it is not even really a compound, but a word that must be memorized as a unit with a meaning that is completely unpredictable from its parts and is not analyzed as being a head plus something added on. The result is the same: whatever is involved in *walkman*, the head is not *man*, and thus no irregular inflectional information can be associated with it, and the regular inflection wins out.

How does morphology fit into the grammar?

We've now talked a bit about two components or modules of the grammar, syntax and morphology, and we might wonder a bit about how the modules work together. This is perhaps the best time to talk about this question because morphology is actually fairly weird and certain respects, and it is far from clear how it should fit in.

The peculiar nature of morphology

From a logical point of view, **morphology** is the oddest of the levels of linguistic analysis. Given the basic design of human spoken language, the levels of phonology, syntax, semantics and pragmatics are arguably unavoidable. They needn't look exactly the way that they do, perhaps, but there has to be something to do the work of each of these levels.

But morphology is basically gratuitous: anything that a language does with morphology, it usually could also do with syntax; and there is always some other language that does the same thing with syntax.

For instance, English morphology inflects nouns to specify plurality: thus *dogs* means "more than one dog". This inflection lets us be specific, in a compact way, about the distinction between one and more-than-one. Of course, we could always say the same thing in a more elaborated way, using the resources of syntax rather than morphology: *more than one dog*. If we want to be vague, we have to be long winded: *one or more dogs*.

Modern Standard Chinese (also known as "Mandarin" or "Putonghua") makes exactly the opposite choice: there is no morphological marking for plurality, so we can be succinctly vague about whether we mean one or more of something, while we need to be more long-winded if we want to be specific. Thus (in Pinyin orthography with tone numbers after each syllable):

1. na4er5 you3 gou3
   there *have* *dog*
   "there's a dog or dogs there."

2. na4er5 you3 ji3 zhi1 gou3
   there have several CLASSIFIER dog
   "there's dogs there"

As an example of another kind of morphological packaging, English can make *iconify* from *icon* and *-ify*, meaning "make into an icon." Perhaps it's nice to have a single word for it, but we could always have said "make into an icon." And many languages lack any general way to turn a noun X into a verb meaning "to make into (an) X", and so must use the longer-winded mode of expression. Indeed, the process in English is rather erratic: we say *vaporize* not **vaporify*, and *emulsify* not **emulsionify*, and so on.

In fact, one of the ways that morphology typically differs from syntax is its combinatoric irregularity. Words are mostly combined logically and systematically. So when you exchange money for something you can be said to "buy" it or to "purchase" it -- we'd be surprised if (say) groceries, telephones and timepieces could only be "purchased," while clothing, automobiles and pencils could only be "bought," and things denoted by words of one syllable could only be "acquired in exchange for money."

Yet irrational combinatoric nonsense of this type happens all the time in morphology. Consider the adjectival forms of the names of countries or regions in English. There are at least a half a dozen different endings, and also many variations in how much of the name of the country is retained before the ending is added:

-ese	Bhutanese, Chinese, Guyanese, Japanese, Lebanese, Maltese, Portuguese, Taiwanese
-an	African, Alaskan, American, Angolan, Cuban, Jamaican, Mexican, Nicaraguan
-ian	Argentinian, Armenian, Australian, Brazilian, Canadian, Egyptian, Ethiopian, Iranian, Jordanian, Palestinian, Serbian
-ish	Irish, British, Flemish, Polish, Scottish, Swedish
-i	Afghani, Iraqi, Israeli, Kuwaiti, Pakistani
-?	French, German, Greek

And you can't mix 'n match stems and endings here: **Taiwanian*, **Egyptese*, and so on just don't work.

To make it worse, the word for ***citizen of X*** and the general adjectival form meaning ***associated with locality X*** are usually but not always the same. Exceptions include *Pole/Polish*, *Swede/Swedish*, *Scot/Scottish*, *Greenlandic/Greenlander*. And there are

some oddities about pluralization: we talk about "the French" and "the Chinese" but "the Greeks" and "the Canadians". The plural forms "the Frenches" and "the Chineses" are not even possible, and the singular forms "the Greek" and "the Canadian" mean something entirely different.

What a mess!

It's worse in some ways than having to memorize a completely different word in every case (like "The Netherlands" and "Dutch"), because there are just enough partial regularities to be confusing.

This brings up George W. Bush. For years, there has been a web feature at Slate magazine devoted to "Bushisms", many if not most of them arising from his individual approach to English morphology. Some of the early and famous examples, from the 1999 presidential campaign, focus on the particular case under discussion here:

> "If the East Timorians decide to revolt, I'm sure I'll have a statement."
> Quoted by Maureen Dowd in the New YorkTimes, June 16, 1999
>
> "Keep good relations with the Grecians."Quoted in the Economist, June 12, 1999
>
> "Kosovians can move back in."CNN Inside Politics, April 9, 1999

President Bush, if these quotes are accurate, quite sensibly decided that *-ian* should be the default ending, after deletion of a final vowel if present. This follows the common model of *Brazil::Brazilians* and *Canada::Canadians*, and gives Bush's *East Timor::East Timorians*, *Greece::Grecians* and *Kosovo::Kosovians*, instead of the correct (but unpredictable) forms *East Timorese*, *Greeks* and *Kosovars*. And why not? The President's method is more logical than the way the English language handles it.

Despite these derivational anfractuosities, English morphology is simple and regular compared to the morphological systems of many other languages. One question we need to ask ourselves is: why do languages inflict morphology on their users -- and their politicians?

There's no easy answer

Distinguishing morphology from syntax

So morphology does a lot of the same things that syntax does, but on a different level. This makes it somewhat difficult at times to draw the line between the two.

For example, we normal consider prepositional phrases like *in the house* and *for the glory* to be constituents put together by the syntax. Yet they often serve precisely the same functions as nouns with case-marking, like that we discussed for Icelandic and Old English.

So while in Modern English we might say *the end **of** the book*, in Old English we would have said *thaet ende thae**s** boce**s***, where we have endings on the noun and determiner instead of a preposition.

Some examples like this aren't all that problematic. There really is a syntactic difference between the two modes of expression.

But with the example at hand, there is some evidence, which would take us too far afield, to indicate that certain PPs are not really syntactic phrases at all, but just funny types of morphology, exactly like nouns with case-endings.

There are also examples going the other way. So, one might think that the possessive *'s* in English is an inflectional suffix that attaches to nouns, just like the plural *s*. After all, the two follow exactly the same rules of pronunciation, depending on the preceding sound:

Noun	Noun + s (plural)	Noun + s (possessive)	Pronunciation (both)
thrush	thrushes	thrush's	iz
toy	toys	toy's	z
block	blocks	block's	s

And neither the plural nor the possessive can be used by itself. So from this point of view, the possessive acts like a part of the noun, just as the plural does. However, the plural and possessive behave very differently in some other ways:

1. If we add a following modifier to a noun, the possessive follows the modifier, but the plural sticks with the head noun:

	Morpheme stays with head noun	Morpheme follows modifier
Plural	The toys I bought yesterday were on sale.	*The toy I bought yesterdays were on sale.
Possessive	*The toy's I bought yesterday price was special.	The toy I bought yesterday's price was special.

   In other words, the plural continues to act like part of the noun, but the possessive acts like a separate word, which follows the whole phrase containing the noun (even though it is merged in terms of sound with the last word of that noun phrase).
2. There are lots of nouns with irregular plurals, but none with irregular possessives:

Plural (irregular in these cases)	Possessive (always regular)
oxen	ox's
spectra	spectrum's
mice	mouse's

So in some ways the possessive is acting like a morphological affix, while in other it is acting like an independent word that is brought together with the NP in the syntax. In-between elements like this are called **clitics**, which comes from the Greek word meaning "to lean". That is, they are like words that can't stand up on their own and have to lean on some other word.

Thus while plural formation in English is clearly morphological, it is not clear whether the addition of the possessive clitic is morphological or syntactic.

The point is, the line between syntax and morphology is somewhat blurred.

The syntax-morphology interface

Assuming that we could somehow come up with a consistent way to draw the line between syntax and morphology, we have to wonder then how the two are related. Since they deal with very similar things, they must be tightly connected, but it is not entirely clear how they should be ordered.

Should the syntax do its work and send it off to the morphology, or vice-versa? Or should the two actually work simultaneously?

We can think about these questions in terms of the following sentence:

> He knows that they like her better.

The two main options of how to analyze the derivation of this sentence are:

1. The syntax puts together the framework of a sentence, putting some abstract version of the words together, and sending the result to the morphology. The morphology then makes sure that the words have the right forms for the positions they show up in. So it will make sure that the 3rd singular masculine pronoun at the beginning has the shape *he*, because it is the subject of its clause, whereas the 3rd singular feminine has the shape *her* because it is the object of its clause, and the verb in the first clause will show up with the form *knows* because it has a 3rd singular subject, while the verb in the second clause will be *like* because the subject is plural.
2. The morphology puts together an unordered list of words, fully formed and inflected, something like {like,that,better,her,they,he,knows}. The syntax then takes this list and tries to built a grammatical sentence out of these words that is consistent with the forms they have. Thus *he* will have to be the subject of *knows* because it is the only 3rd singular (*they* is plural and thus inconsistent with the *-s*) pronoun or noun that is in a subject form (remember, *her* is an object form). Similarly, *her* will have to be the object of *like* because it is the only noun or pronoun in object form. And so forth.

Either of these options is entirely plausible, and in fact both have been entertained at various times. Thus e.g. Chomsky's earliest theories and Panini's theory took the first option, Chomsky's theories from the 70's, 80's and early 90's took the second option, and modern theories vary in which way they go, with some of them avoiding the question entirely by denying that there is a real distinction between syntax and morphology.

Debates of this kind are extremely common within linguistics, and can be found for just about every pair of modules of the grammar. They may seem like chicken-and-egg debates, but they tend to get people very excited.

Morphology FAQ

These questions and answers are based on some patterns of error observed in homeworks and exams in previous years.

Can a word = a morpheme?

Yes, at least in the sense that a word may contain exactly one morpheme:

Word (=Morpheme)	Word Class
car	noun

thank	verb
true	adjective
succotash	noun
gosh	interjection
under	preposition
she	pronoun
so	conjunction
often	adverb

Are there morphemes that are not words?

Yes, none of the following morphemes is a word:

Morpheme	Category
un-	prefix
dis-	prefix
-ness	suffix
-s	suffix
kempt (as in unkempt)	bound morpheme

Can a word = a syllable?

Yes, at least in the sense that a word may consist of exactly one syllable:

Word	Word Class
car	noun
work	verb
in	preposition
whoops	interjection

Are there morphemes that are not syllables?

Yes, some of the following morphemes consist of more than one syllable; some of them are less than a syllable:

Morpheme	Word Class
under	preposition (> a syll.)
spider	noun (> a syll.)
-s	'plural' (< a syll.)

Are there syllables that are not morphemes?

Yes, many syllables are "less" than morphemes. Just because you can break a word into two or more syllables does not mean it must consist of more than one morpheme!

Word	Syllables	Comments
kayak	(ka.yak)	neither **ka** nor **yak** is a morpheme
broccoli	(bro.ko.li) or (brok.li)	neither **bro** nor **brok** nor **ko** nor **li** is a morpheme
angle	(ang.gle)	neither **ang** nor **gle** is a morpheme
jungle	(jung.gle)	neither **jung** nor **gle** is a morpheme

So (if you were wondering -- and yes, some people have trouble with this) there is no necessary relationship between **syllables**, **morphemes**, and **words**. Each is an independent unit of structure. Syllables are actually units of phonological structure, which we will discuss in the next lecture.

What are the major differences between derivational and inflectional affixes?

First, it's worth saying that most linguists today consider this distinction as a piece of convenient descriptive terminology, without any fundamental theoretical status. Then we can point to the basic meanings of the terms: derivational affixes "derive" new words from old ones, while inflectional affixes "inflect" words for certain grammatical or semantic properties.

	derivational	inflectional
position	closer to stem	further from stem
addable on to?	yes	not in English
meaning?	(often) unpredictable	predictable
changes word class?	maybe	no

Are clitics inflectional or derivational morphemes?

The answer would depend on your definitions -- and as we explained earlier, the categories of "inflection" and "derivation" are descriptive terms that really don't have a strong theoretical basis. However, based on comparison to typical examples of inflectional and derivational affixes, the answer seems to be "neither", in that clitics are not really lexical affixes at all.

If time allows, we'll talk about the **distributional method** for determining morpheme boundaries.

[Ling 001 Homepage] [Class Schedule]