23 September 2014
Linguistics 203

Aaron J. Dinkin
aaron.dinkin@utoronto.ca

Lecture 3
Morphology and word formation

How is the **meaning** of a word derived from or related to its **structure**?

At base, the relationship of a word's **form** and **meaning** is **arbitrary**—
there's no **fundamental** connection between the two.
The word *dog* means 'dog' just because **speakers of English all agree** that it does;
the sound "dog" **itself** has nothing to do with the animal,
and other languages of course use totally different sounds for that meaning.

But many words **do** have **internal structure** that meaning can be derived from!
Dog is a **simplex** word that has no such internal structure;
but the plural ***dogs*** does!
Dogs consists of :

- *dog*, which contributes the main meaning (the type of animal),
- *-s*, which makes the word **plural**.

So the meaning of *dog* can be derived systematically from smaller components.
These components are called **morphs**.
Morphology is the study of these structural units,
and **how they combine** to form words.

The **morph** is the smallest unit of word structure:
i.e., the smallest unit that affects a word's **meaning or stuctural behavior**
and **cannot be divided** into smaller structural parts.
A simplex word like *dog* consists of **smaller units**—
e.g., the consonants /d/ and /g/, the vowel /ɒ/—
but they **don't contribute** to the word's meaning or behavior.

Another simple example: ***un-***.
it appears in **many words** with the same **simple meaning**, 'not':
un-desirable, *un-likely*, *un-inspired*, *un-happy*, *un-sophisticated*, etc.
A slightly subtler example: ***phon-***.
it means 'sound', and recognizably carries that meaning in many words—
phone, *phonetics*, *telephone*, *phonology*, *euphony*, *symphony*, *phonics*, etc.—
even if it's not as straightforward as *un-* is in how it combines.

morph vs. morpheme

We've used the term **morpheme** already; it refers to a **slightly more abstract** entity:
a **morph** is a **specific form** that appears as a structural unit in a given word;
a **morpheme** consists of **one or more morphs** with the same origin and function.
Thus the **morpheme** *phil-*, meaning 'liking', appears as the **morph** *phil-* in *philanthropy*,
but as the slightly different morph *-phile* in *Anglophile*.
We'll talk more about how different **morphs** of the same **morpheme** relate to each other
over the next few classes.

A morph in English can be of **any length**, e.g.—
- one consonant: *-s*, as in *dog-s* and *cat-s*
- one syllable: *dog*, *cat*, *un-*
- two syllables: *anti-*, as in *anti-freeze*, *anti-tank*
- many syllables: *Madagascar*

Even though *mad*, *gas*, and *car* are all morphs that appear in some English words,
Madagascar is only **one morph long**—*mad*, *gas*, and *car* **don't play a role** in its meaning, structure, or etymology; the similarity is coincidental.

An English word can have **any number of morphs**:
- *boy*
- *boy-ish*
- *boy-ish-ness*
- *gentle-man-li-ness*
- *un-gentle-man-li-ness*
- … etc.!

Some morphs have meanings or functions that are **easy to recognize**:
e.g., *phon-* means 'sound', and it recognizably carries that meaning in *phone*, *phonetics*, *telephone*, *phonology*, *euphony*, *symphony*, *phonics*, etc.
Others seem harder to attach a consistent meaning to:
e.g., *-ceive* appears in *receive*, *conceive*, *perceive*, *deceive*—
we can tell it's a morph because of its consistent **structural behavior**,
but we **can't tell at first glance** what **meaning** it contributes.
Some morphs contribute **only structural function**, not meaning—
e.g., *-ty* in *dignity*, *rapidity*, *variety* means just that the word is a **noun**.

How far do we go in dividing words into structural components?

It depends **what our goal is** in performing the analysis.

Consider the word ***appendectomy***.
most English speakers would be able to indentify about **two morphs** here:
- *append-* meaning 'appendix'
- *-ectomy*, meaning 'surgical removal'.

From the standpoint of **synchronic grammar**—
i.e., the grammar of the language **as it exists in speakers' minds**—
that might be all there is.

But we're more concerned with **etymology** in this class—
the **origin and history** of words and how they come to mean what they do.
So we'll use a **more expansive concept** of what counts as a morph—
the **smallest** unit that has an indentifiable structural contribution
in the context of the history of the word and the vocabulary as a whole,
even if it's not obvious to the average speaker.

From the **etymological** point of view, *appendectomy* consists of **five** morphs:

ap-	'to': from Latin
pend-	'hang': from Latin
ec-	'out': from Greek
tom-	'cut': from Greek
-y	noun-forming ending

One of the goals of this course:
to learn to **recognize** and **interpret** such morphs and their meanings,
in order to be able to understand how they relate to the meanings of words.
Once you know the common morphs and how they combine,
it becomes easier to **parse** English vocabulary elements
and understand the origin of their meaning.

E.g., *concurrent* can be analyzed as *con-curr-ent*:
- *con-* means 'together', as in *con-tract* ('pull together'), *con-fuse* ('melt together')
- *curr-* means 'run', as in *re-curr-ing* ('running back), *curr-ent* ('running')
- *-ent* forms adjectives

So the meaning of *concurrent* **originates** as 'running together'.

Categorization of morphs (and morphemes)

Free vs. **bound**:
- **Free** morphs are those that can **stand on their own** as a word:
  boy, *dog*, *cat*, *kick*, *berry*
- **Bound** morphs are those are always **attached to other morphemes**:
  -s in *dog-s*; *-ing* in *kick-ing*; *cran-* in *cran-berry*, *dign-* in *dign-ity*.

(A **hyphen** is usually used when citing a bound morph in isolation,
to indicate **where it usually attaches** to the rest of the word.)

Content vs. **function**:
- **Content** morphs (a.k.a. **lexical** morphs) have **referential meaning**:
  they usually refer to a specific thing, concept, action, quality, etc.
  - e.g., *dog*, *kick*, *purple*, *tennis*, *dign-*
- **Function** morphs usually embody **structural** or **grammatical** information about content morphs, or **modify** or **specify** shades of meaning:
  - *-ness* indicates a word is a noun; *and* connects phrases
  - *-ed* means 'in the past'; *un-* means 'opposite'; *-s* means 'more than one'

Most function morphemes are **bound**, but **free function morphemes** exist too:
the, *of*, *and*, *be*, etc.
English possesses both **free and bound** content morphemes in large numbers:
the focus of this class is **bound content morphemes** from Greek and Latin.
English makes **much more use** of **free morphemes** than many languages do;
in Latin and Greek, for instance, **almost all content morphemes** were **bound**.

Complex words usually consist of a **root** plus one or more **affixes**.

The **root** is a content morph usually contributing the core of the meaning:
dog is the root of *dog-s*; *black* is the root of *black-en-ing*;
-ceive is the root of *re-ceive* (even though *-ceive*'s own meaning is vague).

An **affix** is typically a **bound function morpheme** contributing information about finer shades of meaning, structural relationships, etc.
Types of affixes:
- **prefixes** are attached before the root: *un-* and *fore-* in *unforeseen*
- **suffixes** are attached after the root: *-en* and *-ing* in *blackening*

We'll use the term **base** for **what an affix is added to**:
this may be just the root itself, or the root plus some affixes already added.
E.g., in *blackening*, *-ing* is attached to the **base** *blacken*,
and *blacken* itself consists of *-en* attached to the base *black* (which is the **root**).

Some languages use **infixes** as well—
affixes that are attached **inside** their base, rather than before or after.
Latin used an infix *-n-* for the **present tense** of **some** verbs:
the stem *vic-* 'conquer' had the present tense *vinc-ō* 'I conquer'.
English morphs borrowed from such Latin verbs **may or may not** include the *-n-*:
compare English *con-vince* and *con-vic-tion*.

A **very small number** of native English words **also** show traces of infix *-n-* in their present tenses:
compare *stand~stood*, *bring~brought*, *think~thought*.
This is a relic from one of the earliest ancestors of English,
and obviously not part of the active grammar anymore.
A few forms in English **can** be used productively as **infixes**, as in *abso-fucking-lutely*.

In English, **suffixes** often determine what **grammatical category** a word is; e.g.:
- *-ty* makes nouns (*digni-ty*, *varie-ty*, *rapidi-ty*)
- *-ize* makes verbs (*personal-ize*, *organ-ize*, *sympath-ize*)
- *-ous* makes adjectives (*cauti-ous*, *cavern-ous*, *furi-ous*)

Prefixes **usually** don't change the part of speech, but there are some exceptions:
de- attaches to nouns and produces verbs (*de-claw*, *de-louse*, *de-bug* etc.).

Many affixes only attach to bases of a **specific category**.
For instance, *un-* attaches to **adjectives** and **verbs**, but not usually **nouns**:
unaware and *undo* are words, but not *unknowledge* or *undeed*.
You can sometimes figure out the **internal structure** of a word based on this:

Does *unhappiness*, consist of *un-* attached to the base *happiness*,
or *-ness* to the base *unhappy*?
It **must be** the latter—*unhappy+ness*—because *un-* doesn't attach to nouns.

Some words can go both ways!
e.g., *unlockable* could be **either** *un+lockable* **or** *unlock+able*.
Note that the two structures have **different meanings**:
'not lockable' vs. 'able to be unlocked'.

How are new words **created**?
What processes of word formation exist in English?

Derivation is one of the most common:
adding one or more affixes to a base to create a word with a **different meaning** and/or different **category** than the base.
Affixes that have this function are called **derivational affixes**;
in English, the large majority of affixes are derivational,
including most of the examples above: *-ty*, *-ize*, *-ous*, *de-*, *un-*, *-ness*, etc.

English has **two classes of derivational affixes**:

Class 1:
- can trigger **changes in the pronunciation** of the base
- often affects placement of **syllable stress**
- can combine with **bound roots**
- mostly consists of affixes **borrowed from Greek and Latin**

Examples:
- *-(i)ty*:
  - *sane* → *sanity*: changes the vowel in the base
  - *public* → *publicity*: changes /k/ to /s/, and syllable stress
  - *dign-* → *dignity*: base is not a free morpheme
- *-(t)ive*:
  - *destroy* → *destructive*: changes pronunciation of base
  - *fug-* → *fugitive*: base is not a free morpheme

Class 2:
- has **no effect on pronunciation** of the base
- **rarely** affects which syllable is stressed
- mainly combines only with bases that are themselves **standalone words**
- mostly consists of **native** English affixes

Examples:
- *-ness*: *promptness, happiness, strangeness*
- *-ful*: *hopeful, delightful, wonderful*
- *-less*: *hairless, baseless, friendless*

Some derivational affixes do have **specific meanings**—
-less means something is absent,
-ful means something is present (to a large degree)
post- means 'after',
un- means 'not', or reversal of an action—
while others only indicate **grammatical** functions—
-ness and *-ty* turn adjectives into nouns, but don't add any further meaning.
Either way, the contribution of the **root** is in general **very important**.

Some morphs are **empty**—they seem to have no meaning or structural function,
but just provide a **transition** from one morph to the next.
E.g., the *-a-* in *flotation* connects the root *float* with the suffix *-tion*.

Inflectional affixes contrast with derivational affixes:
inflection creates **different forms of the same word,** indicating **specific grammatical information**, but **not changing the core meaning**.

- *govern*, *governs*, *governed*, and *governing* differ in **inflectional** suffixes.
- *govern*, *governor*, and *government* differ in derivational suffixes.

Inflectional affixes represent **grammatical markers**,
encoding information about verb tense, number, gender, etc.

English only has about seven inflectional morphemes:

• third-person singular present tense	*She wait-s.*
• past tense	*She wait-ed.*
• progressive/present participle	*She is wait-ing.*
• passive/perfect participle	*She has eat-en the cookies.*
• plural noun	*She has eaten the cookie-s.*
• comparative adjective	*I am old-er than you.*
• superlative adjective	*I am the old-est.*

Inflectional morphology is **paradigmatic**—
in general, **every** word of a given category can be **inflected** for all the categories the language uses inflection for.
I.e., every noun has a plural, every verb has a past tense, etc.,
and part of **knowing a word** is **knowing its inflected forms**.

We thus can think of *wait*, *waits*, *waited*, *waiting* as **forms of the same word**—
this sense of "word" is also called a **lexeme**.
Intuitively, a lexeme may be thought of as a "dictionary entry".
Derivation creates **different lexemes**; inflection, various **forms** of one lexeme.
Derivational affixes are almost always **closer to the root** than inflection.

Some words use **irregular** morphs for their inflected forms—
e.g., the plural of *ox* is *oxen*, using the morph *-en* instead of *-s* to mark plural.

Some words go even farther with irregular inflection—
they **alter** or **replace** the **root**, instead of (or in addition to!) adding an affix.
Changing a vowel in the stem is called **ablaut**:
sing~sang~sung; *drive~drove*; *get~got*; *mouse~mice*; *man~men*
(Ancient Greek used ablaut **derivationally**; we'll see this in some loanwords from Greek.)

Some lexemes **replace** the root with an **unrelated** root morph for inflection:
the past tense of *go* is *went*; the comparative of *bad* is *worse*.
This process is called **suppletion**.

In some cases, the inflected root morph **isn't totally unrelated** to the basic root,
but it doesn't seem to have any **systematic** relationship to it—
e.g., *bring~brought*, *was~were*.
This is sometimes called **partial suppletion**.

Compounding

Compounding: creating a word with **two or more lexical roots**
(and possibly some number of affixes as well).
In compounding, instead of having a **base** and adding a subordinate **affix**,
two bases are combined on more or less equal terms.

Sometimes the bases are **bound morphemes**, as in words like *philanthropy*—
that's fundamentally a **compound** of *phil-* 'liking' and *anthrop-* 'human'.
Greek and Latin compounds often **separate** their roots with **empty morphs**:
usually *-i-* for Latin (*ped-i-cure*) and *-o-* for Greek (*bi-o-sphere*).

But many compounds are composed of two **free morphemes**—
classroom, *whitewash*, *bluebird*, *spoonfeed*, *caveman*, *smokestack*, etc.—
or even **morphologically complex** bases:
action movie, *election campaign*, *biology department*, etc.

The bases of a compound can even be compounds themselves—
weekend concert series is a compound of two compounds.

Note that some compounds are **written** as as single word (with no space),
and others are written as two words with a space (or hyphen) between them.
This is **just an artifact of spelling**!
Structurally, they're still compounds either way.
Usually **more commonly used compounds** and those made up of **shorter bases**
are more likely to be written as a single word.

How can you **tell the difference** between a compound and a **two-word phrase**?
Sometimes it's easy: English doesn't allow a phrase consisting of **two nouns**,
so if it's noun+noun, it must be a **compound**: *caveman*, *action movie*.
What if it's a combination that **could** be **either** a phrase or a compound?

- The compound's **meaning** may be **less directly** based on its components than a phrase's meaning.
- **Most** compounds have primary **stress** on the first stem; phrases may be stressed on the second word.

Examples:

- *green house* just means 'a house that is green', and is stressed on *house*; *greenhouse* is stressed on *green* and means 'glass-enclosed garden'.
- *wet suit* means 'suit that is wet'; *wetsuit* means 'diving outfit'.

Most two-word compounds in English are noun+noun producing another noun;
but lots of other possibilities exist!

N+N = N: *classroom*	A+N = N: *greenhouse*	V+N = N: *pickpocket*
N+V = V: *spoon-feed*	Prep+N = V: *downgrade*	N+A = A: *pitch-black*
V+V = V: *stir-fry*	V+Prep = A: *tumbledown*	V+P = N: *sit-in*

…and many more!

Many compounds refer to a **sub-type** of what one of their **bases** refers to;
that base is called the **head**.
E.g.: a *blackboard* is a **kind of board**, so *board* is the **head** of *blackboard*;
lemon-yellow is is shade of yellow, so *yellow* is the head of that.
Compounds of this type are called **endocentric**.

Some compounds don't have a head—
a *redneck* isn't a type of neck; a *pickpocket* isn't a type of pick or pocket; etc.
These are called **exocentric**.

Other methods of word formation

Conversion: assigning an **existing** word to a **new syntactic category**—
noun *paper* becomes a verb, as in *They papered the walls.*
verb *drink* becomes a noun, as in *I need a drink.*
adjective *dirty* becomes a verb, as in *Don't dirty the floor.*
This is also known as **zero derivation**—
it's like adding a derivational suffix, except the suffix is "zero" (i.e., not there).

Eponyms are related to zero derivation:
converting a **proper name** into a common word.
E.g., *sandwich* is named for the Earl of Sandwich, who supposedly invented them;
boycott is named for Charles Boycott, who was the target of one.

Back-formation: creating a word by **removing what looks like an affix** from an existing word, even if it wasn't originally an affix.
E.g., *pea* and *cherry* are derived from **removing the final *-s*** from *pease* and *cherise*:
it was **misinterpreted** as the plural *-s*, and removed to create singulars.
Escalator was originally a **brand name**; *escalate* was created by removing the *-or*.

Blending: creating a word by **combining parts of two words**:
breakfast + lunch = brunch; *smoke + fog = smog*; *chuckle + snort = chortle*.
They're also called **portmanteau words**, from a joke in *Through the Looking-Glass*.

Acronyms: creating a word from the first letters or syllables of a phrase:
radar comes from ***ra**dio **d**etection **a**nd **r**anging*;
scuba comes from ***s**elf-**c**ontained **u**nderwater **b**reathing **a**pparatus*.
Acronyms are relatively rare in English before the 20th century or so,
though there are lots of **fake etymologies** claiming that various older words are acronyms.

Onomatopoeia: creating a word to refer to and **imitate** some sound:
buzz, *hiss*, *beep*, *zip*, etc.