30 September 2014
Linguistics 203

Aaron J. Dinkin
aaron.dinkin@utoronto.ca

Lecture 4
Introduction to allomorphy

A **morph** is a minimal structural component of a **particular word**.
But often several similar morphs have **the same function and origin**,
so that they seem to be **slightly different versions** of a more **abstract** entity.
That entity is the **morpheme**.

E.g., consider the words *dogs* and *foxes*.
They each end with a **plural suffix**: *dogs* with *-s*, and *foxes* with *-es*.
These suffixes are **different morphs**!
But it makes sense to think of them as just **different forms** of the **same suffix**.
Dogs and *foxes* end with different **morphs**, but the same **morpheme**:
the plural **morpheme** *-s* can appear as **different morphs** in **different words**.
These are the **allomorphs** of the morpheme.
(We often refer to a morpheme just by the name of its **most common** allomorph: *-s*, in this case.)
Thus the **morpheme** *-s* has the **allomorphs** *-s* and *-es*.

Another example: the prefix *in-* meaning 'not'. Its **allomorphs** include:
- *in-* : *in-delicate*, *in-tangible*, *in-controvertible*, *in-accessible*
- *im-* : *im-possible*, *im-material*, *im-penetrable*
- *il-* : *il-legible*, *il-legal*
- *ir-* : *ir-responsible*, *ir-regular*
- …and one or two others.

These are all **different forms** of the **same basic prefix** *in-*.

In some cases, the allomorph that appears in a given word depends on **basic properties of the pronunciation** of the **context** the morpheme appears in.
A simple example: the **Greek** prefix *a-* meaning 'not' or 'without'
has the **allomorph** *an-* when it's **followed by** a vowel (or *h*).
- *a-* : *a-theism*, *a-moral*, *a-pnea*
- *an-*: *an-oynmous*, *an-hydrous*, *an-aesthetic*

In many cases, however, allomorphy is more **idiosyncratic**,
and has to be learned on a **case-by-case basis**.

For instance, *wise* appears as the allomorph *wis-* in *wisdom*.
This is just a **special fact** about the morpheme *wise* and the word *wisdom*—
other roots don't change before the suffix *-dom*, for instance.

Note that both **root morphemes** and **affixes** can have multiple allomorphs!

How can we tell whether two **morphs** are **allomorphs** of the **same morpheme**?
Three basic principles.

1. Allomorphs of the same morpheme have the **same function** (and, usually, the same approximate **meaning**).
2. The choice of **which allomorph to use** depends on **other morphemes** in the same word.
3. Allomorphs share a **common history** and related **pronunciation**.

NB: we define morphemes from a **historical and etymological** perspective,
not a **synchronic and psychological** perspective.
A typical native speaker of English **might not be aware** of the historical allomorphic relationships between different words;
this class helps you **go beyond** the knowledge that typical English speakers have of the structure of their vocabulary.

Principle #1 **can be tricky**, because meanings can **change over time**,
so a morpheme **might not have the same meaning** in every word.
But we can use principles #2 and #3 to help clarify.

E.g., the prefixes *in-*, *im-*, and *un-* have the same **meaning and function**.
But *in-* and *im-* are allomorphs of the **same morpheme**, and ***un-* isn't**.
How can you tell?
The choice between *in-* and *im-* depends on **what stem it's attached to**—
specifically, what sound the stem begins with. (More on this later.)
There are words like *impossible* and *indelicate*, but not *in-possible* or *im-delicate*.
But the availability of *un-* **isn't dictated** in that way—
in fact, *un-* can sometimes be attached to the **same** bases as *in-*/*im-*:
inescapable and *unescapable* are both words!
So *un-* isn't an allomorph of the same morpheme as *in-* and *im-*.

E.g., consider the morph *path-* in *sympathy* and *pathology*.
In *sympathy* it means 'feeling'; in *pathology* it means 'illness'.
These don't seem to have the same meaning! Are they the **same morpheme**?
From an etymological standpoint, **yes they are**; they have a **common origin**.
In Ancient Greek, the morpheme *path-* meant 'suffering';
it's easy to see how 'feeling' and 'illness' are **both related** to that meaning.
However, the *path* in *footpath* is **a different morpheme**;
it **has the same form**, but an **unrelated meaning and origin**.
(Note that this means **different morphemes** can end up having the same form!
Another example: the noun plural suffix *-s* has the same form as the third-person singular verb suffix *-s*, but they are **unrelated** and are **different morphemes**.)

Common origin is our **principal guideline** in this class for morpheme identity.
the textbook defines morphs as belonging to the same morpheme
if they have the **same origin** in Latin or Greek.

Sometimes we can identify morphemes **through** patterns of allomorphy.

E.g., consider the words *permit*, *commit*, *submit*, *emit*, etc.
They all contain the syllable *-mit*.
Is *-mit* **the same morpheme** in these words? Or are they **coincidentally** similar?
There **isn't** a clear element of **meaning** that they all share.

We can tell *-mit* is a single morpheme **because of its allomorphy**:
it has the **allomorph** *miss-* when the **suffix** *-ion* is added, in **all** of these words:
permission, *commission*, *submission*, *emission*, etc.
The fact that we see the **same alternation** between *mit* and *miss* in these words
indicates that they all **share** the same morpheme.
On the other hand, *vomit* **doesn't contain** the morpheme *-mit*;
vomit + *-ion* produces *vomition*, not "*vomission*".

(The morpheme *mit-* comes from the Latin for 'send',
and we can see traces of that meaning in many of the derived English words:
e-mit = 'send out', *trans-mit* = 'send across', etc.—
But only really if we **already know** what we're looking for.)

A more complicated example:
the morph *-ceive* appears in *perceive*, *receive*, *deceive*, and so on.
It comes from the Latin for 'take', and we can see shades of that meaning:
re-ceive = 'take back', *per-ceive* 'take through'.
Its allomorph is *cep-* when *-tion* is added: *perception*, *reception*, *deception*.
(Note that *-tion* and *-ion* are allomorphs of the same morpheme as well!)
This alternation is fairly **distinctive**—
similar morphemes **don't** show similar allomorphy;
the relationship between *-ceive* and *cep-* is **idiosyncratic**.

There are **many other morphs** in English derived from the **same Latin root**,
though they don't all have **synchronic** relationships as clear as *ceive~cep*.
We consider them all **allomorphs** of the same morpheme anyway.

- *cip-*, as in *re-cip-e*, *re-cip-ient*, *in-cip-ient*
- *cap-*, as in *cap-ture*, *cap-tive*
- *cup-*, as in *oc-cup-y*, *con-cup-iscent*
- *cei(p)-*, as in *re-ceip-t*, *de-cei-t*

Knowing that this morpheme means 'take' in Latin will help you recognize it,
and understand **why** the words containing it mean what they do.

Note that we can combine the **same morphemes** in **different combinations**;
not all of the possible words this produces are actual English words,
but most of them are:

re-ceive	*re-cip-ient*	*re-cep-tion*	*re-ceip-t*
de-ceive		*de-cep-tion*	*de-cei-t*
con-ceive		*con-cep-tion*	*con-cei-t*
	in-cipient	*in-ception*	

Why does allomorphy occur?

Why don't morphemes **have a constant form**, in all words they appear in?

There are several reasons.

Phonology-based allomorphy

Every language has **rules** about **what sequences of sounds** are possible words.

(We'll talk more about this in the next class.)

Some sequences of **morphs** would produce **illegal combinations of sounds**,
and so a **different allomorph** of the relevant morpheme must be used
in order for the word to have a permissible pronunciation.

E.g., consider the allomorphy of the **regular plural suffix**: *cats*, *dogs*, *foxes*.

Foxes obviously has a different allomorph than the other two.

Cats and *dogs* have different allomorphs too!

In *cats*, the *-s* is pronounced as /s/; in *dogs*, it's pronounced /z/.

Why all this allomorphy?

Because some morphs would create **illegal sound sequences** on some stems.

E.g., what if we tried to use the allomorph /z/ for **all** stems:

cat + *-s* = /kætz/;

fox + *-s* = /fɒksz/.

But English **doesn't allow** words to end with the sequences /tz/ or /sz/!

(In fact, if you try to say them, they'll probably come out as /ts/ and /səz/ anyway!)

So we use **different allomorphs** of the plural suffix with these words.

The allomorphy serves to **repair** problematic sequences of sounds.

Some allomorphy is due to phonology from a **previous stage** of the language—
repairing sequences of sounds that **used to be** illegal,
even if at the **present** stage of the language the repair is no longer useful.

E.g., why does *wise* have a different allomorph in *wis-dom*?

Saying *wise-dom* **wouldn't** involve any illegal sound sequences.

The answer to this goes back to **Old English** sound patterns.

In Old English, the word *wise* had the form *wīs*, with a **long /ī/** vowel.

The noun *wisdom* thus **should have been** *wīs-dom*.

But at some point, a long vowel + 2 consonants became an **illegal sequence**.

To repair this sequence, the **vowel in the root** was **shortened**, creating *wis-dom*.

The **shortened allomorph** *wis-* has **remained** in *wisdom* since then,
even though English **no longer has a rule** against long vowels in that context.

(Recall that the **pronunciation** of the **long** vowel in *wise* changed due to the **Great Vowel Shift**,
but the pronunciation of the short vowel in *wisdom* didn't change.)

There are **several examples** of words that show allomorphy of this type,
because they go back to Old English and have **held on** to their old patterns.

Deep ~ *dep-th* and *wide* ~ *wid-th* are two more.

But **not all** morphemes that go back to Old English will exhibit this allomorphy;
since the **rule about sound sequences** itself is not longer part of English,
many morphemes have **lost** this alternation in some or all contexts.

Allomorphy in **borrowed** morphemes often goes back to the sound system rules of the **language they were borrowed from**.
This is the origin of the *in-* ~ *im-* ~ *il-* ~ *ir-* allomorphy discussed above:
Sequences like /np/, /nl/, /nr/ were illegal in Latin,
so *in-* had **allomorphs in Latin** to repair these forbidden sequences.
English borrowed the prefix *in-*, and the **allomorphs along with it**,
even though English **has never had** a rule against sequences like those.
Other prefixes borrowed from Latin show the **same** pattern of allomorphy, e.g.:
con-duct, *com-pare*, *cor-rect*, *col-lide*.
But the **native** English prefix *un-* **doesn't** have any similar allomorphic behavior:
un-demanding, *un-professional*, *un-restricted*, *un-limited*

Sometimes the **borrowed phonology** is very **similar** to native phonology,
because it's caused by the same **underlying forces** affecting pronunciation.
E.g., **both Greek and English** have morphemes with allomorphs *a* and *an*,
with *a* used before consonants and *an* before vowels:
the English **indefinite article** and the Greek prefix *a-* 'not, without':

an *apple*, ***an*** *orange*	***a*** *pear*, ***a*** *thimble*
an*-archy*, ***an****-onymous*	***a****-political*, ***a****-theism*

The **native** and **borrowed** morphemes show **similar** allomorphy,
because of a **general tendency** across languages to avoid sequences of multiple consecutive vowels, when possible.
The allomorphic patterns aren't **exactly** the same, though—
the Greek prefix in *an-* before **/h/**, but the English article is *a* before /h/:

an*-hydrous*, ***an****-hedonic*	*a hydrant*, *a hedonist*

(Some people do occasionally use *an* for the article in English before /h/: "*an historian*".
But this practice is rare and more or less obsolete.)

Sometimes the phonological motivation of allomorphy is **obscured**
by **other changes** that took place after the allomorphy originated.
Latin vowel weakening is a pattern of allomorphy that we will see a lot of.

In **early Latin**, the **first syllable** of a word was always **stressed**;
and there were rules **against** certain vowels appearing in unstressed syllables.
So morphemes containing /a/ would be "weakened" to /e/ or /i/ in mid-word:

- /e/ before **two** consonants (or /r/)
- /i/ before **one** consonant.

This is the source of some of the allomorphy of *cap-*, described above:

- *cap-* in the **first syllable**: *captive*, *capture*
- *cep-* in a **non-initial** syllable before a consonant: *con-cep-tion*, *de-cep-tive*
- *cip-* in a non-initial syllable before a **vowel**: *in-cip-ient*, *re-cip-e*

…Even though classical Latin **no longer** always stressed the first syllable,
so in many of these words it's the "**weakened**" syllable that was stressed!

Morpheme merger

Sometimes a **sequence** of morphs gets reinterpreted as a **single** morph,
often because one of the two has no particular meaning.

E.g., the Latin morph *corp-* 'body' often occurred with a **suffix** *-us* or *-or*
that had no particular meaning but just created a noun stem.
The *-or*/*-us* suffixes were **reinterpreted** as **part of the same morph** as *corp-*.
This leads to a situation where *corp-* has **allomorphs** *corpor-*, *corpus*:

corp-ulent *corpor-ate*, *corpor-eal* *corpus*, *corpus-cle*

So *corp-or-* was **originally** a sequence of two morphs;
but now it makes more sense to consider it just an **allomorph** of *corp-*.
This is an example of an **extended allomorph**.

Other examples of extended allomorphs:

ten- ~ *tend-* 'stretch': *ten-uous*, *ex-tend*
noc- ~ *nox-* 'harm': *in-noc-ent*, *nox-ious*
(Note that *nox-* is just a way of spelling *noc-s-*!)

Latin occasionally used an **infix** *-n-* or *-m-* to mark **present tense** of a verb.
So in *vincō* 'I conquer', the *-n-* was **originally** a **separate morpheme** from *vic-*.
But it soon was reinterpreted as **part of the root morph**—
so *vinc-* and *vic-* became **allomorphs** of each other.
We see this allomorphy in English morphemes borrowed from those Latin ones:

vic- ~ *vinc-* 'conquer': *con-vince*, *pro-vince*; *con-vic-t*, *vic-tor*
cub- ~ *cumb-* 'lie down': *in-cumb-ent*, *suc-cumb*; *in-cub-ate*, *suc-cub-us*

These **nasal infixes** produce another kind of **extended allomorph**.

Ablaut

Recall that English occasionally uses **ablaut**—i.e., **changing the vowel** in a root—
to express **verb tenses**: *sing~sang~sung*; *drive~drove~driven*, etc.
Latin and Greek **inherited** some use of ablaut from their **ancestor** language
("**Proto–Indo-European**"; also an ancestor of the Germanic languages, including English!)
but its purpose was often **less transparent** than simply forming verb tenses.
Thus the different **ablaut "grades"** of Greek and Latin roots—
i.e., the forms of the roots with **different vowels**—
must simply be memorized as idiosyncratic **allomorphs** of the roots.

The main **ablaut grades** of Greek and Latin are ***e*-grade**, ***o*-grade**, and **zero-grade**
(though for various reasons some words have ablaut grades with other vowels).

Examples of ablaut allomorphy:
gen- 'birth' has ***o*-grade** *gon-* and **zero-grade** *-gn-*.
(This morpheme exists in **both Greek and Latin** loanwords!)
These appear in *gen-ital*, *gon-ad*, and *pre-gn-ant*.

Sometimes a **combination** of factors can create allomorphs:
gen- also has the allomorph *na-*, as in *na-tive*, *pre-na-tal*, etc.
Why is *na-* an allomorph of *gen-*? Because…
- **ablaut** gives *gen-* the zero-grade allomorph *-gn-*;
- **extension** creates the allomorph *-gna-*;
- **phonological repair** turns *gna-* into *na-* at the beginning of a word.

Doublets

Since Greek, Latin, and English are all (distantly) **related languages**,
they have some similar-looking morphemes in common just because they all **inherited them** from their common **prehistoric ancestor**.
E.g., English *mother*, Latin *matr-*, and Greek *metr-* look similar **not** because of any borrowing, but because they're **all inherited from the same source**.
Such words in related languages are called **cognates**.

Often English borrows a morpheme that it **already has** a cognate for,
or **borrows two cognates** from related languages.
E.g., Latin *matr-* and Greek *metr-* are cognates of **each other** and English *mother*, but English borrowed them anyway.
When **a single** language contains morphs that are **cognates of each other**, borrowed from **related sources**, they are called **doublets**.
So the morphemes *mother*, *matr-*, and *metr-* are **doublets** in English.

when both of cognates appear in English. we call them doublets

Sometimes doublets are **similar enough to each other** that **within English**, they can be treated as **allomorphs** of the **same morpheme**.
E.g., *gen-* 'birth' above:
the root *gen-* means birth in **both Latin and Greek** (as cognates),
and English borrowed words with *gen-* from both languages.
So arguably English has a single morpheme *gen-* with both Greek and Latin allomorphs.

Some doublets are **recognizable through their systematic similarities**,
even if they are treated as distinct morphemes within English.
E.g., at the beginning of a word, Latin /s/ often corresponds to Greek /h/:

semi- ~ *hemi-* 'half':	*semi-circle*, *hemi-sphere*
sex- ~ *hex-* 'six':	*sex-tet*, *hex-agon*
serp- ~ *herp-* 'creep, reptile':	*serp-ent*, *herp-etology*

Other doublets in English:

Latin *ped-* ~ Greek *pod-* 'foot'	*ped-icure*, *pod-iatrist*
Latin *nomin-* ~ Greek *-onym* 'name'	*nomin-ation*, *pseud-onym*
Latin *dent-* ~ Greek *odont-* 'tooth'	*dent-ist*, *peri-odont-ist*
Latin *vin-* ~ Greek *oen-* 'win'	*vin-egar*, *oen-ophile*

Just as **Latinate** words are usually more **formal** or **specialized** than native words,
Greek words are often more formal or specialized than Latin words.
Cf. *odontologist* vs. *dentist*, or *oenologist* vs. *vintner*.

Another source of doublet morphs in English is **Latin and French**.
Middle English borrowed the **Norman French** form of a morpheme,
and then later borrowed the **ancestor** of the **same morpheme** from **Latin**.
(Recall that Latin is the **direct ancestor** of French!)

This is the source of *-ceive*, the allomorph of *cap-* in *receive*, *conceive*, etc.:
the Latin morph *-cep-* **became *-ceive* in French** via language evolution,
and **both** versions of the morph ended up borrowed into English.
French/Latin doublets **very often** function as allomorphs within English.

(Note that the textbook **usually does not show** French allomorphs of the morphemes it lists!
It's still a **very good idea** to know the common French variants.)

A very common pattern in English is:
the **last** morpheme in a word often appears in its **French** form,
while **other** morphemes use the Latin form.
This is what yields alternations like *receive~reception*:
when *cap-* is the **last** morpheme in the word, it appears as French *-ceive*;
when you add a suffix **after** it, it becomes the Latin *-cep-*.

Other examples:

French *-fy*	~	Latin *fac-~-fec-~-fic-*	'do, make':
French *-le*	~	Latin *-ul-*	'little'
French *-tain*	~	Latin *ten-~-tin*	'hold'

glori-fy	*glori-fic-ation*
ang-le	*ang-ul-ar*
re-tain	*re-ten-tion*

On the other hand, there are many French/Latin doublets that **don't** participate
in any **synchronic** alternations in English, and **aren't very similar**—
E.g., the Latin root *cad-* 'fall' appears in *cad-ence*, *ac-cid-ent*, and other words.
French *chance* is **also derived from** this root—in fact, it's a doublet for *cadence*.
But we wouldn't call *cha-* in *chance* an **allomorph** of *cad-*;
it doesn't have the sort of **systematic structural relationships** that make
allomorphy a useful concept.