Handout 4:
Morphology

September 29 - October 13, 2016

Morphology is the study of sequences of sounds that are associated with meaning, and how these sequences of sounds combine and interact with each other. In exploring morphology, we will essentially be studying the internal structure of words.

- In particular, morphologists aim to characterize the system that underlies a native speaker's knowledge of their language's morphological system.

Our plan over the next two weeks:

- What's a morpheme?
- Syntactic/morphological categories
- Lexical entries
- Word structure and ambiguity ("unopenable")
- Morphological analysis
- Allomorphy

1 What is a morpheme?

A little puzzle: How do you say "my" in Cree?

- tʃiːmaːn = "canoe"
- nitʃiːmaːn = "my canoe"
- soːniya = "money"
- nisoːniya = "my money"
- wiyaːʃ = "meat"

→ What piece above means "my" in Cree?

→ How do you say "my meat" in Cree?

A **morpheme** is the smallest phonological unit of language that can be associated with a meaning or grammatical function.

- How many morphemes are there in the following English words?
  - cat
  - cats
  - lens
  - dancing
  - celebration
  - unbearable
  - untie
  - under
  - pirate

A little Turkish:

- el = hand
- elim = my hand
- elimde = in my hand
- eller = hands
- ellerimde = in my hands
- elleremiz = our hands
- ellereniz = your hands
- ellerenize = to your hands

→ Identify all the morphemes you can.

"House" in Turkish is *ev*. How would you say "my house", "in your houses", "to our house"?

More about morphemes and words

- A **word** is a morpheme or series of morphemes that can stand alone as a well-formed and complete unit of language.
- Simple words vs. complex words
  - Simple = Words that consist of a single morpheme (e.g., *dog, man, party*)
  - Complex = Words that consist of two or more morphemes (e.g., *dog-s, parti-er, anti-dis-establish-ment-ary-an-ism*)
- Free vs. bound morphemes
  - Free morpheme = a morpheme that can stand alone as an entire word.
    - ◇ Such a morpheme may be able to have other morphemes attached to it, but these are not required.
    - ◇ E.g., *the, cat, run, pretty*
  - Bound morpheme = a morpheme that cannot stand on its own as a word, but rather must be attached to another morpheme.
    - ◇ Never found in isolation.
    - ◇ E.g., *re-, un-, -est, -er*, also *-ceive, -venge*
- How to identify a morpheme:
  - Does it contribute part of the meaning of a word?
  - Is it reusable in other words?

  → Practice with: *investment, rewrite, impossibility*

- Roots and affixes
  - Root = contributes the core meaning of a word, often free (can stand alone as a word), one per word (usually)
    - ◇ What's the root of *investment, rewrite, impossibility*?
    - ◇ Roots can be bound: *fiss*-ure, *fiss*-ion, *cens*-ure, *cens*-or
  - Affix = bound morpheme that attaches to a root, does not contribute the core meaning of a word, can be many per word
    - ◇ Prefix = precedes the root (e.g., *re*-write)
    - ◇ Suffix = follows the root (e.g., writ-*ing*)
    - ◇ Infix = inside the root (e.g., abso-*fucking*-lutely)
    - ◇ Circumfix = surrounds the root (e.g., *a*-hunt-*ing*)

Why should we look at subparts of words, rather than saying that speakers of a language memorize whole words?

- Morphology is *predictable*.
  - *-er* attaches to verbs and gives you a noun that means 'one who verbs'.
    - ◇ writer
    - ◇ drummer
    - ◇ drinker
    - ◇ ...
- Morphology is *productive*.
  - I can give you a made-up word, and you can put morphology on it in a meaningful way. Let's say, for example, that I tell you *wug* is a verb.
    - ◇ wuggable = able to be wugged
    - ◇ wugged = wug in the past tense
    - ◇ wugging = wug in the progressive
    - ◇ unwug = undo the wugging
    - ◇ rewug = wug again

There is huge crosslinguistic variation in morphology.

- Some languages have no (or very little) morphology (e.g., Mandarin, Vietnamese), i.e., there are no affixes.
- Some languages make entire sentences using morphology (e.g., Eskimo-Aleut, Neo-Aramaic).
  - *tm-kasw-ox-waa-laa* = 'We wrote it a long time ago.' (Senaya, Neo-Aramaic)
  - *tusaa-tsia-runna-nngit-tu-alu-u-junga* = 'I can't hear very well.' (Inuit, Eskimo-Aleut)
- Languages differ as to which functions/concepts are expressed by separate morphemes.
  - In English, we typically do not encode gender morphologically (e.g., *child* can be feminine or masculine).
  - In Spanish, gender is usually encoded morphologically (e.g., *chic-o* and *chic-a* both mean "child" but differ in gender).
- Languages differ as to what's an affix and what's a free morpheme.
  - In English, future tense is expressed with a free morpheme, *will*.
  - In Spanish, future tense is expressed with a bound morpheme, *-á* (with different variants for gender and person), e.g., *hablar-á* = 'he/she will talk'

2 Syntactic/Morphological Categories ("parts of speech")

To do morphology, we'll need to be able to identify a word's **syntactic category** (the linguistic term for *part of speech*).

2.1 The wrong way to define syntactic categories: semantically

What your middle school English teacher taught you:

- Noun = person, place, or thing
  - What about... *disappearance, year, honesty, thought, swimming...*
- Verb = action or state of being
  - What about... *is, became, seems, took (a long time)...*
- Adjective = describes a noun
  - What about... *former, mere, fake...*
- Adverb = describes a verb, adverb, or adjective
  - What about... *really, so, yet...*
- Preposition = displays the relationship between words in a sentence
  - What about... *of*?

Further, lots of words can appear as multiple categories:

(1) a. The man will man the manning station.
b. I used Google to Google something and got a Google search result.
c. Zeitoun's claim ought to claim that he is innocent.

Further still, we don't need to have any clue about meaning to determine syntactic category.

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogoves,
And the mome raths outgrabe.
(Lewis Carroll, *Jabberwocky*)

Conclusion: Semantics is not the way to go about identifying syntactic categories.

2.2 The right way to define syntactic categories: substitution classes

Let's see how far we can get by defining syntactic categories based on their environment – in a sentence or even in a word.

- Substitution Class = A group of words/morphemes that have the same distribution
- Distribution = A word's position in a phrase/sentence (syntactic distribution) or a morpheme's position in a word (morphological distribution)

Let's take a random position and see what we get: the _____ exists

⟶ The words that can sit in this position all belong to the same substitution class, which we traditionally call "noun".

(2) a. The girl exists.
b. The lamp exists.
c. The cat exists.
d. The boxcar exists.

⟶ The words that can't sit in this position do not belong to the class "noun".

(3) a. *The happy exists.
b. *The ridiculous exists.
c. *The sleepy exists.
d. *The fluffy exists.

(4) a. *The in exists.
b. *The out exists.
c. *The from exists.
d. *The within exists.

(5) a. *The sit exists.
b. *The sat exists.
c. *The eat exists.
d. *The ate exists.

(6) a. *The quickly exists.
b. *The exuberantly exists.
c. *The often exists.
d. *The always exists.

It turns out, there are distinguishing environments for each traditional "part of speech":

(7) a. VERB: to _____ there (arrive, dance, speak, retire...)
b. NOUN: the _____ exists (cat, chocolate, sand, book...)
c. DETERMINER: in _____ thing (the, every, some, few, a, no...)
d. ADJECTIVE: the _____ thing (blue, insane, silly, lovely, reversed...)
e. ADVERB: have _____ eaten (slowly, hungrily, poorly...)
f. PREPOSITION: dance _____ it (on, around, through, at, with...)

The labels noun, verb, etc. are just names for the subsets of vocabulary items partitioned into substitution classes.

Note that substitution-classes are not fool-proof, and that we should be talking much more generally than I made it look in (7). We won't worry about this for now. This section is aimed just at helping you identify syntactic categories/parts of speech.

2.3 Morphological distribution

Not only can we provide a distinguishing environment for a substitution class by using the surrounding words, but also by using the surrounding morphemes.

(8) VERB
- a. _____-able
- b. _____-s (3rd singular present)
- c. _____-ing
- d. un-_____ (undo X)

(9) ADJECTIVE
- a. _____-ness
- b. _____-ly
- c. un-_____ (not X)

(10) NOUN
- a. _____-s (plural)
- b. _____-ship

→ What we're essentially doing in (8)-(10) is showing that words of different syntactic categories go with different affixes. This is another step in figuring out what a native speaker knows about their language's morphemes.

- *-able* wants to combine with a verb = *-able* **selects** for a verb
- **Selection** = When X requires the presence of Y, then we say that X selects Y.

3 Properties of morphemes

What do you need to have memorized about the following morphemes, to use them like a native English speaker would?

- *cat*

- *-ly* (as in happily, quickly, easily)

We know from our discussion above that morphemes have various properties:

- Phonological component
- Semantic component
- If it's a root...
  - Free or bound
  - Category
- If it's an affix...
  - Position (prefix, suffix, infix, circumfix)
  - Category it selects (what category the affix attaches to)
  - Category it derives (what category the resulting word is)
  - Inflectional or derivational (we'll come back to this later)

All of these properties of a morpheme are encoded in a **lexical entry**, which you can think of as an entry in your mental dictionary.

→ The lexical entry tells a native speaker all they need to know to use a morpheme.

What's the lexical entry for...

- *-s* (as in opens, dances, writes)

The function of an affix: Derivational vs. inflectional

- A **derivational affix** is one that significantly changes the meaning of the morpheme it attaches to, OR changes the syntactic category of the morpheme it attaches to.
- An **inflectional affix** is one that does not significantly change the meaning of the morpheme it attaches to, AND does not change the syntactic category of the morpheme it attaches to.
  - Rather, an inflectional affix reflects other properties of the sentence, e.g., the gender of the subject, the tense of the clause, etc.
- Practice: *-ment* (as in abandonment, government) *un-* (as in undo, unopen), *-ed* (as in walked, asked, cooked)
- What do you notice about the relative position of inflectional and derivational affixes in words that have both, like *Americanizations* and *beautified*?

Some pitfalls to avoid

- Sometimes words that look like they are multi-morphemic are not.
  - Option = opt+ion
  - Potion ≠ pot+ion
  - Burial = bury+al
  - Canal ≠ can+al
- Remember that English spelling is unreliable, so don't worry about minor changes in spelling.
  - party + er = partier
  - fragile + ity = fragility
- Remember that adjacent sounds can influence/change each other, so don't worry about minor changes in sound.
  - pollute + ion = pollution
  - content + ious = contentious

4 Word structure

Words have structure, because affixes attach one at a time.

To find the structure, (i) separate the morphemes, (ii) find the root(s) and label with category, (iii) find affixes and label with category/Aff, (iv) combine!

- Dogs (see also: cats, chairs, cups)

- Watched (see also: danced, walked, opened)

- Rewatched (see also: reopen, rewrote)

- Establishments (see also: government, abandonment)

- Unhappily (see also: happily, easily, unhappy, uneasy)

- Decentralization (see also: deactivate, demystify, pivotal, facial, national, americanize, idolize, organization, digitalization)

- Unopenable (see also: unopen, undo, openable, drinkable, unhappy, uneasy)

Two new important terms: **constituent** and **compositionality**

5 Word formation

There are two major word-forming processes:

- Affixation (derivational or inflectional): combining an affix and a stem (which may be a root or a root with one or more affixes already attached to it)

  (11) a. establish + ment = establishment
  b. establish.ment + s = establishments

- Compounding: combining two or more roots
  - Usually ends up with the category of one of the roots
  - May significantly change the meaning of the individual roots

  (12) a. spoon + feed = spoonfeed
  b. dog + house = dog house
  c. wash + cloth = washcloth
  d. out + house = outhouse

  - Compounds have a **head** (root that contributes category and core meaning)
  - How to tell if a combination of roots in English is a compound word (rather than, say, an adjective modifying a noun):
    - ◇ Stress on the leftmost root: compound (one word)
    - ◇ Stress on the rightmost root: not a compound (separate words)

      (13) a. **green**house
      b. green **house**
      c. **black**board
      d. black **board**
      e. **wet** suit
      f. wet **suit**

    - ◇ Idiomatic/noncompositional meaning: compound
    - ◇ Predictable/compositional meaning: not a compound

Here's an English compound that is many-ways ambiguous: *Japanese kitten pocket watch*

- How many different meanings (and corresponding structures) can you come up with?

Other word-formation processes apart from affixation and compounding:

- Root and pattern: The root is a series of consonants, and the "affix" is a series of vowels which intersperse with the root in a fixed way.

  (14) Arabic (root: ktb "write" (V))
  - a. kataba = "he wrote" (pattern: CaCaCa)
  - b. kaataba = "corresponded" (pattern: CaaCaCa)
  - c. kutiba = "it was written" (pattern: CuCiCa)
  - d. kitaab = "book" (pattern: CiCaaC)
  - e. kutub = "books" (pattern: CuCuC)
  - f. kaatib = "writer" or "writing" (pattern: CaaCiC)
  - g. uktub = "write!" (pattern: uCCuC)

  (15) Neo-Aramaic
  - a. ptox = "open!"
  - b. patɨx = "open(s)"
  - c. ptix = "opened"
  - d. ptaxa = "opening"

  → What root is represented in (15)? And what patterns are represented?

- Internal change ("ablaut"): A vowel of the root is changed to derive a new word.

  (16) English past tense (cf. dance → danced)
  - a. sing → sung
  - b. sink → sank
  - c. drive → drove

  (17) English plurals (cf. hand → hands)
  - a. foot → feet
  - b. goose → geese

- Suppletion: Complete replacement of one morpheme with another.

  (18) English past tense (cf. dance → danced)
  - a. go → went
  - b. be → was/were

(19) English comparative (cf. fast → faster)
   a. good → better
   b. bad → worse

(20) Spanish forms of *ir* (root for "to go")
   a. voy/vas/va/vamos/van (present tense)
   b. fui/fuiste/fue/fuimos/fueron (past tense)

- Stress and tone shift: Movement of stress or tone derives a new word.

(21) English verbs becoming nouns
   a. im**plant** → **im**plant
   b. im**port** → **im**port
   c. pre**sent** → **pre**sent
   d. sub**ject** → **sub**ject
   e. con**test** → **con**test

(22) Mono-Bili (Zaire)
   a. dá (spanked) → dà (will spank)
   b. zɨ́ ("ate") → zɨ̀ "will eat"
   c. wó ("killed") → wò ("will kill")

→ What are the roots in (22)? And what is the morphological process that derives a new word?

- Reduplication: Part of the root is copied and added on to the root to derive a new word.
  - Full reduplication: Whole root repeated to derive a new word.

(23) English emphasis through reduplication
   a. V: like → like like
   b. N: party → party party
   c. Adj: hot → hot hot

(24) Turkish adverbial intensification
   a. tʃabuk "quickly" → tʃabuk tʃabuk " very quickly"
   b. iji "well" → iji iji "very well"
   c. gyzel "beautifully" → gyzel gyzel "very beautifully"

– Partial reduplication: Part of the root repeated to derive a new word.

(25) Tagalog
a. takbuh "run" → tatakbuh "will run"
b. lakad "walk" → lalakad "will walk"
c. pili? "choose" → pipili? "will choose"

→ What is reduplicated in (25) and where does it go with relation to the root?

- Conversion: A new word is derived with no overt change to the word.

(26)
a. Mark is a great **father**.
b. Mark will **father** three children.
c. I will write a **report**.
d. I will **report** my findings.

– Is there a way to tell whether the process above involves a verb becoming a noun or a noun becoming a verb?

– What's the best way to model this?

– Another place we may want to appeal to null (silent) morphology:

(27) English present tense
a. I dance-∅.
b. You dance-∅.
c. She/he dance-s.
d. They dance-∅.

(28) English plurals
a. dog (sg) → dog-s
b. child (sg) → child-ren
c. fish (sg) → fish-∅ (pl)
d. moose (sg) → moose-∅ (pl)

6 Allomorphy

Just like phonemes, morphemes can be pronounced differently depending on their context.

- These variations in the pronunciation of morphemes are called allomorphs.
  - **Morphemes** are the underlying forms stored in your brain. (∼ Phonemes)
  - **Allomorphs** are not stored in the brain, but rather can be predicted by morphological rules (which are just like phonological rules). (∼ Allophones)

6.1 Surface allomorphy

It is often impossible to do morphological segmentation without also doing some phonology.

→ This is because a single morpheme may have more than one surface phonological form.

(29) The English plural morpheme

a.	b.	c.
kæt ∼ kæts	dɑg ∼ dɑgz	bʌs ∼ bʌsəz
dɑk ∼ dɑks	hawnd ∼ hawndz	mejz ∼ mejzəz
pɑt ∼ pɑts	blɑb ∼ blɑbz	tʃərtʃ ∼ tʃərtʃəz
kʌf ∼ kʌfs	kowv ∼ kowvz	gəraʒ ∼ gəraʒəz
bæθ ∼ bæθs	paj ∼ pajz	wɪʃ ∼ wɪʃəz

- Write a lexical entry for the three meaning:phonology pairs in the table below:

phonology	a. /s/	b. /z/	c. /əz/
meaning			
category			
position			
selection			

- So are there three different plural morphemes in English? If not, why not?

- What conditions the choice of the phonological variant? And what is the elsewhere form?

What we have seen above is a form of **allomorphy**:

- 1 morpheme (lexical entry) may be related to multiple surface realizations (= **allomorphs**)
- Complementary distribution!
- When the allomorphs can be related to each other by (plausible) phonological rules, then it is an instance of **surface allomorphy**.
- The English plural:
  - Three surface allomorphs: {s, z, əz}.
  - Resembles the behavior of other morphophonological variation in English:

possessive	‘’s’	Jack’/s/	Dan’/z/	Nash’/əz/
3rd singular	‘-s’	attack/s/	ban/z/	rush/əz/
past tense	‘-ed’	walk/t/	live/d/	edit/əd/

Surface allomorphy is phonology.

- Recall that lexical entries only include information *that is not predictable from other parts of a language’s grammar.*
  - The phonological form stored in the lexical entry (the “phonology” component) is called the **Underlying Representation (UR)**, chosen from the elsewhere allomorph.
  - The variation in the surface realization of the plural morpheme is predictable from the **phonology** of English, given the UR of the plural morpheme.
  - Analysis of surface allomorphy = lexical entry + phonology
- Let’s write a final lexical entry for the English plural morpheme and formalize the surface allomorphy of the English plural in two phonological rules.

(30)

phonology	
meaning	
category	
position	
selection	

6.2 True allomorphy

Not all allomorphy is surface allomorphy, i.e., not all allomorphy is predictable based on the phonology of a language. In just this case, allomorphs *do* need to be included in the lexical entry.

Hungarian

i:r-	'write'	i:r-**s**	'you write'
seret-	'love'	seret-**s**	'you love'
fog-	'will'	fog-**s**	'you will'
olvɒʃ-	'ask'	olvɒʃ-**ol**	'you ask'

- Is this surface allomorphy? If not, why not?

- TRUE ALLOMORPHY must be listed as part of the lexical entry, i.e., is not predictable by a phonological rule. True allomorphy is an ordered list of phonologically unpredictable URs.
- Let's write a lexical entry for 2nd singular subject agreement in Hungarian:

(31)

label	
phonology and true allomorphy	
meaning	
category	
position	
selection	

  - The ordering of allomorphs in a lexical entry is important: The list is **disjunctive** and consulted **top to bottom**, with the elsewhere allomorph at the bottom.
  - Note that just because a choice of allomorph is phonologically-conditioned does NOT mean that it is predictable by phonological rule.

True allomorphy in English?

A computer; a spire; a basilisk...
An apple, an obelisk, an eel...

True allomorphy in Tzeltal

k'ab 'hand'	lumal 'land'	k'op 'language'
hk'ab 'my hand'	alumal 'your land'	sk'op 'his language'
akan 'leg'	inam 'wife '	at'el 'work'
kakan 'my leg'	awinam 'your wife'	yat'el 'his work'

label	
phonology and true allomorphy	
meaning	my
category	
position	
selection	

label	
phonology and true allomorphy	
meaning	your
category	
position	
selection	

label	
phonology and true allomorphy	
meaning	his
category	
position	
selection	

Like segmentation into morphemes, our positing of true allomorphy is always considered tentative. Consideration of further data might lead us to reclassify some particular alternation as phonology (surface allomorphy).