McFetridge, Chapter 5: The structure of words

All three of our course themes come out in this chapter:

— change

— structure

— representation

We can start by asking "What is a word?"

This seems obvious at first: A word is the thing between spaces on a page.

The problem with this is two-fold:

i) writing on a page is not speech and speech tends to be continuous for varying lengths of time,

ii) the predecessors of English, i.e., Greek and Latin, did not have spaces. The notion of a space between written constituents is a recent thing.

But we can surely do better! Try this:

A word is a minimal free form.

This definition entails the following:

i) a word is a linguistic unit

ii) a word can be used by itself (it's free). Other linguistic units like the "-s" ending, e.g., mutants, cannot stand on their own

iii) a word is the smallest unit (it's minimal) that can be used by itself, as opposed to, say, **phrases**, which are built up from words.

To put this another way, words are intermediate between phrases and **morphemes** (which are used to create words).

Of course it's possible to come up with counterexamples, e.g.,

"Did you say clear or unclear?"

"Un."

But we won't let these little examples of truncated speech detain us.

"Trunc - what?"

Consider the word "word". It's from Old English. Early meanings were different from ours above: "word" meant speech as in

"by word of mouth" = orally

"in the beginning was the word" = the deity speaks the world into being

"to have words with someone" = to argue with

Other languages don't have a word for "word".

How are words represented?

Well, the Greek & Roman orthographies did not delineate words.

— the space is a recent invention

— the Thai language still does not use spaces

Note the ambivalence that we still have with "**compound words**".

— these are words made up of linguistic units that might also be words in their own right

— spelling isn't necessarily a guide, e.g., "set screw", "set-screw", "setscrew"

— consider a commonplace word like "wastebasket"

One diagnostic for compound status is the **stress pattern**.

— English is a so-called "stress-timed" language

In multisyllabic words this means that primary stress is first and secondary stress is on a later syllable.

— for example, in the word "greenhouse" referring to the kind of structure where plants are grown under controlled conditions, the stress is on "green"

— but if I were referring to a house that is the colour green, both the linguistic units "green" and "house" would receive equal stress.

All of this is clearer if we take a word we know to be consistently treated as a single unit: "hot dog". If I were to say each element with equal stress it would refer to an actual dog whose temperature is elevated … rather than the revolting snack "food" that has helped obesity to become public enemy #1.

Other examples: blackbird = name of a bird vs. black bird = a bird that is black; police car; mouse pad…

All of the preceding also touches on the issue of how many words there are in a language.

— are "compute", "computed", and "computing" three separate words or just forms of a single word?

There are (at least) two ways of thinking about this:

(i) a word is the thing that we speak or read

(ii) a word is the thing in our mind; the abstract thing that we find in a dictionary

— we'll use the term word to describe what we speak and read

— we'll use the term lexeme to describe the abstract entity in our minds

A word is an inflected lexeme … which raises this issue: what's inflection?

Inflection is a process that adds grammatical information to a word. Such things as plural, tense, comparative and superlative markers and other things.

> Not all word classes can be inflected, but we won't worry about that annoyance for now.

Adjectives inflect. For example,

positive	red
comparative	redder
superlative	reddest

So do nouns,

singular	rat
plural	rats

So how do we create words? We need rules that build structure.

Word ⇒ Lex + infl

rats $\quad Word_{N\,[+pl]} \Rightarrow Lex_N + s$

rat $\quad Word_{N\,[+sg]} \Rightarrow Lex_N + \varnothing$

So far, so good. Let's give a rule that will generate the "-ing" form of "compute" (computing). This form is called the present participle.

$$Word_{V[pres\ part]} \Rightarrow Lex_V + ing$$

Problem: "computing" does not have the "e" that's in "compute". So if "compute" is the lexeme, we need another rule e + i → i.

The symbol '+' indicates a boundary

So, in a nutshell, inflection creates words, but not new lexemes.

Another process, **derivation**, creates new lexemes.

Derivation is highly productive and English employs dozens of affixes, i.e., prefixes and suffixes, in creating new lexemes.

Unlike inflection, derivation can create lexemes of a different grammatical category than the form on which they are based, e.g.,

> adding the agentive suffix '-er' to the verb "teach" yields the noun "teacher"

Also unlike inflection, derivation can change meaning, e.g.,

> adding the negative prefix 'un-' to the adjective "happy" yields "unhappy"; a rather different meaning!

This kind of derivational morphology is constrained by rules that limit which affixes can apply to any given lexical category, e.g., *uncar.

How do inflection and derivation interact?

Once a lexeme is inflected, we can't inflect it any further, e.g.,

— computes ✔

— computed ✔

— computesed ✕

We can go on applying derivation, however, e.g.,

— govern ✔

— government ✔

— governmental ✔

But once we inflect a derived form, we're blocked, e.g.,

— governments ✔

— governmentsal✕

Thus we say that "inflection appears outside of derivation".

There is much more that could be said about how words enter the lexicon. For example, stress shift can create new lexemes:

> **re**cord vs re**cord** Depending which syllable you stress you'll have either a noun (the first form) or a verb (the second form).

Another process is conversion in which a lexeme of a particular category becomes "repurposed" with an addition category, e.g.,

> "ink" is a noun but can be used as a verb in a sentence like "He inked the contract yesterday."

Finally, consider backformation. The noun "wrinkle" is actually derived from the past participle form "wrinkled". The verb "refrigerate" is derived from the noun "refrigerator". These words are formed "backwards", as it were.

We must leave those interesting details and consider an important piece of terminology: the **morpheme**.

> — morphemes are used to build lexemes and words, i.e., they're the little bits like "un-" and "-able" that are added to a word like "read" to yield "unreadable".

— A morpheme is the smallest meaningful unit in a language.

This statement gives rise to a closer look at "meaning" …

There are at least 3 kinds of meaning:

(i) LEXICAL MEANING - lexemes such as we find in the dictionary carry lexical meaning; so do morphemes such as "un-" and "-able" that must attach to lexemes.

(ii) GRAMMATICAL MEANING - inflectional meanings that are required by grammar, e.g., the plural "-s" or the comparative "-er", also provide a sort of meaning.

(iii) CATEGORICAL MEANING - endings like "-ly" that don't carry lexical meaning but do change category. Thus "-ly" creates adverbs, e.g., "quickly".

The root is the morpheme that carries the core or central meaning of a lexeme, e.g., the "read" part of "unreadable". The other morphemes, one a prefix and the other a suffix, attach to the root.

So what about this "-able"?

— we can't attach it just anywhere *carable, *beautifulable

— it's not the same thing as the linguistic string "able" in "table"

— it looks like we can attach it to verbs to form adjectives

$$Lex_A \Rightarrow Lex_V + able$$

This is a **derivational rule**.

Now, consider this table (V.5 from your text)…

Derived word	Root	Category
unclean	clean	adjective
uncover	cover	verb
unfold	fold	verb
unfriendly	friendly	adjective
unholy	holy	adjective
unbind	bind	verb

So what can the prefix "un-" be added to?

If we add it to adjectives we create meanings like "not clean", "not friendly", and "not holy" - a negative meaning.

But when we add it to verbs we create meanings like "remove the cover", "remove the fold", and "remove the binding" - a reversative meaning.

Now, an essential component of a morpheme is its meaning.

— if two identical sequences have different meanings, then they are not instances of the same morpheme. Rather they are two different morphemes.

Bottom line: there are two different "un-"s in English. They give rise to the following derivational rules:

— Negative: $Lex_A \Rightarrow$ un + Lex_A

— Reversative: $Lex_V \Rightarrow$ un + Lex_V

So to create "unreadable" we apply the following rules:

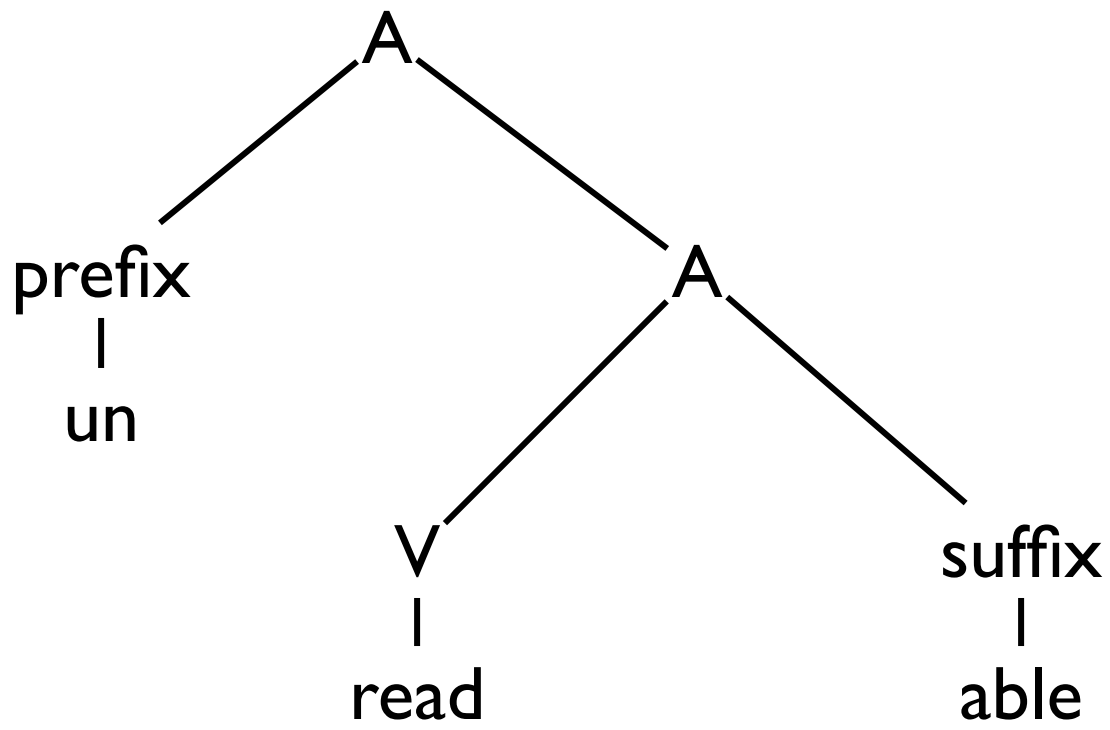
1. $Lex_A \Rightarrow Lex_V$ + able (this gives us an adjective)

2. $Lex_A \Rightarrow$ un + Lex_A (this gives us the final form "unreadable")

Can those two rules be applied in either order?

NO

The negative "un-" can only be attached to adjectives. "read" is a verb so negative "un-" cannot be attached to it directly. Rather we have to create the adjective "readable" first and THEN attach the prefix.

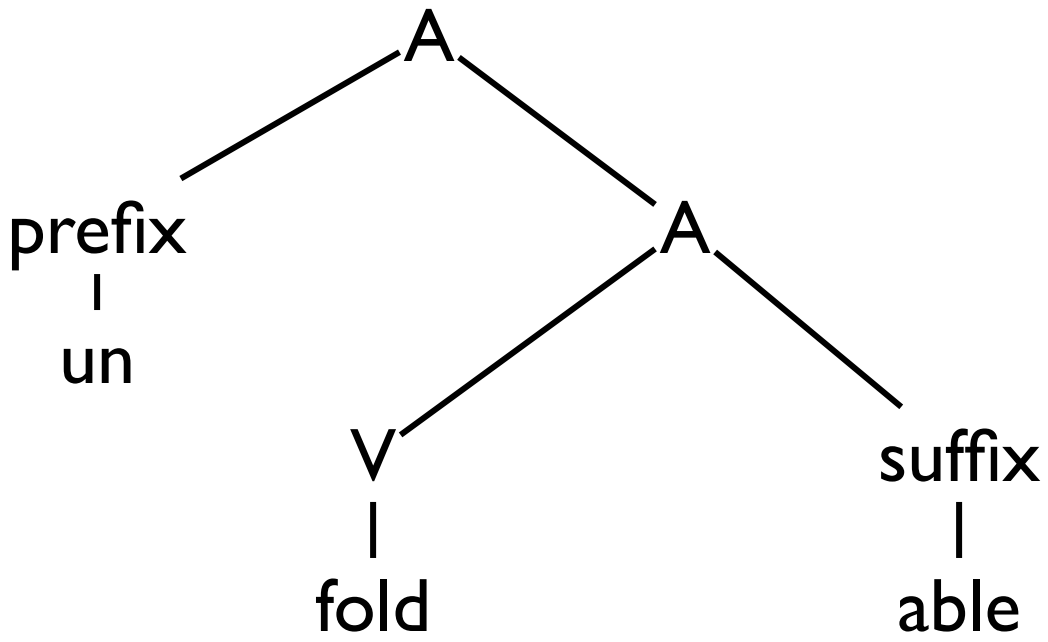


Consider now "unfoldable". This word is ambiguous. It can mean:

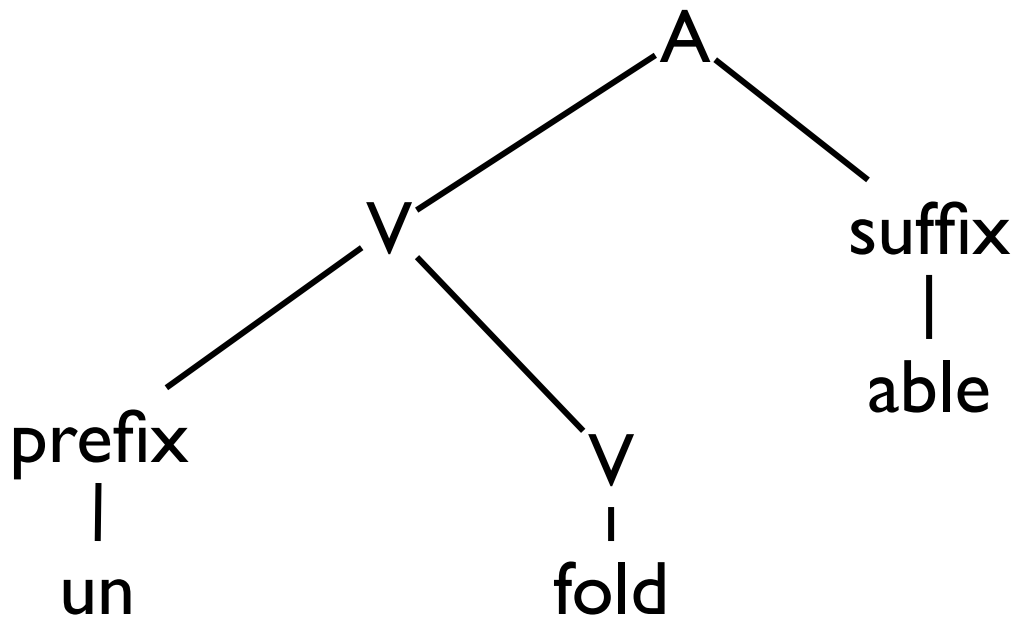
1) not capable of being folded, or

2) capable of being unfolded

This word has the same structure as "unreadable" on the face of it. But because it is capable of two distinct interpretations, we can draw its trees in two distinct ways. The first meaning requires the "negative -un":



The second meaning requires the "reversative un-":



So, we've learned to distinguish between "lexeme" and "word" and also that words have structure. Structure is built through morphological rules that apply in a particular order, thus demonstrating that language is constrained by rules. This is useful to know because otherwise language would be a random affair based on the whims of individual speakers.

McFetridge, Ch. 6: Integration

We've seen how languages change. New words come into being all the time, just as others die out. But new words aren't always just made up from nothing. Frequently they're borrowed from other languages.

We need to explore now how morphemes change as they are added to one another. This allows us to analyze structure and the nature of borrowing in even more depth.

A good place to begin is with **plural** forms.

If you stop to think about it, the plural of some pretty common words is NOT formed with the usual morpheme "-s" at the end .

In this regard, look at Table VI.I from your text … to which I've added a couple of rows…

Latin	Greek	Latin or Greek Plural	English Plural
	phenomenon	phenomena	
	criterion	criteria	
formula		formulae	formulas
medium		media	mediums
alumnus		alumni	
pendulum		pendula	pendulums
vacuum		vacua	vacuums
	thesis	theses	
	index	indices	indexes
datum		data	
stadium	(stadion)	stadia	stadiums

One thing to note from that table is that the foreign plurals are widely used in CE, but a few words have become naturalized and do not apply the foreign pluralization rules.

This means that we'll have to modify our word-building rules to indicate the language that rules come from.

> But is this just the case with inflection (of which pluralization is an example) or does it apply to derivation as well? We can test this by adding the derivational suffix "-ion" to various verbs to form nouns.

Verb	Noun
opine	opinion
complete	completion
produce	production
prevent	prevention

Those result suggest a rule like this:

$$Lex_N \Rightarrow Lex_V + ion$$

But now look at the following table:

Verb	Noun
open	*openion
break	*breakion
fall	*fallion

What gives? It turns out that all of the words that can successfully take the "-ion" suffix are of Latin origin. The words that fail when we attempt to add that suffix are native English words. So we have to refine our rule: $[Lex_N \Rightarrow Lex_V + ion]_{Latin}$

The point is that in some cases English has borrowed not only words, but also the rules that created those words.

Hybrids

We all know that English has borrowed from several languages and is creating new words nonstop. As a result, some words may end up being put together from parts of different languages.

For example, in the word "hypertension" the "hyper-" part is Greek and the "tension" part is Latin. Same thing with "dysfunctional".

> — the reasons for hybrids can be rather interesting. Consider the word "amoral". This came about in the 18th century when philosophers were seeking a word that meant neither "moral" nor "immoral". Back to the Greeks!

Affixes can become naturalized once they've been used in a language long enough. The suffix "-able" is Latin, but it can be added now to English words like "work" and "laugh".

On this final slide for today, I refer back to slide 8 and the inflection of nouns for plural.

— It is not the case that one adds "-s" to the singular to create the plural. In fact, English parallels Latin in the way it forms plurals.

— a Latin plural is formed by adding "-a" to the lexeme and the singular is formed by adding "-um"

— an English plural is formed by adding "-s" to the lexeme and the singular is formed by adding "∅".

The rules have different content, but the processes are the same — as you might expect from two Indo-European languages. The key point is that pluralization is a process that operates on lexemes, and not words.