

Chapter VI: Integration

Introduction

- Previously we have been introduced to 3 concepts that are central to our study:
 1. Change.
 2. Structure.
 3. Borrowing
- In this chapter we are going to integrate these 3 concepts to provide the tools we need for the analysis of the English lexicon.

Structure & Borrowing

- In search of the proper plural form of a word:
- Is *criteria* a singular or a plural?
- What is the correct plural of *medium*: *media* or *mediums*?
- Why is *formulas* the plural of *formula*, except in logic and mathematics where it is *formulae*?
- What is the plural of *alumnus*?
- Let us look at the origins of these words.

Foreign & English Plurals

Latin	Greek	Foreign Pl	English
	phenomenon	phenomena	
	criterion	criteria	
formula		formulae	formulas
medium		media	mediums
alumnus		alumni	
pendulum		pendula	pendulums
vacuum		vacua	vacuums
	thesis	theses	
	index	indices	indexes

Foreign & English Plurals

- Two things are apparent:
 1. The true plural form is that from the original language.
 2. Some words have become naturalized so that they are now receiving the native English plural.
- Naturalization
 1. In some cases, the naturalization form will replace the foreign form. Examples: vacuums & formulas- outside of the specialized discourse of logic and mathematics.
 2. In other cases, both the foreign and naturalized forms continue to persist although with different meanings. Example: *media* & *mediums*.
 3. Finally, some words resist naturalization. Example: *thesis*

Word Building Revision

- Previously, one module for inflection.
- There are however, different ways to inflect a noun for plural depending on whether it is Greek, Latin, English or naturalized.
- Need separate modules for inflecting words depending on their origins.
- If we are revising the inflectional rules, how about the derivational rules?
- Consider the rule:
 $Lex_N \Rightarrow Lex_V + ion$
- This rule will create nouns from verbs by adding *-ion*.

Table VI.2

<u>Verb</u>	<u>Noun</u>
opine	opinion
complete	completion
produce	production
prevent	prevention
open	*openion
break	*breakion
fall	*fallion

Table VI.2

<u>Verb</u>	<u>Noun</u>
opine	opinion
complete	completion
produce	production
prevent	prevention
open	*openion
break	*breakion
fall	*fallion

Derivational Analysis

- Analysis: Whenever it is possible to apply the *-ion* rule, the lexeme that takes the suffix is Latinate.
- The lexemes that can not take the suffix are of English heritage.
- There are many reasons why a lexeme cannot take a this suffix, but the one constant is that if a lexeme does take this suffix it must be Latinate.
- Therefore, we need to create blocks of rules that apply only to lexemes of particular languages.
- The rules must be indexed so that they will be applied only when appropriate.

$$[\text{Lex}_N \Rightarrow \text{Lex}_V + \textit{ion}]_{\text{Latin}}$$

Hybrids

- Derivationally pure: words should be derivationally Latin, Greek or English but not a mixture.
- Note: For most derivational processes this is true.
- However, occasionally hybrids are created.
- A hybrid is a lexeme that contains elements from more than one language.
- If we want to create hybrids we will need to include a mechanism that directs lexemes to other derivational modules.

Hybrids con't

- Hybrids are created for many of the reasons we described in the discussion of borrowings.
 1. A new concept is created and no word exists in any language.
 2. One of the elements has been naturalized. The longer an element has been in a language, the more it resembles native words. They look like English because they have been in the language much longer than others and have been in much more common use. As a consequence, their rough edges have been smoothed off and they have come to resemble English lexemes.

Plural vs. Singular

- Propose that there are 2 rules of nouns in our English inflection module.
 1. One is the rule that creates plural nouns from lexemes by adding *-s*:
- 2. The second rule will create a singular noun from a lexeme by adding nothing at all:

$$[\text{Word}_{\text{N}[\text{Pl}]} \Rightarrow \text{Lex}_{\text{N}} + s]_{\text{English}}$$

$$[\text{Word}_{\text{N}[\text{Sg}]} \Rightarrow \text{Lex}_{\text{N}} + \emptyset]_{\text{English}}$$

Latin Plurals

<u>Singular</u>	<u>Plural</u>
medium	media
pendulum	pendula
vacuum	vacua

Latin Plurals

- We note that the singulars all end in *-um*.
- All the plurals end in *-a*.
- It appears that the lexemes are: *medi-*, *pendul-*, and *vacu-*.
- To generate the singular and plural words, the Latin inflectional module requires 2 rules:
 1. $[\text{Word}_{\text{N}[\text{Pl}]} \Rightarrow \text{Lex}_{\text{N}} + a]_{\text{Latin}}$
 2. $[\text{Word}_{\text{N}[\text{Sg}]} \Rightarrow \text{Lex}_{\text{N}} + um]_{\text{Latin}}$
- Contemporary English differs from Latin in that it no longer adds anything to the lexeme to form the singular, but both have rules that create singular and plural words from lexemes.

Summarize

- The essential point of this section is that when we investigate word structure, we must be aware of the language of origin.
- Conversely, if we want to discover the parent language of a word, recognizing parts of the word is very helpful.

Change & Structure

- Cleaning up: *computeing vs. computing
- To generate the correct form, we proposed a rule:
 $e + i \rightarrow i$ to delete the 'e'.
- This rule is really just a spelling rule.
- Good reason to keep the idea of this kind of rule in our computer program.
- Consider the negative prefixes: *-in*, *-il*, *-ir*

Negative Prefix

in <u>a</u> ctive	il <u>l</u> egal	ir <u>r</u> eplaceable
in <u>t</u> olerant	il <u>l</u> egible	ir <u>r</u> edeemable
in <u>c</u> apable	il <u>l</u> icit	ir <u>r</u> egular
in <u>f</u> lexible	il <u>l</u> egitimate	ir <u>r</u> elevant

Negative Prefixes con't

- All the words are formed by adding a prefix to a lexeme.
- This prefix in each column negates the adjective to which it has been added.
- We could propose 3 rules:
 - $\text{Lex}_A \Rightarrow in + \text{Lex}_A$
 - $\text{Lex}_A \Rightarrow il + \text{Lex}_A$
 - $\text{Lex}_A \Rightarrow ir + \text{Lex}_A$
- There is a problem. How can we constrain these rules so that they apply only to the right lexemes?
- For example, the *il-* rule never applies to lexemes that begin with *t* like **ilterminable* instead of *interminable*

Negative Prefixes con't

- Is there a pattern?
- *il-* appears only when the lexeme to which it is added begins with *l* and *ir-* appears only when the lexeme begins with *r*.
- Have no way of expressing this.
- Do we want to be able to?
- By proposing 3 different rules, we are claiming that there are 3 different negative morphemes.
- Recall: that an important property of morphemes is the persistence of meaning.

Negative Prefixes con't

- Consequently, this property suggests that when we discover different forms with the same meaning we must consider whether they are synonyms, or different forms of the same morpheme.
- What is important to note in this specific case is that the variants are very similar in form and their differences are completely predictable.
- The *il-* and *ir-* forms are predictable.
- The *in-* form seems to occur in a variety of unpredictable places (before *a, t, c, f, etc.*) that have nothing in common.

Negative Prefixes con't

- What we want to propose is a single morpheme with a single lexeme building rule.
- Since we can not predict where the *in-* variant will occur (widest distribution) we select it as representative of the morpheme.
- We will generate the other forms of the morpheme by using the '→' rules that we have previously suggested are necessary for getting spelling right.
- From now on we will refer to these rules as *phonological rules*. Phonological rules are responsible for adjusting how morphemes are pronounced given the context that morphological rules have created.

Negative Prefixes con't

- The proposed negative prefix rule:

$$\text{Lex}_A \Rightarrow in + \text{Lex}_A$$

In addition, two rules to generate the variants of *in-*:

$$n + l \rightarrow l + l$$

$$n + r \rightarrow r + r$$

- The lexeme building rules apply first.
- After the lexeme (and word) have been built, the phonological rules apply to readjust how components of the morphemes are pronounced given the new context they are found in.

Deriving Negative Adjectives

active	legal	regular	
<i>in</i> + active	<i>in</i> + legal	<i>in</i> + regular	$\text{Lex}_A \Rightarrow \textit{in} + \text{Lex}_A$
“	<i>il</i> + legal	“	$n + l \rightarrow l + l$
“	“	<i>ir</i> + regular	$n + r \rightarrow r + r$
inactive	illegal	irregular	Remove ‘+’

Assimilation Rule

- Assimilation Rules: one sound becomes similar or identical to a neighboring sound.
- In the rules that we have just proposed:
n assimilates to *l* and *r*
- In these rules, *n* assimilates completely and so the sound becomes identical to *l* and *r*.
- There are other cases where a sound assimilates only partially.
- Let's consider the Negative Prefix *im-*

Negative Prefix *im-*

moral	<i>im</i> <u>m</u> oral
measurable	<i>im</i> <u>m</u> easurable
mobile	<i>im</i> <u>m</u> obile
mature	<i>im</i> <u>m</u> ature
possible	<i>im</i> <u>p</u> ossible
practical	<i>im</i> <u>p</u> ractical
precise	<i>im</i> <u>p</u> recise
potent	<i>im</i> <u>p</u> otent

The Negative Prefix *im-*

- Note the prefix *im-*
- This prefix is added to adjectives to create an adjective whose meaning is the negative or the original.
- The similarity in form and meaning to the *in-* prefix suggests that this new prefix is somehow related to it.
- If we assume that the *immature* is *in + mature*, then it is clear that we require the phonological rule:
$$n + m \rightarrow m + m$$
- An assimilation rule: *n* has completely assimilated to a following *m*.

The Negative Prefix *im-* con't

- But what about *impossible*?
- The word seems to contain the prefix *im-*
- Again, this prefix is added to adjectives to create adjectives whose meaning negates the original, so we assume that it is another form of the prefix *in-*
- The rule would look like this:
$$n + p \rightarrow m + p$$
- Note: both *m* and *p* are labials, while *n* is a dental nasal.
- Thus, this is an example of partial assimilation: *n* has partially assimilated to the following *p*. It is still a nasal, but is now articulated in the same position as *p*.

The Negative Prefix *im-* con't

- Phonological rules are supposed to apply regularly throughout the language.
- Do these rules apply in other places. This would justify the analysis that we have proposed.
- Take a look at the Prefix *con-*

The Prefix *con-*

form	con form	respond	cor respond	mission	com mission
duct	con duct	relate	cor relate	measure	com measureable
test	con test	lateral	col lateral	press	com press
genial	con genial	lapse	col lapse	pact	com pact
strict	con strict	labor	col laborate	patriot	com patriot
sign	con sign			passion	com passion
join	con join				
figure	con figure				
verge	con verge				

The Prefix *con-*

- The prefix *con-* is added to a variety of different grammatical classes.
- There are instances of *col-*, *cor-*, *com-*
- These other instances appear in just the environments that our phonological rules predict.
- Therefore just need one morphological rule:
$$\text{Lex} \Rightarrow \text{con} + \text{Lex}$$
- The phonological rules that were created to account for the various forms of *in-* will look after the rest.
- Our rules are justified. (Provide independent evidence.)

Summary

- The function of morphology is to create lexemes and words.
- By combining morphemes, morphological rules bring together sounds that may create contexts for change.
- The rules that govern this change are phonological rules.
- These rules adjust the pronunciation of morphemes given their new contexts and apply after morphological rules have constructed lexemes and words.

Change and Borrowing

- A phonological rule potentially applies to every word in the language.
- When would it not apply?
- When the sounds which it is sensitive to do not appear in the word.
- Initially, it appears that numerous examples show that the phonological rules that we have been proposing are incorrect.
- Let's look at the prefix *un-*.

The Negative Prefix *un-*

		happy	unhappy
		conditional	unconditional
		deniable	undeniable
legal	illegal	lawful	unlawful
redeemable	irredeemable	readable	unreadable
		ripe	unripe
mature	immature	manly	unmanly
possible	impossible	pleasant	unpleasant

The prefix *un-*

- The prefix *un-* is added to adjectives to create an adjective whose meaning negates that of the original.
- It looks like *un-* might be a variant of *in-*.
- However, we cannot treat it as such.
- First- there is no way to predict the vowel; that is we cannot propose a rule that changes *i* to *u* in the right circumstances.
- Second- it does not behave the same way: the *n* on *un-* does not assimilate to a following *l*, *r*, *m* or *p*.

The prefix *un-*

- We can propose a morphological rule:

$$[\text{Lex}_A \Rightarrow \textit{un} + \text{Lex}_A]_{\text{English}}$$

- The historical origins of the words are important.
- The prefix *in-* is always attached to words of Latin origin.
- The prefix was borrowed into English when the Latin words to which it is attached were borrowed.
- The prefix *un-* is an English prefix and is added to English words or borrowed words that have been naturalized.
- Phonological rules can be categorized by their language of origin.

English Rules

- The assimilation rules that we have been examining are Latin rules and as a consequence do not apply to English morphemes, such as *un-*.
- To provide a further example, look at the Greek Negative Prefix *an-*.
- The forms on the left are the base.
- First few examples suggest that the prefix is *an-* and that it is added to adjectives to create an adjective whose meaning negates the original.

$$\text{Lex}_A \Rightarrow an + \text{Lex}_A$$

Greek Negative Prefix *an-*

aerobic	anaerobic	theist	atheist
matriarchy	anarchy	gnostic	agnostic
isotropic	anisotropic	pathetic	apathetic
oxygen	anoxic	phonic	aphonic
isometric	anisometric	rhythmic	arrhythmic
urine	anuria	static	astatic

Greek Negative Prefix *an-* con't

- The *n* of this prefix does not assimilate.
- There is a variant *a-*.
- The prefix *an-* is before vowels.
- The prefix *a-* before consonants.
- We need a rule that will delete *n* before consonants.
- Using C to represent any consonant:
$$n + C \rightarrow + C$$
- This rule creates further problems.
- Why do some instances of *n* assimilate, some do nothing, and some delete?

Greek Negative Prefix *an-* con't

- The answer lies in the language of origin.
- The assimilating *n* is Latinate.
- The inert *n* is English.
- The *n* that deletes is Greek.
- The deletion rule should be thus:
$$[\text{Lex}_A \Rightarrow an + \text{Lex}_A]_{\text{Greek}}$$
$$[n + C \rightarrow + C]_{\text{Greek}}$$
- To completely account for sensitivity to language origins, a final revision to the structure of our computer program is necessary.

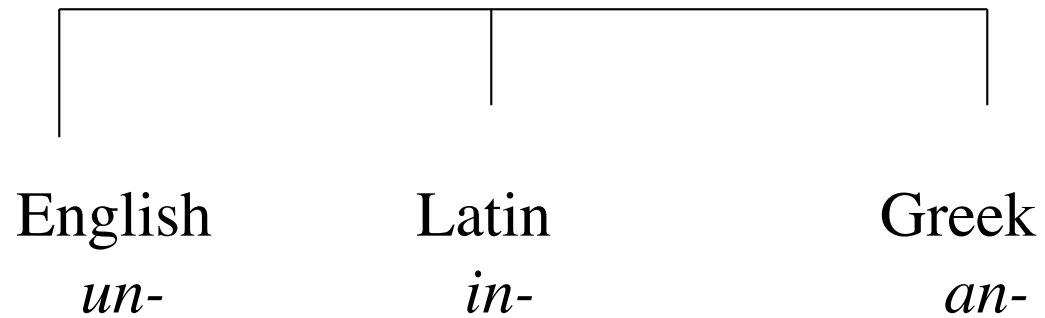
The Prefixes *in-*, *un-*, & *an-*

- Why do Latin, Greek, and English all have a negative prefix that is added to adjectives?
- Why are these prefixes so similar in form?
- These are all sister languages.
- These prefixes are cognates.
- They are descended from a single PIE prefix.

The Prefixes *in-*, *un-*, & *an-* con't

Proto-Indo European

n



Analysis of the Negative Prefixes

- Consider the contrast between *noble* and *ignoble*.
- There is a negative prefix, but what is it?
- Is the prefix *ig-*.
- *Noble* was borrowed from Latin.
- It is subject to Latin morphological rules.
- A possible rule would look like this:

$$[\text{Lex}_A \Rightarrow ig + \text{Lex}_A]_{\text{Latin}}$$

- Problematic: We cannot tell when to use the rule and when to use the rule that we have previously proposed.

$$[\text{Lex}_A \Rightarrow in + \text{Lex}_A]_{\text{Latin}}$$

Latin Negatives II

^a From *gnomon*, “one who knows”.

noble	ignoble
<u>norm</u> ^a	ignorant

Deriving *ignoble*

- Apply our phonological rule:

$$[n + n \rightarrow g + n]_{\text{Latin}}$$

- This isn't right.
- There are examples when $n + n$ does not convert to $g + n$.
- *nocent* (which means “guilty”) *innocent* “not guilty”, not **ignocent*.
- The same root appears in *innocuous* ‘not harmful’ without the predicted g .
- What is the next step?
- Review our assumptions.

Deriving *ignoble* I

noble	noble	
“	in + noble	$[\text{Lex}_A \Rightarrow in + \text{Lex}_A]_{\text{Latin}}$
“	ig + noble	$[n + n \rightarrow g + n]_{\text{Latin}}$
<i>noble</i>	<i>ignoble</i>	remove ‘+’

Re-examining *noble*

- Suppose that our assumption that the prefix is added to *noble* is incorrect.
- The lexeme instead is *gnoble*.
- Suppose instead that the lexeme is instead: *gnoble*.
- We need a rule then that creates *noble* from *gnoble*.
- *g* is deleted when it appears with *n* at the beginning of a word.
- $[\# \text{gn} \rightarrow \# \text{n}]_{\text{Latin}}$
- Note: the $\#$ means that there is no material before, it must be at the beginning of a word.

Considering *ignoble*

- If the lexeme to which *in-* is added is *gnoble*, then the structure of *ignoble* is *in + gnoble*.
- We need a rule that will delete the *n* of *in-*.
- Must be restricted to just those lexemes in which the prefix is added to a lexeme that begins with *gn*.

$[n + gn \rightarrow + gn]_{\text{Latin}}$

- Combining these rules we can derive both *noble* and *ignoble*.
- 2 Types of evidence re-affirm our analysis:
 1. From other Latin borrowings.
 2. From cognates in Greek and Latin

Deriving *ignoble* II

gnoble	gnoble	
“	in + gnoble	$[\text{Lex}_A \Rightarrow in + \text{Lex}_A]_{\text{Latin}}$
noble	“	$[\# gn \rightarrow \# n]_{\text{Latin}}$
“	i + gnoble	$[n + gn \rightarrow + gn]_{\text{Latin}}$
<i>noble</i>	<i>ignoble</i>	Remove ‘+’

Latin Borrowings

- The lexemes *ignoble* and *ignorant* are based on Latin form *gn*¹.
- Used to create words that denoted aspects of knowledge (and, by extension, esteem in *noble*).
- Thus we have: *ignorant* “not knowing”.
- Latin had another form that alternated between *gn* and *gen* and refers to acts of birth and production.
- The *gen* version is apparent in *generate* and *genus* among many other words.
- The *gn*² form is most readily apparent in *pregnant* from *pre* “before” + *gn* “birth”.

Latin Borrowings con't

- Consider the forms *natal* and *cognate*.
- *cognate*- a prefix has been added to *gnate*.
- We know from previous discussions that the prefix is *con-*
con + gnate
- The lexeme *natal* is derived from the form *gn²* (its meaning and similarity of form).
- If the lexeme is *gnatal*, the rule we proposed for *noble* will convert it to the correct *natal*.
- That the rules we need for lexemes created from *gn¹* are also needed for lexemes created from *gn²* is strong evidence that these rules are correct.

Deriving *ignoble* II con't

gnatal	gnate	
“	con + gnate	$[\text{Lex}_A \Rightarrow \text{con} + \text{Lex}_A]_{\text{Latin}}$
natal	“	$[\# \text{gn} \rightarrow \# \text{n}]_{\text{Latin}}$
“	co + gnate	$[\text{n} + \text{gn} \rightarrow + \text{gn}]_{\text{Latin}}$
<i>natal</i>	<i>cognate</i>	Remove ‘+’

Deriving *ignoble* II

- Let's look at cognates of *ignoble* from other languages to verify further our analysis.
- Greek words: *gnostic* and *agnostic*
- *gn^l*: is more noticeable here than it is in the Latin.
- Greek lacks that rule that deletes *g* before *n* at the beginning of words.
- The lack of *n* in *agnostic* is predicted by the deletion we have already proposed for Greek ($an + C \rightarrow +C$).
- The English forms begin with *kn*.
- Recall Grimm's law: voiced stops became voiceless stops in Germanic languages (*g* became *k*).

Deriving *ignoble* II

- English *know* is predicted from Latin *gno* and Greek *gno*.
- However, English *know* is pronounced as [no] and particularly not as [kno].
- English has had in its history an Anglicized version of the rule in Latin that deletes velars before a nasal.

[kn → n]_{English}

- In Latin this rule applies only at the beginning of a word, in English the rule applies everywhere.
- That is why English *know* and *unknowable* do not show the alternation found in *noble* and *ignoble* but instead the root is [no] in both.

Know

Latin	Latin	Greek	Greek	English	English
noble	ignoble	gnostic	agnostic	knowable	unknowable
(norm)	ignorant				

Conclusions

- We need to divide the lexicon into partitions for Latin, Greek, and English words.
- Later we will add a significant partition for French.
- The principles that govern word building are expressed as morphological rules and phonological rules.
- Each partition contains a unique set of morphological and phonological rules.
- The principle of generality requires that each phonological rule applies regularly within the relevant partition.