LECTURES ON POSTSYNTACTIC MORPHOLOGY

Andrew Nevins

Version of July 8, 2015

Contents

1. The Leaves of the Tree are Morphemes

We all know there is regularity in words; for instance, contrast *antidisestablishmentarianism* vs **arianismestablishantidisment*. Speakers of English may not know what either one means (or could mean), but somehow know that both are the same length, both involve the same pieces, and nonetheless that the second one 'breaks the rules' of what can be a word. Where does this regularity come from? A system of rules and constraints, naturally – but what is that system, and how does it interact with the system that puts together words to make phrases?

A traditional view is one called lexicalism, in which, for a sentence such as *I love derivations*, the third word, *derivations*, enters the syntactic tree as a whole word, e.g. as 'derivations'. Where, though, does its regularity in having suffixes in the right place in the right order come from? A lexicon, it is said, and under this view, the leaves of the syntactic tree include preassembled words.

The intuition behind lexicalism is that phonological words correspond to leaves of the tree. But how does this work for items like *I've*? Surely these don't correspond to a single leaf of the tree. (One might try a heavily lexicalist approach, in which *I've* is an auxiliary that specifically requires a null subject. But this wouldn't then tell us why it couldn't undergo subject-aux inversion). The existence of 'contractions' like *I've* already shows us that some part of the phonological packaging into words needs to take place after the syntax is over. The next question would be, is there any reason to think that the phonological packaging of *I've* is of a wholly different nature than that of affixes?

One dead-end would be to say that affixes 'contribute meaning' to their host in a way that cliticized forms do not. A quick inspection of words like *scissor-s* shows this can't work out. We know that the *-s* of scissors is not part of the root, since it can be omitted in compounds and denominal verbs (*to scissor, scissorfight*). But does it contribute any meaning at all? It's surely not pluralizing a singular. It's more like the case of a gender suffix in the Romance languages. The feminine *-a* in Spanish adjectives, as found in *la mesa redonda* 'the round table' does not contribute meaning. It's simply a well-formedness requirement that even inanimate nouns have a gender suffix, but that's hardly a contribution to their 'meaning'. Similarly, we might say that *scissors* belongs to a special 'gender' of nouns, ones that require an *-s* ending even when this ending is not contributing a plural meaning.[1]

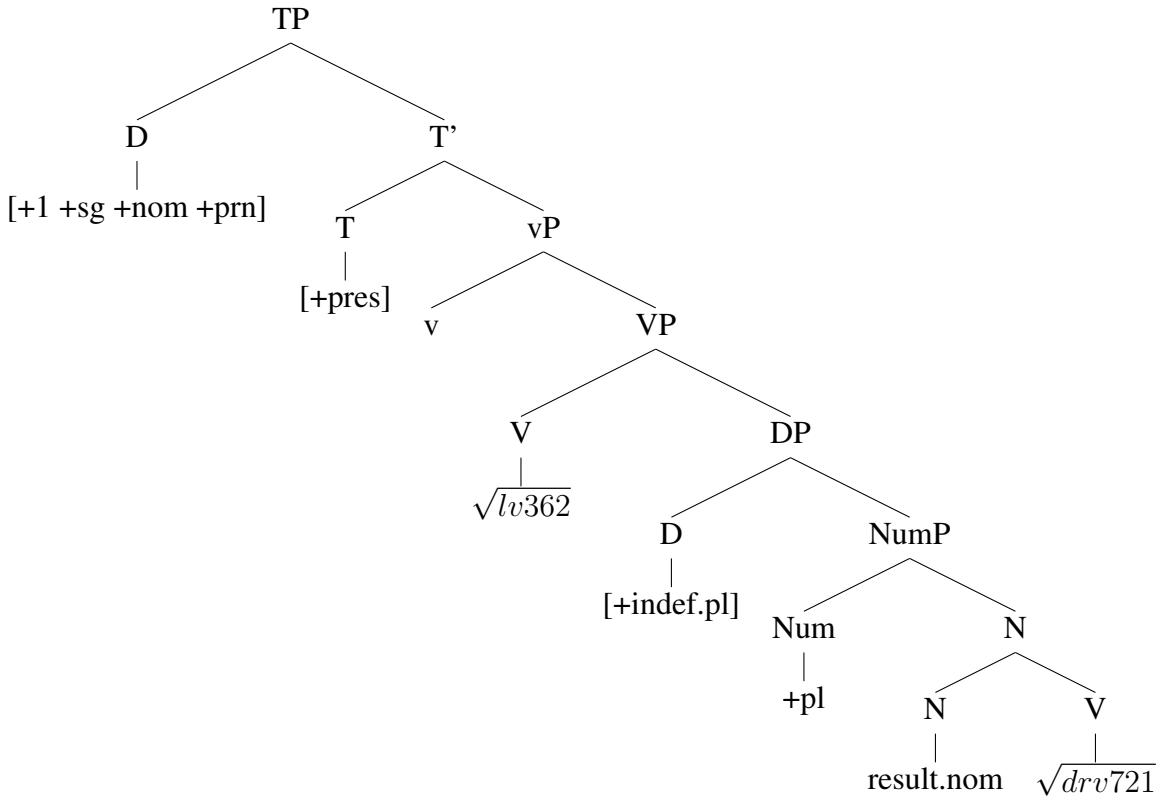
Returning to cliticized forms, such as *I've*, we might think about whether require phonological packaging after the syntax in a way different from affixes? The usual answer is that clitics somehow correspond to their unreduced forms. This may be true etymologically or diachronically, but even the most well-known reducible element, *n't*, does not simply correspond to a reduced form of *not*, and in fact, has many 'affix-like' properties. If *won't* (from *will* + NEG) also doesn't correspond to a single syntactic leaf, shouldn't the packaging of these two pieces into *won't* also take place after the syntax? However, if that is so, we necessarily begin dealing with allomorphy (the choice of *wo* vs *will*) after the syntax is completed.

Let's then return to a clause like *I love derivations*, and how this would look if all allomorphy

[1] It probably will not do to say that the fact that scissors is 'inherently' plural is a necessary fact of its objecthood in the world; in Portuguese, the body-part noun for one's back, *costas*, is inherently plural, and in Hebrew, the noun for bicycle, *ofanayim*, is inherently plural. Like grammatical gender, which often but not always corresponds to biological gender, being a so-called 'pluralia tantum' noun like *scissor-s* is a language-specific choice.

and phonological packaging were done after syntax, a view we'll call 'late phonological insertion'.

(1)



This tree (without head-movement, Agree, or case-assignment shown) represents all of the pieces that will ultimately contribute to the clause *I love derivations*. The leaves of the tree come straight from a list of available atoms in English, which we will call the Formative List. Formatives do not start out with either a fixed sound or a fixed meaning; these are acquired along the way in specific configurations depending on the syntax, similar to the way that phonemes like English /t/ have no fixed sound before they are placed into a phonological environment in which their allophony is determined. Note that these square-root like $\sqrt{lv362}$ elements have no phonology either. They're simply feature-bundles, like the ones in the subject position, that correspond to the formative ultimately expressed as *love*. We'll see how this works below.

It's now time to turn these formatives into phonology. Doing so requires consulting a look-up table, on a list that we will call the Exponent List. The first thing we'll look at is the subject pronoun, which is ultimately realized by the exponent *I*:

(2) a. [+1 +sg +nom +prn]$_D$ $\leftrightarrow$ /aj/
b. [+1 +sg +prn]$_D$ $\leftrightarrow$ /mi/

Note that if the pronoun had not been in Spec,TP, it would not have had nominative case, and wouldn't've been exponed by *I*. Similarly, there might be individuals that simply lack the exponent /aj/ in their List (for example, Cookie Monster), in which case the syntax would have been exactly the same, but the Exponent List would only contain /mi/, and the sentence would be pronounced as *Me love*

derivations.

Plural Number is realized as follows:

(3) a. [+pl]$_{Num}$ ↔ /-z/

The items *v*, T, and the D heading the object noun phrase all happen to be realized as null in this case, but make a contribution to the syntax and the semantics:

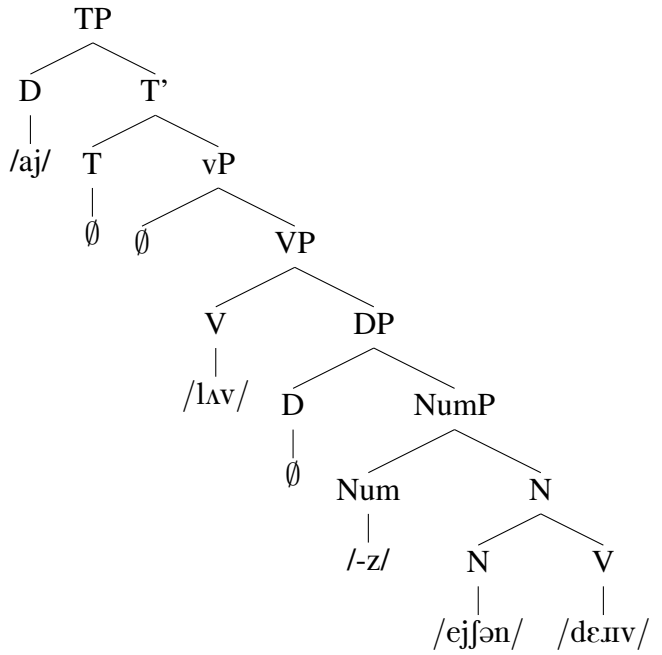
(4) a. [+pres]$_T$ ↔ ∅
b. v ↔ ∅
c. [+indef.pl]$_D$ ↔ ∅

Now what about the remaining formatives? They would be as below:[2]

(5) a. $\sqrt{lv362}$ ↔ /lʌv/
b. $\sqrt{drv721}$ ↔ /dɛɹɪv/
c. [result.nom].N ↔ /ejʃən/

Here's what the resulting tree looks like:

(6)



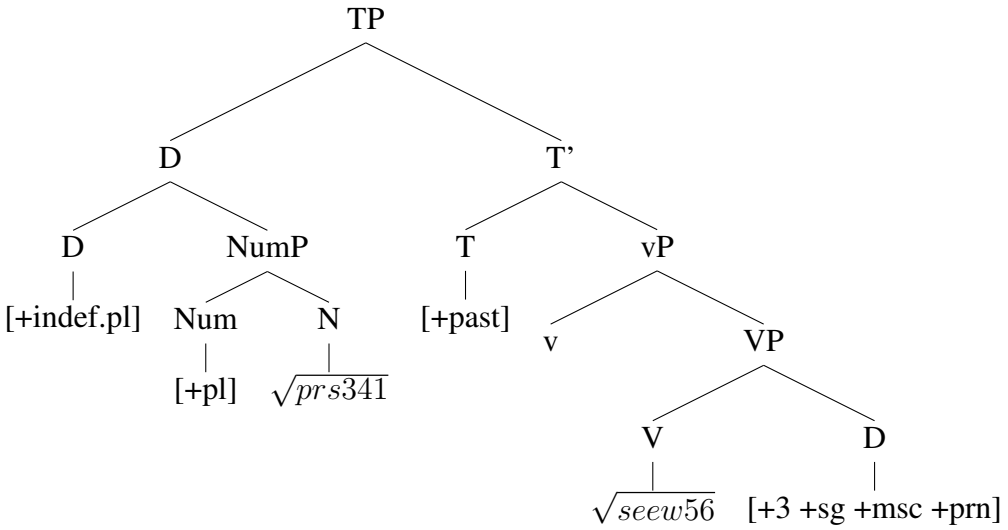
The only question is, how does the hierarchical and left-to-right sequence we've shown, /-z/ /ejʃən/ /dɛɹɪv/, actually enter the phonology as /dɛɹɪvejʃənz/? In other words, there are aspects of what is called the *linearization* of morphemes that must be addressed. Either the syntax must be different than we've shown above, or some other packaging must be taking place. The claim is that the syntax is a little different than we've shown it here, namely that it involves head-movement from V to N and then from [V+N] to Num. This is exactly what we will return to in later sections.

[2]We omit a discussion of the potential internal structure in the suffix /ejʃən/ for expository purposes.

The other question that you might have thought of is, why bother to have *love* dominated by V, if it's a verb? In fact, we will separate the category of V from the content of the root, and slightly revise this tree as well in later sections. In short, the reasoning behind this is that the noun, verb, and adjective *love, lover, loves, lovely* all are 'related'. The easiest way to express this relatedness is to say that they all contain the same root.

Let's reflect – there are three steps in the analysis. The first, the Formative List, provides the Input to the Syntax. Then the Syntax itself happens. When the syntax is done, the Exponent List is consulted. Any analysis done within this framework necessarily involves these three steps. Note, therefore, that word-building hasn't taken place in a 'Generative Lexicon', but rather involves three distinct steps. For this reason, the framework is called *Distributed Morphology*, because as you see, the job of putting words together (much of it overlapping with the job of putting phrases together) is distributed across many different steps and modules of the grammar. Now it's your turn. Write out the Formative List, the Syntax, and the Exponent List for the clause *People saw him*. We'll proceed step-by-step . . .

(7)



Let's start with the easy parts. Again, the items *v*, T, and the object's D all happen to be realized as null in this case, but make a contribution to the syntax and the semantics:

(8) a. [+pres]$_T$ $\leftrightarrow$ $\emptyset$
b. v $\leftrightarrow$ $\emptyset$
c. [+indef.pl]$_D$ $\leftrightarrow$ $\emptyset$

What about the pronoun? Again, it's similar to the case above with the 1st person pronoun. (Note that the features we are using here to represent aspects of pronouns will be revisited in Chapter 2).

(9) a. [+3 +sg +nom +prn]$_D$ $\leftrightarrow$ /hi/
b. [+3 +sg +prn]$_D$ $\leftrightarrow$ /hɪm/

The interesting new twist comes with the realization of the root $\sqrt{prs341}$, which, as you know for English, is *person* in the singular but *people* in the plural. This kind of allomorphy is conditioned

by the Number head, which is adjacent to the root. As such, we will fully pursue the parallelism with *allophony* in phonology, of the type in (10), which is responsible for the affricated form of the obstruent in words such as *tree* [tʃɹiː]:

(10) /t/ → [tʃ] / _ ɹ

Allomorphy rules are a similar kind of context-dependent change in form. In this case, the root $\sqrt{prs341}$ has the following entries:

(11) a. $\sqrt{prs341}$ ↔ /piːp̩l/ / [+pl]$_{Num}$ _
b. $\sqrt{prs341}$ ↔ /pɝsn̩/ / [−pl]$_{Num}$ _

Our Exponent List is almost complete. However, keeping everything as we did from the derivation of *I love derivations* will cause a small problem, namely that [+pl]$_{Num}$ should be /-z/, which would incorrectly yield **Peoples saw him*. However, we know that certain nouns in English have a zero-form for the plural: *sheep, deer, fish, shrimp*. In this case, *people* are no different. The specialized form of the plural suffix as zero must refer to a specific list. Plural Number is realized as follows:

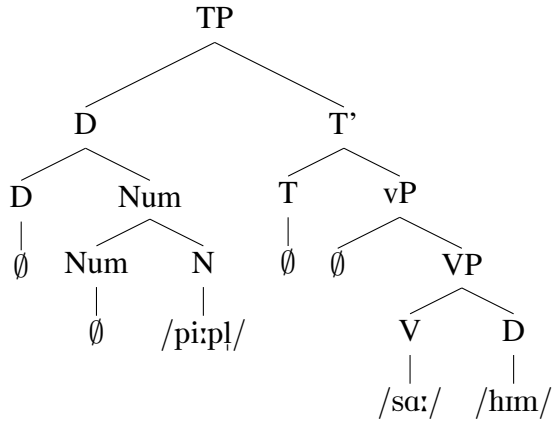
(12) a. [+pl]$_{Num}$ ↔ ∅ / _ {fish, sheep, shrimp, deer, people, . . .}
b. [+pl]$_{Num}$ ↔ /-z/ (everywhere else)

The last part of the tree to be discussed is the verb. In this course, we will adopt, as much as possible, the syntactic analyses put forth in Adger's textbook *Core Syntax*, simply to have a fixed and comprehensive syntax to refer to as a consistent part of the overall architecture. In Adger's book, he proposes that the heads T and *v* undergo the operation *Agree* for tense features. This would mean that the *v* in our tree would also bear [+past]. As a result, the allomorphy rule for the root $\sqrt{seew56}$ adopts a format similar to that of $\sqrt{prs341}$:

(13) a. $\sqrt{seew56}$ ↔ /sɑː/ / [+past]$_v$
b. $\sqrt{seew56}$ ↔ /siː/ / [−past]$_v$

The resulting tree, with all its exponents now inserted, is:

(14)



This concludes the discussion for now. It is important to re-emphasize a few points. Notice that we

do not use the term 'morpheme' in this book, as the morpheme is traditionally called 'the smallest unit that pairs sound and meaning'. First of all, many things that would be called morphemes have no meaning (e.g. the *-s* in scissor-s). Second of all, there is no direct pairing of sound and meaning in the late-phonological-insertion model. Instead, as in the 'inverted Y-model' of the grammar adopted in minimalism since Chomsky (1993), LF (semantic interpretation) and PF (phonological packaging) are independent, orthogonal routes that formatives undergo after the syntax is finished. The term 'morpheme' is a decidedly 'undistributed' notion, and the step-by-step model we've just gone through keeps Formatives (the input to the syntax) independent from Exponents (which are based on the output of the syntax, and hence what ends up next to what in the tree).

There are a few other remarks to make. First, the features we have adopted above have not yet been motivated or discussed; that will be taken up in Section 2. Importantly, a formative is a single leaf of a syntactic tree, but it itself can have a complex internal structure of many features. Secondly, the question of linearization of morphemes has not been discussed; that will be revisited at various places. Finally, it is important to note that not all work done within Distributed Morphology assumes that all roots undergo late phonological insertion; however, the allomorphy found with roots like *people* suggests that it is necessary, and rather than treating roots differently from other formatives (arguably a more complicated theory for the beginning morphologist, and not necessarily the right one!), we will treat them using the same formalism throughout these lectures.

2. Featural Decomposition

What are the elements that make up syntactic terminals? We intuitively know that a pronoun like *we* is something like '1st person plural'. But how are these parts of its contribution to the morphosyntax and morphosemantics structured inside the word? In fact, one of the insights of morphological theory is that terminal nodes (items like *we*, that inhabit a subject position within spec, T in the syntax in a given configuration) have an internal structure, much like the consonants and vowels (or 'segments') of phonology have an internal structure as well. We wouldn't say that *we* corresponds to three *distinct* syntactic terminals (e.g. we wouldn't say that there is a '1st person phrase', a 'plural phrase', and an 'unspecified gender phrase', that all constituted different heads in the syntax), because the syntax seems to treat *we* as a single leaf of the tree, in terms of the fact that it is assigned case as a single element, moves as a single element, can be deleted (in languages and configurations that allow it) as a single element, and does not show any evidence for movement to positions 'inside' it. The fact that it is '1st person plural, unspecified for gender' is a result of a featural structure *internal* to this single terminal node (which is another way of saying, this single leaf of the syntactic tree, which may very well, like a leaf, have cells and other structures inside it).

Similarly, it is a consensus that a segment like /p/ in phonology is actually a combination of various subatomic elements called features, such as [−sonorant, −continuant, −voice, −nasal, +labial], and that these features not only define the subparts of /p/ but also allow the sound to be grouped in natural classes with others. For example, /p/ and /b/ are often identical in terms of the kinds of consonant clusters they can participate in, and hence the fact that both share [−sonorant, −continuant, −nasal, +labial] allows one to characterize that. In yet other phonological patterns, /p,t,k/ pattern alike in triggering devoicing of following sibilants, e.g. in plural formation, and the feature complex [−sonorant, −continuant, −voice, −nasal] allows one to capture these three as a natural class. But the element /p/ patterns as a single segment for processes such as syllabification, deletion, or metathesis, which operate at the level of segments as units.

There are, of course, affricates, which are the union of two distinct segments under a single timing slot – thus /t/ and /ʃ/ can join up to form /t͡ʃ/. Similarly in morphology, certain distinct terminals can join up to form more complex units (the way that atoms can join up to form molecules) via operations like head-movement (also called head-adjunction), which puts two syntactic leaves together, side by side, at which point the two move along together as a complex unit from then onwards. In fact, precisely this kind of head-movement, joining up two atomic terminals to form a more complex one, is what constitutes formation of an *M-word* – a morphological word – a technical concept to which we will return at various points in the discussion.

But for now, let us focus on the structure internal to terminal nodes such as *we*, and ask how features like person, number, and gender are organized. Since at least Aristotle it is common to talk about 'first person, second person, third person' as three categories. However, the morphology of languages may not necessarily be organized in exactly these terms. The Aristotelian view would say that there are three values of the feature 'person', with values ranging over 1,2,3. But linguistic processes are often organized in more fundamental ways. Let's consider, as a parallel, the organization of vowel height in phonological systems, in which there are typically high, mid, and low vowels (as in Slovak, below):

(15)

i	u	[+high, −low]
e	o	[−high, −low]
æ	a	[−high, +low]

The question is whether this should be described in terms of "height 1, height 2, height 3" or in terms of other types of features, and the general consensus in phonology is that even though this is a three-height system, the best description in terms of natural classes is one with the two binary features [± high] and [± low]. Now, these two binary features and their values importantly have a semantics:

(16) [+high] is true if the tongue body is higher than a certain threshold H
[−high] is true if [+high] is false
[+low] is true if the tongue body is lower than a certain threshold L
[−low] is true if [+low] is false

It should be clear, given these semantics, that while a vowel (say a mid vowel) can easily be [−high,−low] (say, if the tongue body is in between H and L, and hence neither above H nor below L) it cannot ever be [+high,+low], as the tongue body could not be simultaneously above H and below L. Thus, although one might initially think that a system of two binary features 'overgenerates' in producing four combinations when in fact (16) has only three, in fact the fourth combination does not arise once we given these features – as we should for all features – a semantics of when they are true or false. We then might ask what properties this particular system of features has, in terms of grouping of natural classes. As you can see, this system groups together the high vowels and mid vowels as a natural class (as they both share [−low]), and it groups together the mid vowels and the low vowels as a natural class (as they both share [−high]). The world's languages show a plethora of processes in which high vowels and mid vowels pattern together to the exclusion of low vowels, and in which mid vowels and low vowels pattern together to the exclusion of high vowels, but crucially there are not processes in which high vowels and low vowels pattern together to the exclusion of mid vowels, and the system of features above delivers this result, as there is no feature-value shared by high vowels and low vowels to the exclusion of mid vowels. The choice of a system of features, therefore, is crucial to describing natural class behavior.

This immediately invites the question, what about person, which also involves three categories, and in which we might think these three categories show certain affinities to one another such that not every natural class grouping is equally natural to express. In fact there are a great deal of syntactic processes that group together 1st and 2nd person to the exclusion of 3rd – one of the most striking ones is the phenomenon of *partial pro-drop* found in languages like Finnish, which, like Italian, allow subject pronouns to be omitted, but differently, allow this only for 1st and 2nd person, and not for 3rd. This system is called 'partial pro-drop' precisely because not all subject pronouns can be dropped. In this case we might say that what 1st and 2nd person have in common is that they are *participants* in the speech event – I'm talking to you, and you're listening to me, whereas *him, her,* and *them* are not part of the conversation. Let us therefore say that 1st and 2nd person are [+participant], whereas 3rd person is not.

There are also a number of phenomena that seem to distinguish the 1st person from all others.

The one that we'll focus on for now is gender specification in pronouns. While many languages (like English) show gender distinctions for third person (e.g. *him* vs. *her*), and yet others show distinctions among the 2nd person (e.g. Arabic and Hebrew), it is quite rare for languages to show gender distinctions in the 1st person (e.g. to have a distinct pronoun, say *I* and *Ia* as grammatical pronouns differentiating masculine versus femine speakers). We can encode the different natural class patterning of 1st person versus the rest in terms of a feature [+author] (using the term neutrally for authors of speech events, signing events, or writing events). The resulting set of three pronominal categories would be as follows; note that under this view, '1,2,3' become descriptive labels without a morphosyntactic reality, which we can use simply for convenience.

(17) 1 [+part, +auth]
2 [+part, −auth]
3 [−part, −auth]

It would seem that we're well on our way to having a set of binary features to describe the three categories of person, and indeed, can attribute a semantics to them, such that [+participant] means that the referent (which could be plurality) includes a participant in the speech event, whereas [+author] means that the referent includes the author of the speech event. Given these semantics, the featural combination [+auth,−participant] is ruled out because it is a contradiction for the referent to include the author of the speech event but not include a participant in the speech event, and hence the two binary features yield precisely the three categories we find in English. Indeed, the system in (17) is all that's needed to describe the pronominal inventory of languages like English. But there are hundreds of languages that include a further distinction, sometimes called a 'clusivity' distinction, between 1st person plurals that are *exclusive* and *inclusive*. For example, suppose your father says to you *We are proud of you* – the pronoun *we* includes the author of the speech event but it does not include you – it is hence what's called a 1st person exclusive. By contrast, when I say to you *Let's go*, I mean that you're intended to go too – and in this case it's what's called a 1st person inclusive. Some descriptive traditions say that the category of 1st person inclusive constitutes a 'fourth person' – but at this point it should become clear that using the ordinal numbers is a bit absurd – why not call it the '1st and a halfth person'. The problem that arises is how to characterize the distinction between 1st person inclusive and exclusive featurally – do we simply introduce another feature? While this solution has been proposed at times, we will revisit this problem for person categories with a wholly new approach to morphosyntactic features, in which features *are not* simply conjoined together and interpreted in terms of logical conjunction (e.g. [+F,+G] means that F is true and that G is true), but rather in a more complex way of combining – to which we return only after an important tangent through the question of what features are used to describe *number* distinctions.

3. Number Distinctions

Many languages make a distinction not only between singular and plural, but also include a third category, dual number, precisely to refer to groups which contain exactly two. We have a minor inkling of what this would be like in the distinction between *all* and *both*, where the former is only

used for groups of three or more (and thus to say *All my parents wear glasses* immediately invites the conclusion that you must have more than two parents, for whatever reason). Now, in languages like Sanskrit, Slovenian, and Classical Arabic, dual number is more than just something like having the word *two* or *both* modifying the noun phrase in question, because the dual category constitutes a distinct set of suffixes on not only the noun in question, but also on verbs and adjectives that *agree* with this noun. As soon as we've seen that something like 'dualness' on a noun can trigger a distinct category of agreement suffixes on verbs and adjectives that themselves are not 'two', we know we're dealing with a morphosyntactic feature above and beyond a simple numeral or numeral quantifier.

So, what kinds of features would make sense for a three-way number system opposing singular, dual, and plural number? Well, first of all we might begin by considering what features lies behind the basic singular-plural distinction. This distinction is, in fact, grounded in the notion of 'singulars', and we will propose that the defining characteristic of number systems is whether something is a singularity or not. What this means, interestingly enough, is that in expressions like *zero dogs* in English, the 'plural' *-s* occurs not because zero is 'plural', but rather because this quantity is *also* nonsingular. In other words, the affix *-s* in English is not for plurals, it's for 'nonsingulars'. Hence, the basic number binary number feature for distinguishing *he* from *they*, *dog* from *dogs*, etc is [± singular]. Under this view, of course, dual and plural are both [−singular] (justifying their patterning as a natural class in many languages, in which, say, adjectival agreement distinguish dual and plural, but in which dual and plural constitute a collapsed category for verbal agreement). So what feature distinguishes dual and plural? The proposal here is that a feature [± augmented] does the job, but we need to discuss the semantics of this feature:

(18)

singular	[+sg, −aug]
dual	[−sg, −aug]
plural	[−sg, +aug]

Again, notice that we have three categories and two binary features above.[3] And while [+singular] is true if the referent is a singularity (and the negative value if not true), what [± augmented] contributes is not simply another true-false predicate of the referent – what it does is *act on another feature*. The intuition is the following: the dual is [−singular], but it is perilously close to being [+singular] (while the plural is comfortably far from being [+singular]). The way this is encoded is by saying that the value of [± augmented] for [−singular] is interested in whether there is a proper subset of the referent that is *still* [−singular].

(19) [+augmented] = $\lambda \mathrm{P} \lambda x \exists y [y \subset x \wedge \mathrm{P}(x) \wedge \mathrm{P}(y)]$

In prose, what [+augmented] means is: given some predicate P that is true of some set x, x is [+augmented] if there is a proper subset of x for which P is also true. Thus, in the case that the predicate P is [+singular], and the reference set contains only one member, this set is [−augmented] for its value of [±singular], because there is no proper subset that is still [+singular].

[3] Don't worry, not all things come in threes – there are two-height and four-height vowel systems, person systems with four categories, as alluded to above, gender systems in twos and threes, and number systems with even more than three categories

Sets of cardinality one are thus [−augmented] for [+singular]. Now consider a set of cardinality 100. This set is [−singular], and there is indeed at least one proper subset (in fact, many of them) that is still [−singular]. Thus, this set is [+augmented] for its value of [±singular]. Finally, consider a set of cardinality 2. This set is [−singular]. However, *there is no proper subset of a set of cardinality 2 that is still [−singular]*. Thus, this set is [−augmented] for its value of [±singular].

Notice, however, that the feature [± augmented] does not just combine with [± singular] as an independent truth-value. Rather, [± singular] is *evaluated first* as true or false, and then [± augmented] applies to the result, and is evaluated as a second feature that states whether the result of applying it to the first feature is true. So, features like [± augmented] are not like phonological features [± voice, ± labial] that simply conjunctively combine.[4] In a sense, what [± augmented] does is, for any feature F, says whether it 'minimally' satisfies it or not: a dual only 'minimally' satisfies [−singular], and so is [−augmented].

Before proceeding any further, you might be wondering if [± augmented] is relevant for languages without a dual, such as English.[5] There are two answers to this question – both depend on one's view of language acquisition, in a sense. One might say that universal grammar is such that all categories are universally available, even if not expressed in a certain language – and thus, say, that Mandarin Chinese has a Tense node in its syntax, even if never morphologically expressed. Pursuing a parallel with phonology, on the other hand, in which learners of a language without a [± lateral] distinction quickly discard this category from their phonology before 9 months of age, we might conclude that [± augmented] is not present in the number system of English once the learner has reached a certain age by which s/he has received no evidence to maintain it in the grammar. Nonetheless, when asked to write out a complete set of number features on exams and problem sets and the like, it might be good to actually write out features like [± augmented], for the sheer reason that once you start working on a language that *does* need this feature, you'll already be in the habit of writing it out and being comfortable with its contribution.

4. Returning to Person

Now let's return to person. The key notion we have just learned is that features can operate on other features. Recall that the problem we are concerned with is how to represent inclusive / exclusive distinctions, without necessarily introducing yet another feature. Instead, we will modify the way that the features compose in terms of their morphosemantics. For person, let's suppose that what binary features do is *add* individuals to the referent or take them out, an idea introduced for person features by Daniel Harbour.

[4] An analogy could perhaps be made with the feature [± delayed release], which only could be true or false once [± continuant] was already false, but this feature has now been abandoned with the autosegmental representation of affricates. Perhaps the closest analogue in phonology are versions of Element Theory that employ headedness, such that {A,I} represents /æ/ while {A, I} represents /ɛ/; in other words, in which the primes do not combine conjunctively, but rather, one is a head/operator on the other.

[5] While we will not show it here, recursive use of the feature [± augmented] can account for languages which in addition to singular, dual, and plural have a *trial* – an inflectional category for sets of cardinality three. The fact that typologically, a language will only have a trial if it already has a dual can be derived from the fact that in a system with [± augmented], having a distinction between trial and non-trial necessarily involves recursively building on the dual/non-dual distinction.

For this to begin to make sense, let's first define some notation based on set theory, where we introduce some individuals and sets: *i* is an individual that is the author, *u* refers to the addressee, and *iu* is a 'plural individual' (a kind of set) containing both the speaker and the addressee. What about those who are not the author or the addresse? Let's call these *o, o', o",* ... for all of the 'others'.

(20) Now, let's define a shorthand for a range of possible sets:

a. i_o: This includes all the possibilities of *i* plus zero or more others.
b. u_o: This includes all the possibilities of *u* plus zero or more others.
c. iu_o: This includes all the possibilities of *iu* plus zero or more others.
d. o_o: This includes all the possibilities of zero or more others.

We can now turn to what the plus or minus values of features do in this new system. Suppose that [+ auth] adds i to possible referent in a denotation, while [−auth] takes it away from each one. Similarly, [+ participant] adds each of $\{i, u, iu\}$ to every possible referent in a denotation, while [−participant] takes it away from each one.

Finally, let's assume that all references start out 'by default' with the referential possibilities of o_o: they can refer to zero or more others. Let's now see what these features do when applied to this set of possible referential denotations:

(21) a. [+author] applied to $\{o_o\} \rightarrow \{i_o\}$
b. [−author] applied to $\{o_o\} \rightarrow \{o_o\}$

(22) a. [+participant] applied to $\{o_o\} \rightarrow \{i_o, u_o, iu_o\}$
b. [−participant] applied to $\{o_o\} \rightarrow \{o_o\}$

The effect of these features on their own is simply to add or delete possibilities for the denotational possibilities that a pronominal item can contain. So far, these carve out essentially the same natural classes as we saw in (17) above. But what about the combinations? Now the order matters. Specifically, if we apply one first (let's put this inside parentheses) and the other one afterwards (let's put this outside parentheses, to the left), then the second one will modify the result of what happened by applying the first. Now, while there are four combinations, all of which are shown below, it will be useful to go through them one by one in the following order for explanatory purposes.

(23) a. +author(+participant): $\{o_o\} \rightarrow \{i_o, iu_o, u_o\} \rightarrow \{i_o, iu_o\}$
b. −author(-participant): $\{o_o\} \rightarrow \{o_o\} \rightarrow \{o_o\}$
c. −author(+participant): $\{o_o\} \rightarrow \{i_o, iu_o, u_o\} \rightarrow \{o_o, u_o\}$
d. +author(-participant): $\{o_o\} \rightarrow \{o_o\} \rightarrow \{i_o\}$

Let's go through these one by one. (23-a) adds *i,u,iu* to every possibility first, and then it takes the result of that and adds *i* to every possibility. The result is $\{i_o, iu_o\}$, which means it can refer to the speaker plus zero or more others, and the speaker and hearer plus zero or more others. When such a reference set is [−singular], this means a pronoun with these features can refer to the first person plus others, or the first and second person plus others – just like *we* in English. In

fact, when [+singular], the only individual that would satisfy this cardinality within the referential possibilities is i. Thus (23-a) entirely corresponds to the first person in languages like English.

(23-b) takes out all participants, and then takes out the author (which wasn't there anyway), leaving only others. This is relatively straightforward, and corresponds to third person, whether [± singular], in languages like English.

(23-c) adds in all participants to each plural individual, but then takes the author *i* away from these. Taking *i* away from i_o has the result of leaving o_o as a referential possibility – meaning that a pronoun with these features can refer to either u_o or o_o. Given, however, that we have just seen that (23-b) is reserved for o_o, we must invoke a principle of lexical complementarity. Specifically, for semantic cases where there are two items, and one, M has a strict subset of referential possibilities of the other, N, we may say that the use of N is confined to cases not covered by M. In the case at hand, this means that since (23-c) can refer in principle to either u_o or o_o, but (23-b) only to o_o, then (23-c) will be restricted to u_o. With this principle in hand, (23-c) corresponds to the second person, whether [± singular], in languages like English.

We finally turn to (23-d) – a feature combination which we originally began looking at specifically for languages like Warlpiri (spoken in Australia) or Kadiwéu (spoken in Brazil), which include a specific distinction between exclusive and inclusive 'we'. (23-d) first takes out all participants, leaving only o_o, and then adds in the author, hence creating i_o. Here's where the interesting payoff of the Lexical Complementarity principle just invoked above comes in. In a language which allows both combinations (23-a) and (23-d), since the former can denote either i_o or iu_o, and the latter only i_o, when both are present in a language, the former will be restricted to iu_o. This will exactly deliver the result that (23-a) is restricted to inclusive 'we' only when the language also has (23-d), the strictly exclusive 'we', competing in the language as a distinct category. Otherwise, (23-d) will not be present, and (23-a) can mean either exclusive or inclusive 'we'.

Taking a step back, essentially what these features have done with these semantics and their combination is allowed [+author,−participant] to "make sense" without being a contradiction. This is precisely because instead of each feature being independently evaluated as true or not, one feature acts on the result of the other. In this sense, rather than as an unordered set, we can represent these features as ordered sets, e.g. <+a(-p)> for (23-d). Finally, notice that we say nothing special about the 'other' per se, as it is always present from the beginning in the denotational possibility of a reference sets, kind of the way that 'schwa' is a carrier signal that is modulated by adding or subtracting certain acoustic signatures in Element Theory.

It's important to emphasize that the feature combination <+a(-p)> is not distinguished from <+a(+p)> in English; *we* and *us* in English seem to basically always refer to any set containing the author, without specifically excluding the addressee (unless the pragmatics make it salient one way or the other, in expressions such as *let's go*, as most imperatives cannot really exclude their addressee). But there is no need for a dedicated exclusive category in English, as in the 'marked' combination of (23-d), and so we might say that once a learner acquires English morphosyntax by the relevant age, the feature [±participant] is not contrastive as an internal argument of [+ author].

5. Unvalued Features, Agreement, and Concord

The features on an auxiliary like *am* or *were* include tense features, such as [± past, ± finite], where [−past, +fin] are exponed by forms like *is*, and [−fin] for forms like *be*. Crucially, however, they also include agreement features, such that *am* is used for [−past, +fin, +part, +auth, +sg] terminal nodes. At this point it is important to make a distinction between the features on such nodes that are inherent to them (such as [−past]), and those that have been acquired through the course of the derivation. This draws on the notion from Minimalist syntax of *unvalued* features – namely those features that a formative comes with that are indicated by a *u*, meaning that they have no value at all yet, and need to acquire one within the syntax. Thus, the formative node that will be eventually realized as *am* has the following structure *before it is merged into the syntax*:

(24) Formative: [−past, +fin, $u$part, $u$auth, $u$sg]

Features such as [$u$auth] need to be valued within the syntax, and are done so by the operation *Agree*, extensively developed in Minimalist models of syntax. This operation, enacted by the T node in the syntax, will Probe and Copy the values [$u$part, $u$auth, $u$sg] from a Goal that it c-commands. Thus, in a sentence such as *I am going to the store*, the T node will Probe and Copy the values [+part, +auth, +sg] from the subject DP that starts below it, is Agreed with, and subsequently comes to sit in its specifier (the specifier of TP). In other words, it is predictable, once the syntactic configuration is given, what values the features [$u$part, $u$auth, $u$sg] will have, and hence these are not stored within the Formative. On the other hand, features such as [−past] on the auxiliary are unpredictable, and must be stored as part of the formative. This is roughly akin to the notion of underlying forms in generative phonology, whereby the unpredictable information is stored in the underlying form whereas predictable features that will be acquired allophonically are not. For this reason, it is sometimes useful to terminologically distinguish *formatives*, which are the input to syntax, and may bear unvalued features, from *terminal nodes*, which may acquire new feature-values and properties during the syntax (through Agree, case-assignment, movement, and so forth). It is the latter which are sent to PF to be turned into phonological content via the Exponent List.

The agreement operation outlined above for Tense nodes also holds for what is called Concord, i.e. agreement that takes place within the determiner phrase. While there isn't much overt evidence for Concord in English, there is in fact number agreement borne by demonstratives, as evidenced by the contrast between *that* and *those*. For languages such as English with a proximal/distal distinction in demonstratives, we would say that the formatives are [± distal, u singular]. As the result of an Agree operation between D and the Num head, a terminal node that comes to acquire the features [+distal, −singular] would be spelled out as /ðowz/.

6. Gender and Case Features

The approach described above for Number and Person, in which features combine in ways beyond simple symmetric conjunction, but rather apply in specific orders, is associated with the work of

Daniel Harbour, and is a promising and active area of research in person and number.[6] As it has not been applied to or explored for gender or case systems, and there is no immediate empirical need (of the sort there was with inclusive versus exclusive we) to go beyond the simple symmetric conjunction for these, we will present the features for gender and case here for completeness of exposition.

Gender features do not correspond directly to real-world semantics as easily as person and number features, because in many languages gender applies to inanimate nouns like carpet, spoon, and dirt, in which there is not even a division into masculine and feminine pairs. However, binary opposition and natural classes may still be found. In English while it is not obvious whether [± feminine] or [± masculine] would be the right binary division to make, the fact that certain inanimates (e.g. ships, nations) may be assigned feminine gender in English suggests that the right predicate might be [± feminine]. Moreover, in languages with three genders (e.g. Greek, Russian), the masculine and neuter often pattern together, suggesting that these latter two are both [− feminine]. We thus will adopt the proposal that two-gender systems are [± feminine], whereas three-gender systems subsequently activate an additional feature [± masculine].[7] In fact, such a feature may be needed for English *it*, which is arguably neuter (although could be coded as simply inanimate) and for uses of singular *they*, e.g. *Ask one of your classmates if they can be your partner for the assignment to be done in pairs*, where it clearly refers to a singular, animate individual but with unspecified gender. We thus will adopt the view that *it* and *they* are both [−masc,−fem] but differ in the feature [± animate].

Let's return to where we began, looking at the internal structure of a pronoun like *he*. What features correspond to the pronouns *he, him, his*? Up until now, besides case, we have the following:

(25) [−auth, −part, +sg, −aug, −fem, +anim]

Now, suppose we treat case features like 3-vowel heights. [± nominative, ± genitive]. Their semantics would boil down to predicates true of syntactic configurations: [± nominative] would be true of a noun phrase that was assigned case in Spec, TP while [± genitive] would be true of a noun phrase assigned case in spec, DP. The negative values would mean neither of those respective predicates are true. We can write down the exponent list as follows:

(26) /hiː/ ↔ [−auth, -part, +sg, -aug, -fem, +anim, +nom, -gen]
/hɪm/ ↔ [−auth, -part, +sg, -aug, -fem, +anim, -nom, -gen]
/hɪz/ ↔ [−auth, -part, +sg, -aug, -fem, +anim, -nom, +gen]

Now, for the feminine versions of these pronouns, it seems there is a syncretism:

[6] Ackema & Neeleman (2013) have also recently developed a theory of person features in which the features are operations that may successively add or remove elements from sets in an ordered fashion, with different operations and thus some different effects from those of Harbour, though not of direct consequence for us in the present discussion.

[7] Languages with many genders (e.g. Bantu noun classes) would have a richer feature structure. Note however, that classifier systems like those of Chinese are probably different in nature, since a 'default' classifer can be used (albeit it sounds less specific) while using the wrong gender in a gender language is simply ungrammatical.

(27) /ʃiː/ ↔ [−auth, -part, +sg, -aug, +fem, +anim, +nom, -gen]
/hər/ ↔ [−auth, -part, +sg, -aug, +fem, +anim, -nom, -gen]
/hər/ ↔ [−auth, -part, +sg, -aug, +fem, +anim, -nom, +gen]

Rather than writing down the same exponent twice, we can redo things as follows:

(28) /ʃiː/ ↔ [−auth, -part, +sg, -aug, +fem, +anim, +nom, -gen]
/hər/ ↔ [−auth, -part, +sg, -aug, +fem, +anim]

In the formulation above, the exponent /hər/ is underspecified – that is *un*specified – for case features, meaning it is in principle compatible with any terminal node coming from the syntax. The manner in which underspecified exponents are chosen to be inserted at a terminal node given a competition with ones that are not underspecified is the topic of the next two sections of lecture notes.

7. Allomorphy and the Maximal Subset Condition

Case features are acquired by a terminal node in the syntax. They do not come from the formative list, but rather are acquired in certain syntactic configurations. Thus, nominative case is assigned by finite T, and a terminal node in spec T acquires [+nom,−gen]. Accusative case is assigned by *v*, and a terminal node would thus acquire [−nom,−gen]. Finally, genitive case is assigned by Spec, D, and a terminal node in this position would acquire [−nom,+gen]. Now, on the exponent list, if the gender is [−fem], there are distinct exponents for all three of these: *he, him, his* respectively. But if the gender is [+fem], then a certain exponent, namely *her*, is used for two distinct sets of case features in the syntax. Thus, there is one exponent that corresponds to two morphosyntactically distinct terminal nodes. We capture this by saying that this exponent is *underspecified*.

Recall the question of the case features of *her* on the exponent list:

(29) [-part, +sg, +fem] ↔ /hɔɪ/

The way we've written them, what blocks use of *her* in the nominative context as well? It is perfectly compatible with a terminal that exits the syntax with the feature [+nom]. So why can't one use this exponent to realize a syntactic terminal that acquired [+nom] in the course of the syntax? It must be the notion that there is a more specific form, namely /ʃiː/, that blocks the use of *her* in this context. In other words, given two items which have a similar set of features, the more specific one will win the competition. This is what's called the Elsewhere Principle in the literature, but which we will call the Maximal Subset principle in order to reflect more clearly the logic of how it works. In order to use an exponent E in order to realize a terminal node T, the features on E must be a subset of (this includes being identical to) the features on T:

(30) f(E) ⊆ f(T).

But suppose that there are two exponents, E1 and E2, the featues of both of which constitute subsets of the features on the terminal node T to be realized. How is the choice made of which one to insert?

(31) f(E1) ⊂ f(T), f(E2) ⊂ f(T), and f(E1) ⊂ f(E2) → E2 wins the competition

In other words, while E1 and E2 both contain subsets of the features of T, E2 contains the *maximal* subset – more features than E1. To take a concrete case at hand with a nominative third-person feminine subject pronoun, in the following scenario, E2 wins the Maximal Subset competition:

(32) Terminal node:[-part, +sg, +fem, +nom,−gen]
E1 [-part, +sg, +fem] ↔ /hɔɪ/
E2: [-part, +sg, +fem, +nom] ↔ /ʃiː/

Given the fact that the exponent E1 is actually used not only in the one non-nominative context of the accusative, but in fact in *all* non-nominative contexts (i.e. the genitive as well, in addition to environments like fragment answers in which potentially no case at all is being assigned), it appears to be the default, non-specialized exponent for third-person feminine singular pronominal

terminals. Putting things differently, ↔ /ʃiː/ is used in the specific combination of nominative features, whereas ↔ /həɪ/ is used everywhere else, or 'elsewhere'. The notion of an Elsewhere Item, then, is the allomorph that has the most heterogeneous distribution – somewhat akin to listing the allophones of a specific phonemes and referring to the one with the most widespread distribution as the default form of the allophone, but one which is blocked from usage when more specific forms override it.

This principle of competition, whereby the most specific item wins, is found with stem allomorphs such as the English copula system, but also found in affixes as well. The reason that forms like **childs* or **hitted* sound odd is because they are using the 'wrong' allomorph – in the former case, a more specific allomorph /-ɪɛn/ wins over the elsewhere /-z/; in the latter because the more specific allomorph /-∅/ wins over the elsewhere /-(ə)d/.

Elsewhere allomorphs in morphology thus tend to be the most productive exponent of the terminal in question. While this seems relatively straightforward for cases like the plural suffix in English (which has the highly specific allomorphs found in plurals like *children, stimuli, stigmata*) – new nouns and loanwords productively take the default plural suffix /-z/ instead of these – it can also be found even for auxiliaries – consider the case of the past tense copula in English:

(33)

	sg	pl
1	was	were
2	were	were
3	was	were

We will represent these in terms of the three entries below (note that there is homophony of the two forms of *was*, a fact we can come back to later, which does not affect the exposition about *were* at this point).

(34) $[+\text{cop}, +\text{past}, +\text{auth}, +\text{sg}]_{Aux} \leftrightarrow /\text{wʌz}/$
$[+\text{cop}, +\text{past}, -\text{part}, +\text{sg}]_{Aux} \leftrightarrow /\text{wʌz}/$
$[+\text{cop}, +\text{past}]_{Aux} \leftrightarrow /\text{wəɪ}/$

Given the two allomorphs found distributed across these six person/number combinations, the elsewhere form is the one with the most widespread distribution – and this is clearly *were*, found in a motley array of four of the six, and one with no person or number specifications at all. In fact, there is evidence that *were* is the default form, about which nothing special needs to be said – it bears no features at all specific to person and number on this view – and is found everywhere that a more specific form can't be used. Strikingly enough this even includes subjunctive forms of the verb, in which person distinctions are lost – but lost in favor of the default – hence *Were I taller,...* with a mismatch in the presence of a singular pronoun, and not **Was we taller,...* with a mismatch in the presence of a plural pronoun; in a sense, this shows the productivity of the elsewhere form in (34), *were*, beyond its original contexts.

The notion of an elsewhere item, like /həɪ/, which is found as the exponent of two distinct terminals even though they have different features – or *were*, which expones four featurally-different terminals – is crucially couched within models involving late phonological realization of syntactic structures; ones in which the morphology and phonology *interpret* the output of syntax, and do not

play any role in the formative list that constitutes the input to syntax.

On this view, the fact that there are four phonologically 'homophonous' versions of the word *were* is completely independent of the fact that the formatives and terminals in the syntax of these four respective environments are indeed distinct, and that in another language with another set of exponents, they might indeed receive distinct realizations. The fact that they do not have distinct realizations in English has no effect on their syntax, which by hypothesis operates independently of the eventual phonological form of these items.

Compare how this would work in an early-insertion framework (i.e. lexicalist, 'checking' theories), in which words enter the syntax as fully-phonologically specified formatives. What would enter the syntax in the case of *were*? On the one hand, one could include four homophonous forms for *were*, instead of the one single one in (34). This would clearly be missing a generalization; it would be akin to listing the rule of voicing assimilation as a three separate rules in English (e.g. *hat-s/head-z, cup-s/cub-z, lock-s/leg-z*: one for coronals, one for labials, one for dorsals, even though they all have the same effect and can be reduced to a single statement in the grammar. So how could a pre-syntactic, lexicalist phonology get away with having only one entry for *were*? Perhaps it could negatively specify where it appears, e.g. "*were* is compatible with all past tense forms except those of the 1st and 3rd singular". This, however, misses its sudden extension to subjunctive contexts (e.g. *Were I taller,...*), as well as having to state *twice* the distribution of *was* (once in the specification of *was*, and again in the specification of *were*, to avoid the latter being used in the precisely the context of the former). Summarizing, in early-insertion approaches (i.e. *unlike* the one we are adopting in this course), there is no way to guarantee both syncretism (the fact that *were* is used in four different environments) and blocking (the fact that *were* can't be used in all six different environments) in a straightforward way.

Let's briefly return to the two separate entries for *was* in (34). Is there a way to unify them? In fact, there is, given what we know about the pronoun *you*, which notoriously shows no singular/plural distinction (and in fact, historically, comes from a plural form, later generalized to the singular as well). If we say that 2nd person pronouns in English simply *have no* singular/plural distinction (represented by a zero-value for this feature), we can rewrite the table as follows:

(35)

	[+sg]	[∅sg]	[−sg]
[+auth, +part]	was		were
[−auth, +part]		were	
[−auth, −part]	was		were

(Note also that there is no particular reason to write paradigms in a specific top-to-bottom and left-to-right order, once feature-values are used; we could have easily changed the orders of the columns and rows; paradigms, as such, in the order memorized in second-language classes, have no specific psychological correspondence to the geometric pattern above; and merely correspond to the intersection of certain feature-value combinations with others).

There are now only two exponents: one for [+singular] pronouns, and one for the rest (elsewhere) no matter what their precise specification for number may (or may not be):

(36) $[+\text{cop}, +\text{past}, +\text{sg}]_{Aux} \leftrightarrow /\text{wʌz}/$
$[+\text{cop}, +\text{past}]_{Aux} \leftrightarrow /\text{wəɹ}/$

Incidentally, a question arises as to where the category labels, such as Aux or D, should go in the list of morphosyntactic features on the lefthand side of the exponent list. While we indicate it with a subscript that indicates the label above a terminal node, it could also be indicated as a feature-attribute pair like those for binary-valued features, thus [past: +, Cat: D]. Where left implicit, we stick to the notation above, arguably more familiar from phonology.

Moreover, we have not included a root for the copular verb 'be' above; this may be because it literally is simply a bundle of features, with no root – that is to say, no encyclopedic content above and beyond its features, just like pronouns. What about 'have'? Well, we need to distinguish possessive and auxiliary have – notice that only the latter undergoes subject-verb inversion in my English, where I cannot say **Have you a car?* – the latter again seems to not be based on any encyclopedic root, bearing a relation only of grammaticalization to the former. (Similarly, the verb *tener* in Spanish, which means 'to have', literally meant 'to grasp' in Latin (and still does in Italian), but in Spanish, it has lost all encyclopedic content describing a 'manner of holding' and denotes only possession – one step of the way towards becoming an auxiliary – which in Portuguese, *has* fully happened, in which this verb is used in auxiliary constructions as well).

In (36), we have now stated the distribution of these two allomorphs precisely in terms of one specific and one default item, the choice of which is governed by the Maximal Subset principle on the output of syntax; within the syntax, these terminals are fully-specified for person and number, but the morphology of English simply does not have the resources to fully expone them with distinct items. The syncretism that occurs with the elsewhere item is not like the homophony that occurs with, say bat_1 (for sports) and bat_2 (a flying mammal) – syncretism is non-accidental homophony – the use of the same exponent for different morphosyntactic feature specifications, due to the interaction of the inventory of exponents on the list and the way that Vocabulary Insertion, necessarily after syntax, works.

8. Neutralization and Metasyncretism

We might say that the distinction between singular and plural is never found in the 2nd person pronouns of English, or that the distinction between participant and non-participant persons is never found in the plural in English, even though we know these distinctions to be independently available in other areas of the grammar. Paralleling with phonology, we can say that a potential contrast is *neutralized* – very much like languages with final-devoicing, in which the contrast between [± voice] (e.g. *p,b*) is found in some places in the language (namely, onsets), but is neutralized (and hence reduced to only one dimension) in a specific context (the coda).

When neutralization (another word for syncretism, really), is persistent across a number of distinct environments in the language, however, it looks like it may be a generalization beyond that of a single 'paradigm'. Consider the following patterns in Latin noun declensions (the word declension refers to classes of nouns; note that 'conjugation' refers to verbs), where there are five classes and four cases being considered (nominative, accusative, dative, and ablative):

(37) CLASS I:

case	sg	pl
nom	-a	-ae
acc	-am	-as
dat	-ae	-is
abl	-a	-is

(38) CLASS II:

case	sg	pl
nom	-us	-i
acc	-um	-os
dat	-o	-is
abl	-o	-is

(39) CLASS III:

case	sg	pl
nom	(variable)	-es
acc	-em	-es / -is
dat	-i	-ibus
abl	-e	-ibus

(40) CLASS IV:

case	sg	pl
nom	-us	-us
acc	-um	-us
dat	-ui	-ibus
abl	-u	-ibus

(41) CLASS V:

case	sg	pl
nom	-es	-es
acc	-em	-es
dat	-ei	-ebus
abl	-e	-ebus

In all five classes, the dative plural and the ablative plural have the same ending. This is a neutralization of a potential case distinction (one which *is* contrasted in the singular in many cases), but the fact that these are syncretic in every single one of the declension classes seems like it's a generalization above and beyond the inventory of specific exponents for each class – rather, it seems like there is something in the grammar which specifically *prevents* such a distinction from *ever* being realized, or, put differently, *enforces* this syncretism across all declensions. When the

syncretism occurs independently of the specific exponents (e.g. it's not about *-ebus* in class V or about *-is* in class II; it's about all of them – or put differently, none of them in particular), it's called a metasyncretism – a syncretism pattern above and beyond the facts of specific affixes in specific classes. In next week's lecture notes, we will return to the question of metasyncretisms, and the mechanism responsible for exponent-independent generalizations of this sort in the grammar.

9. Contextual Specification Also Matters in Determining the Most Specific Exponent

Once the syntax is done, the process of consulting the Exponent List and choosing an item to realize a particular terminal is called the process of *Vocabulary Insertion*. One of the key aspects of Vocabulary Insertion is that it respects what is called the Elsewhere Condition, which as we discussed, is a process of choosing the exponent that realizes a Maximal Subset of the features on the terminal to be exponed. We'll assume as a working hypothesis that Vocabulary Insertion proceeds terminal-by-terminal, from the bottom of the syntactic tree up to the top, turning each terminal of morphosyntactic features into a phonological exponent, one by one.

(42) f(E1) $\subset$ f(T), f(E2) $\subset$ f(T), and f(E1) $\subset$ f(E2) $\rightarrow$ E2 wins the competition

If you recall, exponents like *her*, underspecified for case, could never win over *she* in the appropriate (nominative) context, as the latter exponent bears case features, and hence the features borne by *her* are a subset of those of *she*; this is straightforwardly captured by the formalization above, in terms of the intrinsic features of competing exponents.

However, consider cases like **childs*, a form that is impossible, given the more specific way of forming the plural *children*. The two exponents in question are:

(43) $[-\text{sg}]_{num} \leftrightarrow$ /-z/
$[-\text{sg}]_{num} \leftrightarrow$ /-rɛn/ / _ $\{\sqrt{kiddo366}, \sqrt{bro47}\}$

Nothing in the formalization of (42) will guarantee that *ren* wins over *-z* in the context of this root, as both realize the same number of features. Instead, the *contextual specification* of two exponents must also be taken into account in order to compare which is the most specific for purposes of the Elsewhere Condition.

10. So How Could a Less-Specified Item Ever Win?

Consider dialects of English that say

(44) a. I'm not lucky.
b. Am I lucky?
c. Aren't I lucky?
d. *Amn't I lucky?

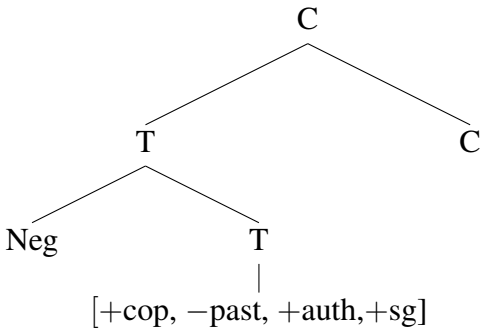
Why should (44-d) be substitued by (44-c), given the fact the exponent list below?

(45) $[+\text{cop}, -\text{past}, +\text{auth}, +\text{sg}]_{Aux} \leftrightarrow$ /æm/
$[+\text{cop}, -\text{past}, +\text{sg}]_{Aux} \leftrightarrow$ /ɪz/
$[+\text{cop}, -\text{past}]_{Aux} \leftrightarrow$ /aːɹ/

Assuming that the auxiliary has syntactically copied the features of the subject pronoun during agreement, why would it fail to realize the most-specific exponent in (45-a) specifically in the

context in (44-c-d)? It's not something about the sentence 'generally being negated', as it is fine to say *Am I not lucky?*. Rather it is something specifically about the auxiliary and negation *being part of the same M-Word* (and not just that, but also being part of the same M-word as a [+Q] feature that triggers the movement of the auxiliary to C in yes/no questions). In other words, the specific configuration that has the surprising, unexpected form *are*, rather than *am*, under T, is the following:

(46)



The ban on *amn't* is not phonological (as some dialects of English have it) – note that it is not repaired phonologically (say, by epenthesis of a vowel or deletion of a consonant) but rather *morphologically* – by insertion of the 'wrong' exponent. Recall we are talking about a highly derivational model of grammar, in which Vocabulary Insertion happens after the syntax is done. Suppose that in between these two steps is a specific moment in time during which *morphotactics* are evaluated. Phonotactics define what phonemes can 'touch' each other and the order in which they can do so (you may recognize the root 'tactile') – i.e. form part of the same syllable (as *pn-* was a possible sequence in Ancient Greek but is not in Modern English), and morphotactics defines what morphemes can touch one another, i.e. constitute part of the same M-word. Just as every language (and dialect) may have its own phonotactic restrictions, so may it have its own morphotactic restrictions.

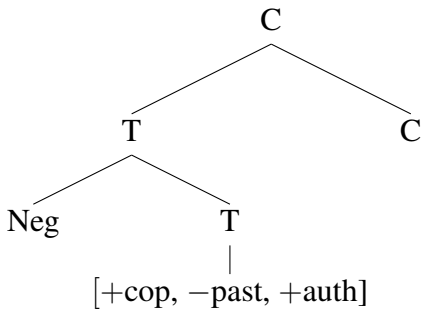
The application of these morphotactic readjustments takes place in another module, after syntax, and before Vocabulary Insertion (and hence way before phonology takes place). They are specific to questions of what can constitute a well-formed word in these language in terms of the kinds of features and terminals that can co-habitate in an M-word. In the dialects in question, (46) is banned, and the way it's repaired is to delete certain features. Specifically, the feature [+sg] seems to be deleted in this context *before the Exponent List is even consulted*, which means that the exponent *am* cannot be used when the time to do Vocabulary Insertion rolls around, as *am* bears a feature which is not present on this terminal node as a result of the morphotactic deletion.

In other words, some of the features that were present on the terminal node when it left the syntax 'got robbed' from it on the way to Vocabulary Insertion.

(47) Delete the feature [+sg] on a terminal node in the configuration in (46)

For this reason, deletion rules of this type are called *impoverishment* rules, as they make the terminal node less rich and less specified than it was when it left the syntax. The result of (47), in the post-syntactic tree before it is exponed, is:

(48)



As a result, an unexpected, less-specified exponent can suddenly win, as the more specified exponent is rendered irrelevant, and knocked out of the competition, by the operation of an impoverishment rule that makes it no longer eligible. In other words, the little envelope that Vocabulary Insertion, personified as an accountant, usually receives in the mail from the syntax, and takes and pastes on phonological exponents according to a diligent and straightforward bean-counting exercise of comparing the exponents using the Maximal Subset principle, does not know that the envelope was hijacked along the way, and that some of the features were removed from the terminals before arriving for Vocabulary Insertion. As a result, the Elsewhere item is the only eligible exponent to use, and *are* is inserted in this context, yielding *Aren't I lucky*.

In fact, a number of dialects of English also show what is called *weren't*-levelling, in which they say *He was lucky* but *He weren't lucky* — thereby demonstrating that impoverishment (called levelling because it 'levels' the distinctions between more specific and less specific forms in favor of the latter) happens rather generally on English auxiliaries in the presence of negation, even though the *amn't* example may be the most widespread and dramatic case. This suggests that negation is putting on a certain degree of morphotactic pressure on the configuration. Why negation? The idea to motivate here is that negation bears a kind of markedness.

Markedness refers to the asymmetric treatment in the grammar of two opposites, which are not treated as two symmetric options. One of them is treated differently. The simplest example might be a right turn on a traffic junction (in the UK). This is clearly more complex than a left turn at a traffic junction. Left and right are not treated as two symmetric options; one of them involves a more complex scenario, and we see that this doesn't necessarily boil down to one of these being more or less frequent options, as some intersections might involve many more right turns than left, but that doesn't change the fact that a right turn is more marked.

Notice that if we compare the two diphthongs in English, /aj/ (as in *mice*) and /aw/ (as in *mouse*), one of them is preferable / less marked than the other, namely /ai/: research in phonology has shown that it's not the case that one has a front glide and the other has a back glide and that these are two symmetric options. In fact /ai/ makes for a better diphthong because the vowel /i/ is also *higher* than /u/, and as a result lends itself to a diphthong in which the offglide is further away from the nucleus /a/, and hence shows an even greater contrast between the two halves of the diphthong. Now, this is a comparison of what's called the *paradigmatic* markedness of /aj/ vs /aw/ simply comparing them in terms of their inherent properties as items in the inventory. But interestingly enough, paradigmatically marked elements are also more limited in the ways that they can combine with *other* elements in more complex units. Thus, while the diphthong /ai/ can cooccur in a syllable with any following coda consonant (e.g. *time, tine, lied, jibe*), the diphthong

/au/ cannot: *town* and loud are possible, but not *towm / taum, jaube / joub* or anything of the sort – this more marked diphthong limits the consonants it combines with to be coronal only. This in fact illustrates what can be called *additive markedness*: while /au/ is allowed in plenty of words (but more marked than /ai/) and /b,g/ are allowed (but more marked than coronals, as extensively shown in a series of papers by Aditi Lahiri), the combination of these two marked categories is disallowed in words like **taum*. Returning to morphological markedness, the same occurs with (46): while [+auth] on its own, although marked, is tolerated, as is negation, putting these two marked categories together under subject-verb inversion instantiates an additive markedness that needs to be repaired.

Interestingly enough, the way that the tree in (47) schematizes the 'forbidden combination', it does not actually prescribe a specific repair, and one might imagine that other repairs besides (48) are possible. In phonotactics, many languages disallow two vowels in sequence (where one is not a glide). They ban what is called 'hiatus', with a constraint like *VV. But languages differ in their repairs – some delete the first vowel in this context, others delete the second, others delete the less sonorous of the two, and yet others fuse/coalesce the two vowels into one. We find the same kind of same-constraint, different repairs in the morphotactic constraint at hand – there are actually dialects of the UK in which instead of saying *Aren't I lucky?*, they say *Isn't I lucky?* (but don't say *I is lucky* – the deletion rule is still specific to the context of an M-word with negation under C). These dialects seem to implement the alternative impoverishment rule in (49):

(49) Delete the feature [+auth] on a terminal node in the configuration in (46)

By application of the elsewhere principle, the next-most specific item, still eligible in this context given the exponents in (2), would be *is*. In other words, as a result of deleting a different kind of feature, the resulting set of items left in the competition may be different.

11. Dissimilatory Deletion

The original motivation for impoverishment came from the following case in Spanish, where there are pronominal clitics that resume topicalized items, kind of like in *The horse, I sold it yesterday*. These pronominal clitics agree in person, number, gender, and case with the argument they're related to. In Spanish, the dative 3rd person singular one is *le*, and the accusative 3rd person singular one is *lo*.

(50) El caballo, lo vendí ayer 'The horse, I sold it yesterday'.
the horse, cl.3.acc sold.1sg yesterday

(51) A Pedro, le vendí el caballo ayer 'Pedro, I sold him the horse yesterday'.
To Pedro, cl.3.dat sold.1sg the horse yesterday

(52) [+pron, −part, +sg, −fem, −nom, −dat] ↔ /lo/
[+pron, −part, +sg, −fem, −nom, +dat] ↔ /le/

Notice that the pronominal clitics thus express different case features depending on whether they resume the direct or indirect object. But now what happens if you topicalize both of these argu-

ments, something like *To Pedro, the horse, I sold him it yesterday*. We'd expect something like the following, which is actually ungrammatical:

(53) *A Pedro, el caballo, le lo vendí ayer 'To Pedro, the horse, I sold him it yesterday'.
To Pedro, the horse, cl.3.dat cl.3.acc sold.1sg yesterday

In fact what you find instead is the 'spurious' clitic *se*. What is that? It's actually the most underspecified, elsewhere pronominal clitic in Spanish, used for impersonals, reflexives, reciprocals, middles – a whole range of heterogenous contexts. By hypothesis it has only the features below:

(54) [+pron] ↔ /se/

This pronominal clitic will never be used when a more specific one is available and matches the terminal to be exponed. Unless of course, that terminal has been mucked with on its way to Vocabulary Insertion, which is precisely what seems to be happening here.

(55) A Pedro, el caballo, se lo vendí ayer 'To Pedro, the horse, I sold him it yesterday'.
To Pedro, the horse, SE cl.3.acc sold.1sg yesterday

Something in the morphotactics doesn't seem to like having these two clitics next to each other. What could it be? In this case, it seems to be that the two of them are 'too similar' – a kind of dissimilation is taking place, and the first one is undergoing an impoverishment, after the syntax, that changes its features, or more specifically, deletes its features. Specifically, we might say that all of the person, number, case, and gender features are deleted. Why? Because they're so similar to those of the neighboring clitic. This is *like* the kind of dissimilation that takes place in phonology (although it has nothing to do with phonology per se; I'm simply drawing a parallel here) in words like *diphthong* [dɪfθaŋg], which in my English is never pronounced this way, but rather as [dɪpθaŋg] – in other words, I enact dissimilation, because having two voiceless fricatives, [f] and [θ] right next to each other, is banned, and must be repaired, in this case, by changing/deleting the feature [+continuant] on the first one.[8] In the morphotactic case at hand, it's entirely parallel: having two [+pron, −part, −nom] elements right next to each other cannot be tolerated, and must be repaired, which is achieved by impoverishing the majority of features on the first one.

12. Lack of Gender on 1st-Person pronouns

Impoverishment is used for cases where syncretism transcend properties of individual exponents, in other words for cases of meta-syncretism. Now recall that 1st person pronouns don't have gender in English. Not only do they not show this in *I*, but also in *me, us, we, my, our* – thus in *every* pronoun with [+author], regardless of case or number, a gender distinction cannot be realized. This is another metasyncretism – it's a generalization about the interaction of gender

[8] A similar principle of fricative dissimilation may underlie the British English pronunciation of *sixth* as [sɪkθ], with deletion of the fricative /s/.

neutralization with 1st person across any and every potential exponent. Rather than saying that all six of these exponents simply fail to be specified for [± fem] (even though we see it's possible in the third person), it seems that there is a systematic and consistent constraint against the expression of gender in any combination with [+author].

(56) Constraint: [*+auth, +fem]. Repaired by deleting [± fem] on terminals shared with [+author]

Now, how do you say 'I am content' in Spanish? Well, either as *Yo estoy contento* or *Yo estoy contenta*. But how does the first person pronoun *yo*, which shows no gender distinction, trigger a different gender on the adjective? Where does the gender on the adjective come from? Well, it must come from the pronoun, but specifically, it must come from the pronoun *during the syntax*, as agreement happens during the syntax. In other words, we must presume that the first person pronoun *does enter the syntax* with the feature [± fem] from the formative list – there is no syntactic problem with having the combination of [+author] and [+feminine] in the syntax. The problem is *morphotactic*. The lack of gender on first singular pronouns is an instance of opacity. A feature f originates on a terminal A in the formative list, is copied over to a terminal B in the syntax, and is deleted from A in the morphotactics, remaining only on B:

(57) $[\ \text{f}\ ]_A\ [\ ]_B$
$[\ \text{f}\ ]_A\ [\ \text{f}\ ]_B$
$[\ ]_A\ [\ \text{f}\ ]_B$

Impoverishment rules can thus often be diagnosed as the result of a 'mismatch' between what is expected from the syntax and what actually happens: certain combinations are no problem in the syntax as the result of merging and agreeing formatives, but certain language-specific morphotactics do not tolerate these combinations, and enact deletion prior to Vocabulary Insertion, thus leading to less-specified exponents winning, or certain exponents never having a chance to win (the case of meta-syncretism at hand).

You might be thinking that the impoverishment rule in (56) only ensures that there will be neutralization of gender in the 1st person, but doesn't guarantee this for the second person. In fact, many languages disallow gender distinctions in *both* the first and second person. Such languages would have the impoverishment rule in (58):

(58) Constraint: [*+part, +fem]. Repaired by deleting [± fem] on terminals shared with [+participant]

However, languages such as Arabic that *do allow* gender distinctions in the second (and third) person, *but not the first*, would have the more limited impoverishment rule in (56). For English, therefore, the rule in (58) is more precise, whereas for Arabic, it would have to be the rule in (56), which does not affect [−author] pronouns.

What about 3rd person plural pronouns? In many languages with fully-fledged gender systems, there is a gender distinction in 3rd plural – for example Spanish *ellos* vs *ellas*. Other languages simply delete all gender distinctions *in the plural*, such as Russian. In Russian, the impoverishment rule would be:

(59) Constraint: [*−sg, +fem]. Repaired by deleting [± fem] on terminals shared with [−sg]

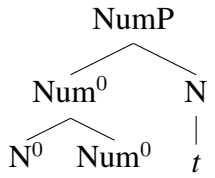
It is quite crosslinguistically common, therefore, for languages to lose gender distinctions either in the 1st person, in all [+participant] persons, and/or in nonsingular numbers. Finally, you might be wondering about English *they* – this too seems to adhere to (59), in losing gender distinctions in the plural.[9]

[9]However, there are some open questions about singular *they*, alluded to in Lecture 2 Notes – it too is decidedly genderless, and seems to be morphologically plural, as in *Ask your partner if they are happy*, with the 'plural' form of *are*. Nonetheless, this form is semantically a singular. The real question about this pronoun, therefore, is the nature of how it comes into the syntax as a formative. It seems to be [+sg] (or perhaps unspecified for number) and either [−fem,−masc] (or underspecified for gender) – and in any of these cases, the question of neutralization of gender in a rule like (59) would not arise.

13. On M-Words, Infixes, and Prefixes

Recall that we have discussed the formation of M-Words; complex items formed by head-movement. Thus when an N moves to Num, the result ends up as *dog-s*:

(60)



The semantics (and hence the syntax) of an expression like *dog-s* however, is one in which Num is clearly higher than N, as *dogs* represents 'a plurality of dog', i.e. the plurality or singularity is predicated of a Noun that must be first named. Mark Baker, in a series of papers in the 1980s, used the term 'Mirror Principle' to explain the fact that even though Num is higher than N in the syntax, in the linear order of morphemes, N comes before Num. Doing head-movement by left-adjunction (i.e. movement to the left) will automatically capture mirror-principle effects: the lower head moves to the left of the higher one, and forms and M-word with it, and as a result, the higher head, if it is affixal, becomes a suffix to what started out as the lower head.

We will proceed with the assumption that, just as movement of phrases (e.g. wh- expressions) is generally to the left, head-movement is also to the left. As a result, most cases of affixation will end up as suffixation, which seems to accord well with typological tendencies. Cutler, Hawkins and Gilligan (1985) call 'The Suffixing Preference': the fact that more languages have more suffixes than one might expect if prefixation and suffixation were equally likely options (or in fact, if they correlated with the phrasal syntax being head-initial or head-final). Cutler et. al chalk up the suffixing preference to extralinguistic factors, specifically processing and lexical access. Nonetheless, the immediate question that comes to mind is, so how is prefixation derived, for cases like the negative adjectival prefix *in-* in English?

No doubt you might be thinking that modifying the way that head-adjunction works could result in prefixation. But before even going down that road, we must take an excursion into *infixation*, a robust phenomenon especially in Semitic languages, and one which, upon inspection, no amount of fiddling with the directionality of head-movement can generate. Consider the following nouns and verbs from Modern Hebrew, all based on the triconsonantal root $\sqrt{sgr119}$

(61)
a. sagar 'to close'
b. hisgir 'to extradite'
c. histager 'to cocoon oneself'
d. seger 'closure'
e. sograyim 'parentheses'
f. misgeret 'frame'
g. misger 'to frame'

These are very much like words based on *-dict-* in English, all of which involve in some way the concept of 'saying', e.g. *dictate, contradict, predict, diction, addict* (the last of which is someone

who 'say yes towards things') Looking at all of these, they share something in common, both in terms of phonology (they all contain the same three consonants always in a set order, although with other elements intervening), and semantics (they all refer to 'closure' in one way or another). The first of these is simply a transitive verb, while the second can be viewed as a causative verb, something like 'to cause to be closed off from'. But unlike the words based on *-dict-*, these don't just involve prefixes and suffixes to the left and right, but more fundamentally, elements tucked in between the root – hence the term *nonconcatenative* morphology. This is illustrated for these particular words in the table below, in a schematic format for all but the last one, to which we will return:

(62)

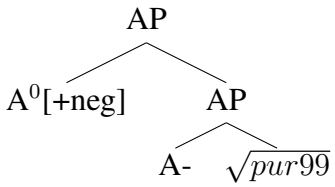
derivational category	infixes	prefixes and suffixes	combination
transitive *v*	a a	none	sagar
causative *v*	∅ i	hi-	hisgir
reflexive *v*	a e	hit-	hitsager (→ histager through later phonology)
result *n*	e e	none	seger
abstract *n*	o ∅	-ayim	sograyim
instrument *n*	∅ e	mis-	misger

Given that it is impossible to derive the infixation of vowels like *e e* inside a head via head adjunction, it looks like the right way to deal with infixation is to make it a property of the exponent list, and not a result of the syntax. That is to say, the exponent list must contain statements like the following:

(63) transitive *v* ↔ /a. . .a/, INFIX

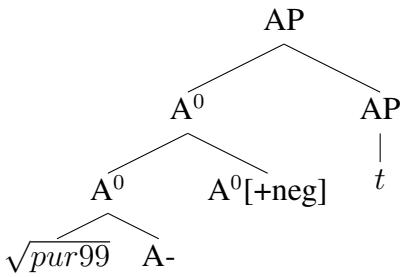
Once we know it is inevitable to include some specification of whether an affix is a prefix, suffix or infix in the exponent list, we can return to the question of prefixes in English. Rather than making them the result of alternative syntactic derivations, we might also include their directionality of affixation in the exponent list. Consider, for example, the negative adjectival prefix *im-*, which we will treat as an A- head (i.e. a head that projects the category adjective, in addition to other syntactic and semantic features it includes), in a word like *impure*:

(64)



As a result of two steps of head-movement, the structure will be:

(65)

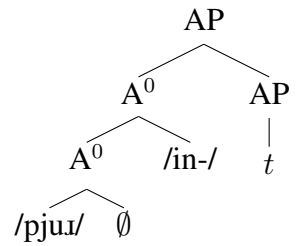


Now it's time to consult the exponent list, where the following entry is relevant:

(66) A^0[+neg] $\leftrightarrow$ /in-/, PREFIX

The result of inserting phonological material with this specification is that it will be added to the *left* side of its host within the M-Word. This can be schematically represented below, as the output of the relevant steps of Vocabulary Insertion:

(67)



The hyphen in the prefixal entry /in-/ indicates here that it will be linearly placed to the left of its host, and once nasal place assimilation happens in the phonology, the result is, voila, [impjuɪ]. Note that, given the suffixing preference, we will adopt the hypothesis that specification as INFIX or PREFIX needs to be indicated on the exponent list, while specification as suffix is a default, and need not be explicitly indicated in the exponent list.[10]

14. Root-Derived vs Category-Derived, the Encyclopedia, and Idioms

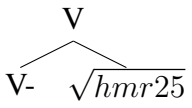
In the discussion above, we mentioned that $\sqrt{dict-}$ means something like 'saying' or 'speech', and that $\sqrt{sgr}$ means something like 'closure' or 'to close'. But these are already linguistically encoded, and the point about roots is that they are (a) categoryless, (b) unpronounceable in isolation and (c) express a linguistically-unpackaged concept. In this sense they are like Schrödinger's Cat – a thought experiment used in quantum physics, in which a cat, inside a box, is simultaneously alive and dead at the same time, a superposition of different states. But if you open the box, you see the cat *either* alive or dead, but not both. A root describes the situation of the closed box: $\sqrt{dict-}$ simultaneously bears the potential to be part of a noun or a verb, and indeed a variety

[10]Some cases of what seem to be prefixes may also result from them starting *lower* than their host. For example, there are reasons to think that the prefix *re-*, meaning 'to do X again' as in *reheat*, starts lower than the V to which it eventually incorporates. If this is correct, its movement up to V and head-adjunction would derive its prefixal status; in this sense, rather than saying *re-* is a prefix on *heat*, *heat* would be a 'suffix' on *re-*, in the sense that it is the host to which a left-adjoining lower element moves.

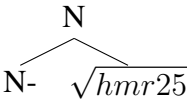
of different ones. But as soon as you try to say what $\sqrt{dict-}$ means, you've opened the box; you've put it into a noun or a verb, and hence you've immediately removed it from its potential of being simultaneously both. Roots have no category, and it is only when they combined with a 'categorizer', i.e. A-, N-, or V-, that they become one of the lexical categories adjective, noun, or verb.

Now, there are noun/verb pairs such as *a hammer* and *to hammer*. At first blush, one might think that *to hammer* means *to use a hammer*, but it doesn't. One can easily say *I hammered in the nail with a shoe*, and in this case, *to hammer* descibres a manner of an action (a specific repetitive motion acting upon an item, e.g. *John hammered at the door with both hands, hoping someone would hear his knocking*). In this case, what we'd want to say is that *a hammer* and *to hammer*, as noun and verb respectively, come from the same root:

(68)



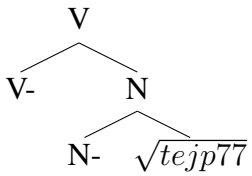
(69)



In other cases of noun/verb pairs (to which we will return), the differences are even clearer, as the phonology is not quite the same: *a record* is a specific vinyl medium for storing music, while *to record* is to capture sound onto any medium, usually *not* a record, these days, given the availability of digital media. Notably, the two have different stress patterns: the noun has stress on the initial syllable, while the verb has stress on the final syllable.

However, not all noun/verb pairs show this independence from one another. Verbs like *to tape* (involving the adhesive, and *not* cassettes!) seem to specifically mean 'to use tape'. Thus 'I taped the note to the wall' must involve some kind of tape, even if it is a band-aid (a kind of pink medicinal tape) or other adhesives. It cannot mean 'I stuck it to the wall', describing a manner independent of any specific instrument for sticking, thus *#I taped the note to the wall with glue/staples/spit* simply does not work, unlike *I hammered in the nail with my shoe*. This is because the verb *to tape* is *not* derived from the root $\sqrt{tejp77}$, but rather from the *noun* tape. In other words, *to tape* means 'to use tape (or something very much like tape) for a specific adhesive purpose', while *to hammer* does not necessarily mean 'to use a hammer'. The derivation of the verb 'to tape', then, necessarily *goes through* the noun-categorized version of the root first:

(70)



The difference between *to hammer* and *to tape*, then, is that the former is a *root-derived* verb, whereas the latter is a *category-derived* verb, in this case, a *noun-derived* verb. The verb *to hammer* is a result of forming a verb directly from the root, and hence need not involve the noun 'a hammer',

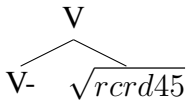
whereas the verb *to tape* is the result of forming a verb *from* the noun 'tape and hence must involve 'tape'. (Of course, an interesting question is *why* one cannot combine the root $\sqrt{tejp77}$ directly with a V-. We will treat this as a selectional restriction in the syntax, akin to the fact that one cannot say **I dined a pizza*. Certain verbs and objects simply cannot combine syntactically, as part of their syntactic restrictions).

There is phonological evidence for the difference between these root-derived noun/verb pairs (in which both are independently derived by combining a categorizer with a root) and verbs that are derived by going *through* an already-categorized noun. Consider the following:

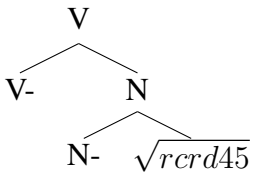
(71) a. a pérmit / to permít
b. a récord / to recórd
c. a dígest / to digést

It is clear that to permít someone to do something does not necessarily involve issuing a pérmit, to recórd something does not necessaritly involve putting in on a récord, and to digést something does not involve producing a dígest (in fact, a dígest usually refers to the collected / condensed works of a body of literature). Nonetheless, all of these nouns, as it turns out, *do* allow verbs to be formed from them, e.g. *to récord*, as in 'to use a record in order to do something', like hitting someone with a record – *Keep pestering me, and I'll record you over the head*, *to pérmit*, as in to issue permits, and *to dígest*, as in to create a digest. The structures of these two types of verbs is as follows, where (72) is *to recórd*, with a semantics in principle independent of the use of a record, while (73) is the *denominal* verb to récord, which necessarily involves the use of a record:

(72)



(73)

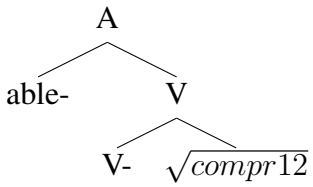


In other words, creating a category from a root allows a potentially idiosyncratic or unpredictable semantics and phonology, whereas creating a category from a noun (or anything already categorized for that matter) necessarily inherits the phonology and semantics of the already-formed category. Let's return to Modern Hebrew, where a parallel phenomenon is found, based on the noun *misgeret* 'a frame'. This can undergo denominal verb formation, the meaning 'to frame' (meaning to put a frame around something), and just like we have seen above for English, it inherits the phonology of the noun, and is thus *misger*, carrying the same prefix, same infixes, and same overall syllable structure of the existing noun, only truncating the last syllable (as transitive verbs stems must be maximally disyllabic). But the root does not enjoy the same potential for independence that it would if it combined directly with a V- head; instead, by combining with an N- head first, that already 'fixes' the meaning and phonology in a deterministic way.

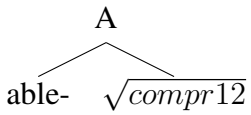
A final pair which illustrates the contrast between root-derived and category-derived forms

rather strikingly is adjective formation based on the suffix *-able*. This adjective-forming suffix is extremely productive when it attaches to verbs; thus *reachable by train, easily downloadable, very readable*, and, of current interest, *comparable*, meaning 'able to be compared'. Thus I can say, after a fishing trip, that anyone could tell that the size of the fish I caught and the one you caught were *comparable*, meaning they could be compared (and presumably the result of this comparison would be that one of us 'won' the competition). Nonetheless, there is another adjective, spelled the same way in English orthography but with a noticeably different stress pattern, namely *cómparable* (with stress on the initial syllable). This word actually does not mean 'able to be compared'; rather it means something like 'equal, equivalent'. Thus, while it is a contradiction to say "the size of these fish can be compared, and indeed they aren't compárable", it is not a contradiction to say "the size of those fish can be compared, and indeed, they aren't cómparable". The structures for the former (a deverbal, hence category-derived, adjective) and the latter (a root-derived adjective) are shown in (74) and (75) respectively:

(74)



(75)



Again, we see that while the former necessarily inherits the semantics (compárable must refer to the notion compáre) and phonology of the verb from which it is derived, the latter has an independent potential for the root to form an adjectival meaning and phonology.[11]

Stepping back from the comparison between *compárable* and *cómparable*, the key point to emphasize is that the latter bears a somewhat 'unpredictable' compositional meaning – if you know what 'compare' means, and what 'able' means, you won't necessarily know what *cómparable* means as a matter of simply combining the two. Rather, the root $\sqrt{compr12}$ combined with the suffix *-able* has what we might call an 'idiomatic' meaning – much like the meaning of *spill the beans*: if you know what *spill* and *beans* mean independently, it still might be unpredictable that the combination means 'to divulge information'. The point to take home here is that there is not really any difference between the idiomaticity of the root $\sqrt{compr12}$ combined with the suffix *-able* and that of *spill* and *beans* – both are idiomatic, in the sense that both involve a special *context-dependent* interpretation. We now introduce the List in which such context-dependent interpretations occur: the Encyclopedic List. Much the same as the Exponent List is the place in which *allomorphy* must be listed for context-dependent phonology, the Encyclopedic List is the place where *allosemy* must be listed for context-sensitive interpretation. We can demonstrate this for roots like *throw*, as follows:

[11]There is independent evidence for an *-able* that attaches to roots from words such as *questionable, fashionable, comfortable, profitable*, where there is no verb in question. Similarly, cognate suffixes in Italian such as *papabile*, based on the root for 'pope', mean something like 'eligible'.

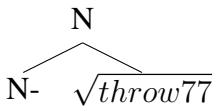
(76) a. $\sqrt{throw77} \leftrightarrow$ vomit / [___up]V
b. $\sqrt{throw77} \leftrightarrow$ light blanket [ ]N
c. $\sqrt{throw77} \leftrightarrow \lambda$ x, λ y, y throws' x / ___elsewhere

Notice that there is an 'elsewhere' interpretation on the LF side, the same way that there is on the PF side. What this means is, just like the plural suffix has a specialized allomorph *-en* in the context of a restricted list of roots, so does the root $\sqrt{throw77}$ have a specialized alloseme in the context of a restricted list of roots.[12]

The notation *throw'* is a stand-in for the language-independent representation of what 'throw' means, i.e. it would be a picture of a throwing event, the same way that *dog'* corresponds to whatever the language-independent meaning of 'dog' is, i.e. a picture of a dog, or a family-resemblance prototype style category that includes everything from greyhounds to poodles. It's important in this sense to say that even the meaning of Italian *gettare* would be *throw'*, rather than English 'throw', i.e. it truly would be the language-independent representation of the meaning of this lexical item.

Notice, of course, that *a throw* can refer to a blanket in (76-b) when it is the root directly combined with an N-, but refers to an event of throwing when it is an N- combined with an already categorized root (as in *that was a good throw*. These are depicted below:

(77)



[12]We may take the parallels between operations on the LF side and the PF side even further, and consider what the structural analogues to impoverishment, prior to consultation of the Exponent List, might be. Consider a sentence such as the following, which is ambiguous:

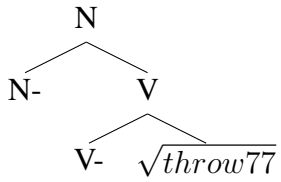
(i) Only I did my homework

One reading of the above sentence means that nobody else helped me on my assignment. But the more interesting reading is the one which means 'nobody else did their homework'. This meaning is more like the following:

(ii) Only I am an x, such that x did x's homework

Given this interpretation, it becomes clear that the *my* in (i) is not interpreted with its phi-features of [+author,−singular]; in fact, these phi-features are never interpreted on this pronoun on the LF branch. This suggests precisely the type of mismatch that implicates the involvement of a feature-changing operation – call it *deprivation* – which removes phi-features before the Encyclopedic List and further LF operations can have a chance to interpret them. Similar uses of 'deprivation' may be involved in the use of second person pronouns for impersonal uses (e.g. *In busy train stations, you should keep an eye on yourself*), which pattern entirely like ordinary second-person pronouns for syntax and PF, but in which the 'addressee' features go uninterpreted on the LF branch. In short, the parallelisms between the Maximal Subset Principle on the Exponent List and the Encyclopedic List, and the use of elsewhere items in both, go even further once we find that mismatch-yielding operations like Impoverishment on the PF branch have structural parallels on the LF branch as well.

(78)



From this perspective, we might adapt the title of a famous paper by Alec Marantz (1996) and say that in light of (77), "*throw* is a phrasal idiom". That is to say, given a distinction between roots and categorizers, even the word *throw* – itself in fact, a structural configuration that yields a phrase – can be idiomatically interpreted according to the encyclopedic entries.

Returning to the discussion of contextually-specific allosemic interpretations in even larger phrasal contexts, consider some of the patterning of idiomatic expressions like *spill the beans* vs. *kick the bucket* (where the latter means 'to expire a life' or perhaps more bluntly, to die). These differ in how they may undergo further syntactic transformations:

(79) a. #the kicking of the bucket, #the bucket was kicked
b. the spilling of the beans, the beans were spilled

We can model this in terms of the way that the Encyclopedic List includes a specification for active (not passive!) voice in the former, but not the latter:

(80) $\sqrt{kick88} \leftrightarrow$ expire / [voice$_{active}$ [___bucket]]

(81) $\sqrt{spill99} \leftrightarrow$ divulge / [___beans]

The fact that the allosemic interpretation necessarily includes an active voice head in the former makes this interpretation impossible when the object is passivized. (Interestingly, it may also block the possibility of an idiomatic reading under noun-verb compounding, e.g. *bucket-kicking*).

We may notice that in *spill the water* and *spill the beans*, the root $\sqrt{spill45}$ enters the syntax the same in both cases (coming from the formative list), 'plays' in the syntax the same way (as a transitive verb, assigning case and theta roles accordingly) in both cases, and is exactly the same on the exponent list and in the phonology as well. It is only on the encyclopedic list that the two diverge; in one case the verb acquires a special interpretation. It is for precisely this reason that we call the present theory one that is *distributed*, rather than bundling all of the lexical information about *spill* into a single 'list'. We've now seen cases of allomorphy, in which the same formative, in two syntactic configurations J and K, might be realized differently on the exponent list in J and K but the interpreted same in the encyclopedic list (e.g. the plural suffix 'means' the same whether it is exponed as /-z/ or /-en/), and of allosemy, in which they might be realized the same on the exponent list but interpreted differently in the encyclopedic list (e.g. *spill*). The potential of the same formative to vary on each of these lists independently is therefore justification for a model of grammar in which the exponent list and the encyclopedic list are independent and distinct entities that never interface with each other directly.

What about the beans and the bucket? They mean something like 'the secret' and 'the life' in these contexts:[13]

[13]There is also a recent coinage, *a bucket list*, which means something like a top-ten list of the things to accomplish

(82) $\sqrt{bean62} \leftrightarrow$ secret / [[−sg] [spill [the ___]]]

(83) $\sqrt{bucket22} \leftrightarrow$ life / [[−sg] [kick [the ___]]]

Interestingly, the non-idiomatic meaning of *spill the beans* is still available, but a common intuition – sometimes explored experimentally – is that it takes some 'work' to suppress the literal meaning; that using the non-specialized alloseme is like using the non-specialized allomorph and saying *oxes* or *goed*.

Now, some roots seem to *only* have an idiomatic meaning. Take *in cahoots*, which means something like 'a conspiracy'. However, the root *cahoot* doesn't occur anywhere but outside this context; you can't say *a cahoot* or even *out of cahoots*. This context-dependence of $\sqrt{cahoot44}$ on its syntactic environment is, by hypothesis, the same as morphemes like *cran-*, which never are found outside of the context of *berry*. Their formal representation on the Encyclopedic List would be as follows:

(84) $\sqrt{cahoot44} \leftrightarrow$ conspiracy / [in [___−sg]]
no elsewhere

Interestingly, the move of saying that on a given list, a root might have *only* a context-dependent interpretation, and no context-independent interpretation (i.e. no default / no elsewhere) is not unique to the Encyclopedic List (i.e. the LF side). It can also be found in certain cases on the PF side. There are a class of verbs in Spanish and Portuguese, for example, called 'defective' verbs, because neither dictionaries nor native speakers can produce any form at all for certain combinations of person, number, tense, and mood. One such example is the verb 'colorir', to color, which has the following set of forms in the present tense and infinitive:

(85)

infinitive	colorír
participle	colorído
1sg	no form, expected colóro
2sg	no form, expected colóres
3sg	no form, expected colóre
1pl	colorímos
2pl	colorís
3pl	no form, expected colóren

What this means is that in the present tense, there is no way for Portuguese speakers to say 'I colour the flower red'. (In other tenses which involve a stressed vowel after the stem, it is fine, but not, as it happens, in the present tense). The phonological pattern seems to be that this root can only be realized when it's followed by a stressed vowel; otherwise it simply has no phonological interpretation (or, putting it more clearly, no realization) outside of this very specifically defined context:

(86) $\sqrt{colr44} \leftrightarrow$ color / ___V́
no elsewhere

or places to visit during one's lifetime, which perhaps reveals an extension of the idiomatic reading of 'bucket' as life.

This entirely parallels the way that $\sqrt{cahoot44}$ has no semantic interpretation outside of a very specific, contextually-defined environment.

15. Impoverishment and Obliteration

Recall that Impoverishement refers to a process such that in a given context a feature is deleted. Say there is F, G, H; and H is deleted but the whole terminal node is still there.

(87) $[\pm \text{H}] \rightarrow \emptyset$ / ___[−G]

The terminal node on which all of these features were found, and on which feature-deletion happened, still undergoes vocabulary insetion, and still according to the maximal subset principle; it's just that the maximal subset principle will no longer take H into account. Thus any potential exponent which mentions H is no longer part of the competiton; it is *overspecified* for this terminal node given what's happened to the latter. Given exponents like *a b c* below, where *a* mentions FGH (be it, say, a plus or minus value of H), *b* mentions FG, and *c* mentions G, *a* will no longer be eligible, but *b* will still win over *c*:

(88) a ↔ [+F, −G, +H]
b ↔ [+F, −G]
c ↔ [−G]

Now there is another type of deletion which is, let us say, more dramatic. This one deletes *the entire node* in a given context. In such case, there is simply no vocabulary insertion at T. Thus under obliteration, none of *a, b, c* will be inserted.

Let's recall the Spanish *se*, which is used for impersonals and for reflexives. It has a close correlate in Italian, where it is pronounced as *si*, and has these same uses as well. Impersonals are the kinds of phrases one sees on signs or hears about what one should and shouldn't do (e.g. one shouldn't smoke here) – these are often rendered with *you* in English, e.g. *When in London, you should have fish and chips*. Reflexives are obligatory ways of specifying certain verbs that involve doing an action to oneself – and we'll focus here on verbs of grooming, (e.g. *Gianni dressed himself ('got dressed') this morning*), or to comb one's hair, to wash one's hands, etc. In Italian, these verbs obligatorily involve a reflexive, the same way that in English, verbs like *to perjure* (meaning 'to lie under oath') do – it is ungrammatical to say **Fred perjured*, one must say *Fred perjured himself*. This is shown below, where, as in many languages, inalienably possessed nouns (e.g. body parts that inherently have an owner) the definite article *the* is used instead of *his*.

(89) Gianni si lava le mani 'Gianni washes his hands'
Gianni REFL washes the hands

(90) Si mangia pizza con il coltello 'One eats pizza with a knife'
IMPERS eats pizza with the knife

We notice that although IMPERS and REFL have the same exponent (i.e. present a case of syncretism), they're really quite different syntactically, as IMPERS can be used as the subject of any verb (just as impersonal *you* can in English), while REFL is lexically restricted to specific verbs. It is a fact particular to Italian that these two uses have the same exponent; in German the impersonal is *Man*, while the reflexive is something like *sich*.

Now under certain circumstances, one might want to use a verb phrase that contains both an impersonal and a reflexive. For example, to say that "One washes one's hands every day", there would be an impersonal subject, as well as a reflexive for the verb-of-grooming. We would expect something like the following, which is ungrammatical:

(91) *si si lava le mani ogni giorno Intended: 'One washes one's hands every day'
IMPERS REFL wash the hands every day

In fact what arises instead is a mismatch, or an unexpected phenomenon, as follows:

(92) ci si lava le mani ogni giorno 'One washes one's hands every day'
ELSW REFL wash the hands every day

What's happening is that the elsewhere item, *ci*, is unexpectedly inserted in place of the first *si*. In Italian, this *ci* is used for 1st person plural objects, and also for locatives! The locative clitic is like in French, where the clitic *y* means 'there':

(93) John y va 'John goes there'
John LOC goes

It is interesting to have a pronoun that refers to a kind of location, and these locative clitics lean on the verb in exactly the same way as pronouns that refer to people. In Italian, there are examples as follows:

(94) Ci vado 'I go there'
LOC go.1sg

Clearly, the LOC usage and the 1sg usage have nothing in common – this exponent is clearly an elsewhere item. It is exactly for this reason that *ci* shows up in (92). It is a clitic of a thousand faces, and hence steps in to do the job once the *si* is removed in (92). Now we know that this is a case of impoverishment, and not obliteration, because a position of exponence still survives. In other words, in the (92) configuration, the features referring to 3rd person on the first clitic are deleted, rendering *si* ineligible for insertion.

(95) /lo/ $\leftrightarrow$/ [+pron, −auth,−part, −nom]
/si/ $\leftrightarrow$/ [+pron, −auth,−part]
/ci/ $\leftrightarrow$/ [+pron]

Notably, while in Spanish the elsewhere clitic is *se*, in Italian it is *ci*. Different languages may end up with different elsewhere items in their inventory. Now it is an interesting question why, in the configuration (92) it is the first one that undergoes impoverishment, as opposed to, say, the second one. This is a generalization that often happens across Romance language clitic clusters – in a banned configuration, the repair often affects the item on the left, not the right. (In fact, parallel questions arise with dissimilation and assimilation in phonology – why it tends to be regressive or progressive in nasal + stop sequences, where /in+pure/ becomes [impure] and not *[inture].)

We can combine locative clitics with other clitics:

(96) Gianni ti ci porta 'Gianni brings you there'
Gianni you LOC brings

But importantly, not with the 1sg clitic, which is also *ci*:

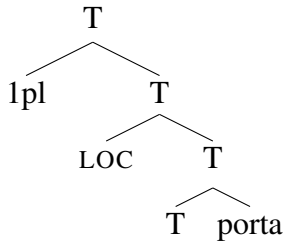
(97) *Gianni ci ci porta Intended: 'Gianni brings us there'
Gianni us LOC brings

Now, normally, when a clitic can't be used, the elsewhere item steps in. But the target that would be impoverished here to give way to the elsewhere item already *is* the elsewhere item! As a result, the whole node must be deleted. In the examples, therefore, it is inappropriate to place a $\emptyset$ node in place of the 1pl. clitic node – instead, the proposal is that the whole node is removed from the tree, in such a way that it never undergoes vocabulary insertion in the first place.

(98) Gianni ci porta Intended: 'Gianni brings us there'
Gianni LOC brings

Again, we return to the question of why the first clitic is the one obliterated. Consider the structure below, in which the two pronominal clitics are hierarchically adjoined to the tensed verb:

(99)



We haven't discussed this before, but let us suppose that Vocabulary Insertion proceeds from bottom-to-top in a tree structure; that is, it starts with the most deeply embedded node N, does Vocabulary Insertion on that node, then proceeds to the next lowest node M, and does Vocabulary Insertion on that node, then to the next node up, say L, and does Vocabulary Insertion on that node. This is called a *cyclic* procedure because it happens one node at a time, as opposed to say all at once. Now, why is the first clitic affected? On this model, we can simply say, because it is higher. The lower clitic can be turned into *ci* by Vocabulary Insertion, but then the later, higher-up clitic is affected by obliteration before Vocabulary Insertion can affect it. (Given a cyclic, bottom-up model, the alternative repair – which doesn't happen – in which the lower node is deleted without the higher node having undergone Vocabulary Insertion yet makes it difficult for the lower clitic to 'know' that the higher one would later cause the problem).

(100) $[+\text{pron}, +\text{auth}, -\text{sg}] \rightarrow \emptyset$ / ___ci

This is a very derivational model, akin to chess – you cannot do castling until you've moved the bishop out of the way, and the higher node. But if the lower node undergoes Vocabulary insertion first, its phonological form can condition obliteration of the higher node in a straightforward

manner.

Let's conclude in comparing Impoverishment versus Obliteration. The two can largely be detected in that the former paves the way for insertion of an elsewhere item, but the latter removes a node altogether. Interestingly, sometimes the two can be hard to tell apart — namely when the elsewhere item can independently be shown to be a zero (phonetically null) item. However, even in these cases, we can diagnose the difference when the presence or absence of such a node would condition allomorphy for *another* terminal node.

16. Allomorphy and Containment

Recall that suppletion is root allomorphy, and the more-specified form should win. Hence comparative forms that fail to use the more specified allomorph, as follows, are ill-formed:

(101) *This is gooder than that

For this case, we could analyze the allomorphy as follows:

(102) $\sqrt{good31} \leftrightarrow$ bɛt / ___CMPR

That is to say, the *-er* on *better* is the same as the ordinary comparative suffix, and *-est* as the superlative.

Interestingly, the comparative in English can be expressed either by the suffix *-er* (sometimes called the 'synthetic' form) or by the freestanding item *more* (sometimes called the 'analytic' form). This is also a case of allomorphy, conditioned by the prosodic shape of the stem (roughly related to how long it is and where the stress falls; it is basically limited to a trochaic foot, with a few lexical exceptions such as **funner*). In fact, we might even say that in English, when the *-er/-est* cannot sit on adjectives which exceed this size, this forces *mo-* support, yielding *more/most*. However, our focus in the present discussion is on the allomorphy of the root itself, not the comparative suffix.

If indeed we analyze *better* as /bɛt/ plus the comparative *-er*, this in turn leads to analyzing the superlative as the result of *bet-st*, with some morphophonology probably turning this into the surface form *best*, and similarly for /wəɹ/ which turns into *worst*. An even clearer illustration of this same root allomorphy in both the comparative and superlative can be found in the related Germanic language Danish, which has the following set of forms:

(103) god, bedre, bedst

The fact that the same allomorph is used in the comparative and the superlative leads to what Bobaljik (2012), in his comprehensive analysis of the morphology of comparative adjectives, calls the ABB pattern. The ABB pattern can also be found in Czech forms of the root for $\sqrt{bad66}$:

(104) špatn-ý, hor-ší, nej-hor-ší

Now, all three of these forms can potentially show different allomorphs, as in Latin:

(105) bonus, melior, optimus

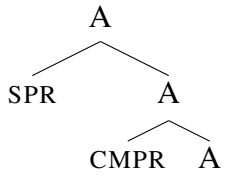
This would exemplify the ABC pattern. Crucially, Bobaljik, in an extensive typological survey, notes that while there is an ABB pattern as in Czech, an ABC pattern as in Latin, and an AAA pattern for cases with no allomorphy at all across the three contexts, there is one logically possible pattern that we don't find – namely, the ABA pattern, which would be something like:

(106) a. *good - better - goodest
b. *špatn-ý, hor-ší, nej-*špatn-vsí

c. *bonus - melior - bonimus

How can the absence of this kind of pattern be derived? The answer comes from a commitment to the syntactic structure:

(107)



This structure depends very much on thinking about the semantics of a superlative. *y is the tallest* doesn't involve a superlative combining directly with *tall*, but rather something like 'for all x, y is taller than x'. In other words, they necessarily involve a comparative as a subpart of their meaning. Speaking formally, (107) means that the superlative *contains* the comparative as a subtree. In fact, this points to a more general conclusion: that morphological analysis requires some convergence with what semanticists say that the compositionality of certain formatives must involve.

Focusing on the root itself, exponents referring to a special allomorph for [CMPR A] will be used in the superlative as well:

(108) $\sqrt{good31} \leftrightarrow$ bɛt / ___CMPR
$\sqrt{good31} \leftrightarrow$ gʊd

There is no way to generate the structure where the A combines directly with the SPRL, and hence no way to state a rule such as the following without reference to the intervening node. In other words, once an allomorph like (108) exists, there will be no way not to use it in the superlative, since the A will *still* be next to CMPR! For this reason, a rule such as (109) would be useless, assuming the containment structure, as A would still be next to CMPR.

(109) Putative superlative allomorphy rule:
$\sqrt{good31} \leftrightarrow$ bʊg / ___SPRL

The *ABA generalization comes from the containment structure assumed above, the fact that allomorphy cannot 'skip' conditioning environments and hence cannot have illegal rules such as (109), and the fact that there cannot be a direct combination of a superlative formative with an adjective. In fact this latter claim has the potential to relate to facts about acquisition, where presumably acquiring the superlative necessarily requires acquiring the comparative first. The important point is that the *ABA ban is actually not the result of the exponent list, but about the fact that the syntax never gives an opportunity to have the Adjective and the SPR as direct sisters.

The *ABA generalization is about patterns of allomorphy across the roots, and not about 'whole forms', or even the suffixes. Of course, when it comes to the exponent of these suffixal formatives themselves, very few languages show both *-er* and *-est* alongside each other in the superlative (or *-ior* and *-imus*), so something must ensure that the CMPR morpheme itself is rarely spelled out in the environment of a superlative (although we do see both affixes co-occurring overtly in cases like Czech above). There is allomorphy of the CMPR morpheme itself, visible in forms like *worse*,

where it is a $\emptyset$, although not in the superlative *worst*.[14]

How about deriving the ABC pattern? The limits of allomorphic context allow reference to the entire surrounding structure, as long as nothing is skipped:

(110) $\sqrt{good31} \leftrightarrow$ opt / ___CMPR SPRL
$\sqrt{good31} \leftrightarrow$ mel / ___CMPR
$\sqrt{good31} \leftrightarrow$ bon

In order to avoid a clever way of making the "C" of the ABC identical with the A, and thereby sneaking in "ABA", Bobaljik proposes the following extremely reasonable additional principle:

(111) Learners avoid positing a contextual allomorph of a morpheme μ that is homophonous with the default exponent of μ.

To conclude, *ABA is ruled out by the principle that a trigger for suppletive allomorphy for a root (whether or not it be a span of two such heads) must be adjacent to the root it conditions. Now, it's important to point out that suppletive allomorphy for comparative forms of adjectives tend to occur more with certain adjectives (e.g. *good, bad* rather than others), but the theory doesn't rule out allomorphy for adjectives like *steep*; it is arguably instead a fact about learning Exponents that in order to consistently acquire allomorphy, the adjectives in question tend to be those which are more frequently uttered and heard.

Now, what about *AAB patterns? These are empirically rare as well (although apparently not quite as rare as ABA). Bobaljik presents two potential ways to rule these out, neither of which follow from the basics of the theory presented above. Donegani (2012) thus conducted an experiment based on the notion of a 'wug test' (although supplying two different wugs and asking for a third) to look into this question, with the following design, using pictures for comparisons:

(112) Superlative Presented, Comparative Held Out: See how zil this N is!
But this one's the wobbest by far.
This one isn't ___than that one.

(113) Comparative Presented, Superlative Held Out: See how zil this N is! This one is wobber than that one. But this ones the ___by far.

Donegani's results showed 84% adherence to *ABA (i.e., in (113), 'wobbest', not 'zillest'). However, he found only 60% adherence to *AAB (i.e. 'ziller', not 'wobber').[15] This suggests, in concord with our discussion above, that *ABA is really the pattern the theory should work harder to rule out. As a further methodological point, it shows the interest of experiments in order to

[14]Interestingly, however, the *ABA generalization also seems to regulate the possibillities of analytic vs. synthetic forms. Thus, a language could not have a pattern with a synthetic comparative, but an analytic superlative, as in a pathological *ABA example *heavy; heaviest; most heavy / *heavier*. Deriving this fact of course depends on how one formalizes the synthetic pattern – whether it be via allomorphy, via head-movement, or via another kind of morphological displacement we will explore later.

[15]An interesting followup would be to show *only* the comparative (as form B) and the superlative (as form C), thereby forcing participants to supply the basic form of the adjective as either BBC (e.g. AAB) or CBC (e.g. ABA). The prediction is that the latter would be less attested.

complement typological research.

17. Suffixes that are Themselves Roots

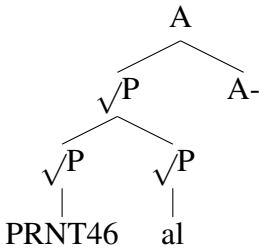
We have discussed the 'blocking effect' of morphemes such as SEP in *persons/people*, CAUS in *shined/shone* or *lie/lay*, and EVAL in *badder/worse* as cases of 'retroregularization', where the locality of specialized contextual allomorphy is blocked, and the root's realization reverts to the regular/elsewhere form. All of this suggests that roots can undergo late insertion too. So now we can begin to consider the possibility that roots are not defined by phonological constancy, but rather by acategoriality. Consider the following suffixes:

(114)

Exponent	Noun	Adjective
al	mammal	normal
an	librarian	reptilian
ant	defendant	defiant
ary	functionary	legendary
ate	consulate	intricate
ic	tunic	magic
ive	incentive	auditive
ory	promontory	rotatory
esque	arabesque	grotesque
y	parsimony	airy
en	warden	golden
ful	handful	colorful

In Lowenstamm (2010), it was proposed that the Distributed Morphology distinction between roots and categorizers should be further refined: rather than the traditional view in which a word like *parental* contains the suffix *-al*, which is an adjectivalizing (or, for short, A-) head, Lowenstamm's proposal is that in fact, *-al* is itself a root, and that the job of turning the structure into an adjective is achieved by another, silent head above these two:

(115) *parental*:

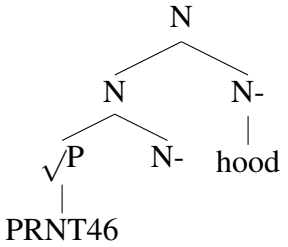


In the structure above, *parental* is never 'turned into a noun' (e.g. *parent*) before then becoming a denominal adjective; rather, it is turned into an adjective well after the addition of the suffixal root *-al*. Under this view, *-al* is a kind of bound root, perhaps no different from the infamous *cran-* found in ericaceous compounds. The reason for claiming that the suffix *-al* is a bound root is that it does not seem to be an A- head – as nouns like *mammal* can be nouns as well. Let us call such suffixes 'ambicategorial', as they form adjectives or nouns (though not verbs, but this may be

related to the properties of the root *-al* or how it combines with V- heads). In reality, the proposal is that, as roots, they have no category at all. Thus, suffixes like *-ant* in *defendant* (noun) and *defiant* (adjective) are acategorial roots.

It is important to contrast suffixes like *-al* with suffixes like *-hood*. The latter always creates nouns. As such, we depart from one detail of Lowenstamm's original proposal, and side with Creemers, Don, & Fenger (2014), who contend that for suffixes such as *-hood*, it makes no sense to separate the root content from the category-fixing head:

(116) *parenthood*:



Let us call suffixes such as *-hood* LP-selecting heads, as their complement is an already-categorized lexical category (e.g. N-,A-,V-), while suffixes such as *-al* are RootP-selecting heads, as their complement is a root phrase. What else does such a proposal achieve, besides providing a more accurate representation of the fact that *-al* does not seem to be an adjectivalizer, after all? The important leverage of this theory comes from the following principle:

(117) a. The complement of an LP-selecting head (call it LPSH) cannot have its stress affected by the addition of this LPSH
b. The complement of a RootP-selecting head (call it RPSH) *can* potentially have its stress affected by the addition of this RPSH

The principle in (117) is a syntactic implementation of the intuition underlying the distinction between Level 1 and Level 2 (also known as root-level and word-level) suffixes in classical Lexical Phonology (or between + and # boundary symbols in earlier generative phonology, which helped raise the question of how their phonology correlates with aspects of their morphology). Word-level suffixes, on the current view, attach to *words* – that is, something with a category. However, there is the added advantage beyond Levels 1 and 2 that (117) is based on an overarching theory of syntactic categories that specifically correlates the stress-changing possibility of the *-al* in *paréntal* with its ambicategorial (or in fact, acategorial) status. Thus, *párenthood*, with the LP-selecting suffix, necessarily inherits the stress of *párent*, while *paréntal*, with the RP-selecting suffix, does not.

Similarly, the suffix *-ian* is a root under this analysis, as it creates both adjectives *reptílian* and nouns *librárian*, and on both of these it has the possibility to alter the stress from what the corresponding nominal formation would be. In fact, this very same *-ian* is a root used in the formation of demonyms, and in the case of *-ian*, the demonym can be either an adjective or a noun:

(118) Maria works in a Panamanian restaurant, and in fact, she is a Panamanian by birth.

Not all demonym-forming suffixes are ambicategorial, as the following examples show (Incidentally, language names always follow the adjectival form, no matter what the nominal form is: *I speak Italian, Scottish, Turkish*, and we put these aside in what follows.)

(119) a. Maria is a Spaniard by birth, and works in *a Spaniard restaurant
b. Maria works in a Scottish restaurant and is *a Scottish by birth

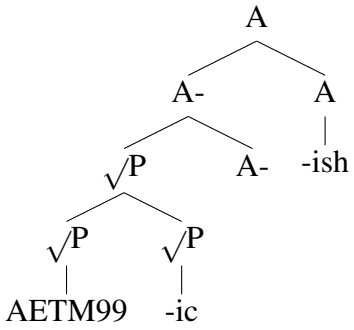
These patterns fall into place once we classify the denonymic suffixes as follows: *-ian* is a root, *-ard* is an N- head, and *-ish* is an A- head.

We can return to some of the contrasts in stress-shift (and indeed, other potential morphophonological changes, such as palatalization) in comparing the suffixes *-ity* and *-ish*. Notice that *atom* has penultimate stress, as does *atomic*, as does *atomicity* (if, as in general in English, final *-y* is unstressable). It is as if, for these cases, stress is reassigned with the addition of each suffix, but in the case of *-ish*, this does not happen.

(120) a. atomic-ish [ərɑ́mɪkɪʃ]
b. atomic-ity [ærəmísɪɾi]

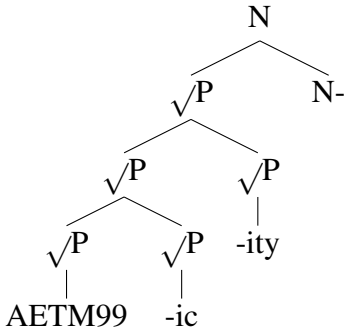
These correspond to very different structures. The suffix *-ish* is LP-selecting, and never changes stress:

(121) *atomicish*:



The suffix *-ity* is RootP-selecting, and in fact its complement *-ic* is still categoryless (cf. *magic, comic*, and other ambicategorial formations). It is only later turned into an N-.

(122) *atomicity*:

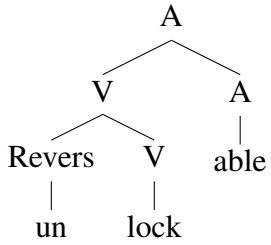


In other words, stress-shift is a misnomer. In words like *atomicity*, stress-assignment is *delayed*. Stress is never assigned until getting to the first category-fixing head. With this in mind, we can now turn to the explanation of the difference between *curiosity* and **gloriosity*. At first blush, the ill-formedness of the latter is surprising, as *curious* and *glorious* are both adjectives. However, recall that in what is developed above, *-ity* is root-selecting. This in turn reveals that *curious* is a root ($\sqrt{}$CUR99), while *glorious* is not – it is the combination of $\sqrt{}$GLOR75 with an N- head (forming *glory* and then an A- head *-ous*. As a result, while $\sqrt{}$CUR99 can combine with $\sqrt{}$ITY22, the already-categorized *glorious* must combine with an LP-selecting *-ness*.

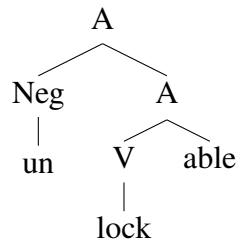
18. Combinations of Prefix and Suffix

What happens when a prefix and a suffix are combined onto a single root? In fact, considering the details of certain cases reveals that the internal hierarchical structure of an M-word may be different, even when the surface linear order is the same. Consider the word *unlockable*, which is ambiguous between two distinct interpretations. One of them – let's call it the 'just closed the trunk' scenario, happens when I close the trunk of my car and you say – hey wait, I need something out of there. I say "No problem, it's unlockable", which means that the trunk of the car is 'able to be unlocked'. In the other, distinct 'rusty bike lock' scenario, you and I are going to leave our bikes somewhere, but I say that my bike lock is rusty and doesn't actually work – that at present 'my bike chain is unlockable', meaning that it is 'not able to be locked'. These two structures are shown in turn below, with the internal structure of *lock* as a root plus categorizer omitted for simplification:

(123)



(124)

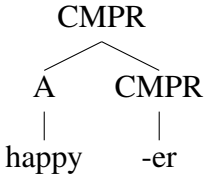


In the first structure, the reversative formative (applying to verbs that are reversible, and hence **to unfry an egg*) combines with the verb, and the structure remains a verb, then categorized as an adjective by the suffix *-able*. In the second structure, the verb plus *able* are first combined to form an adjective, and a distinct formative for negation of adjectives (which happens to have an identical exponent /ʌn/ in English, though is distinct in many other languages) combines to negate the adjective.

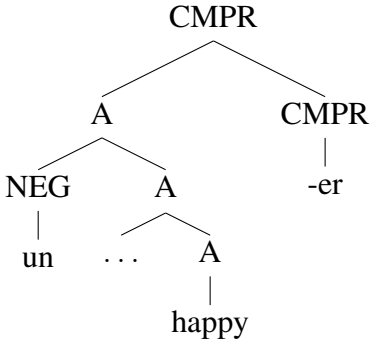
19. Bracketing Paradoxes

We now turn to one of the interactions between derivational affixes that instantiate what are known as bracketing paradoxes. The canonical case to consider is the word *unhappier*, and here's why. The compositional semantics of this word have a meaning of 'more unhappy', e.g. with a syntactic and semantic scope whereby [CMPR [NEG happy']] at LF. However, the morphophonology of this word involve the comparative allomorph *-er*, which we know is limited to disyllabic adjectives (e.g. *prettier,*beautifuller*). From this perspective, it would seem that the allomorphy is determined given a structure at PF like [NEG [CMPR happy']]. Hence the apparent paradox: given the Y-model of syntax and the fact that the same structure must be delivered to both PF and LF, how can the syntax generate something that has the CMPR formative *both* closer and farther to the adjective at the same time? The answer to this question is in fact to say that 'at the same time' is the wrong framework, and that the very notion of a derivational, step-by-step syntax allows for structures whereby the CMPR formative is closer to the adjective first, at which point it is close enough to determine allomorphy as suffixal *-er*, and then at a *later* step, it is *farther* from the adjective, whereby the NEG morpheme has become sandwiched between it and the adjective:

(125) Structure at an earlier point in the derivation, when CMPR and *happy* are local to each other:



(126) Structure at a later point in the derivation, when CMPR is farther away from *happy* than NEG:



In other words, given 'final' representations like [CMPR [NEG happy']] versus[NEG [CMPR happy']], this would look like a bracketing paradox in space, it is in fact no bracketing paradox at all 'in time': these simply represent different stages of the derivation. Well, what operations are responsible for the transformation that occurs in the steps between (125) and (126)? There are two answers in the literature that are both satisfactory and in fact, both of which crucially employ operations already known and independently motivated from phrasal syntax, long before their application to bracketing paradoxes within morphology.

The first of these is to say, as Pesetsky (1985) did, that the CMPR morpheme undergoes *covert movement* above NEG to get scope over it, in a way parallel to what happens in a clause such as two quantifiers: in *Some student has read every book here* can have a reading in which ∀ x, book' (x), ∃ y, student' (y), such that x read y. This quantifier raising involves the phrase *every book* undergoing what is called covert movement to a position where it scopes over *some student*, even though it is pronounced in its lower position. Pesetsky proposed that this kind of quantifier raising can happen for CMPR to generate a structure like the one in (126). The key idea, therefore, is that a derivation which allows merging something low and then interpreting it somewhere higher can be applied to individual formatives as well as phrases. Summarizing, the covert movement approach says that (125) becomes (126) by moving CMPR.

The second possible implementation of the transformation from (125) to (126), for which we will go into a bit more detail, involves the application of an operation called 'Late Adjunction', developed in syntactic theory, to the placement of *un-*. This approach therefore implicates an operation on NEG as the crucial way of disrupting the sisterhood between CMPR and *happy*. In this approach, developed by Newell (2005), a countercyclic operation of tucking in an adjunct phrase to an already merged item in the tree provides a way of expanding the structure not at the top of the tree, but at a lower position that has already been merged. The original motivation for such an operation comes from evidence that Lebeaux (1988) provided that certain adjunct relative clauses do not seem to count for what would otherwise be Binding Condition C violations. Consider for example the following sentence, which shows that when a pronoun that is co-indexed with the R-expression *Rachel* c-commands the latter (i.e. the two are coreferent), the sentence is ungrammatical.

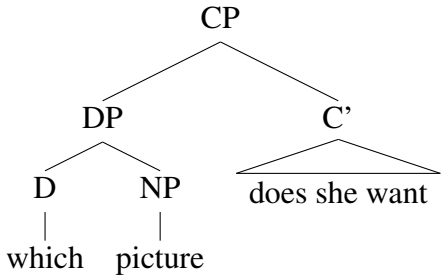
(127) *She$_r$ wants the picture that Rachel$_r$ likes

Now, interestingly, however, doing what looks like simple *wh-* movement of the object noun phrase doesn't end up creating a problem:

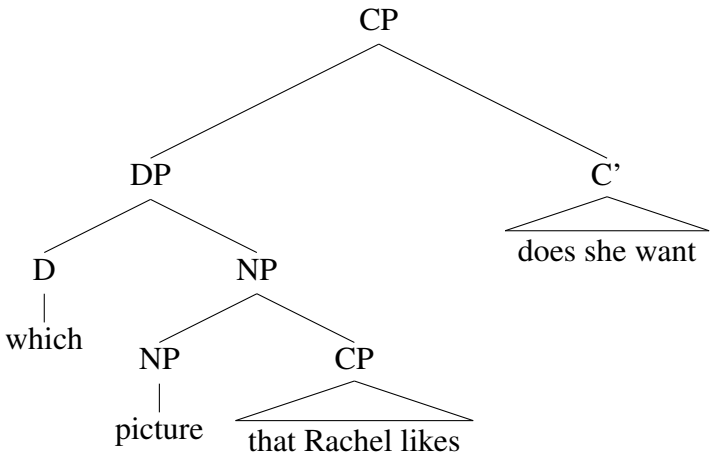
(128) Which picture that Rachel$_r$ likes does she$_r$ want _?

This is surprising on the view that (128) comes from an underlying structure like (127). But Nissenbaum (2000) argued that it doesn't – rather, that (128) comes from an underlying structure that simply involves *which picture* as the base-generated form of the object noun phrase (and hence the form it has when it moves). According to Nissenbaum, what happens thereafter is that the adjunct *that Rachel likes* is late-merged to the wh- phrase in a countercyclic method. Thus (129) becomes (130):

(129)

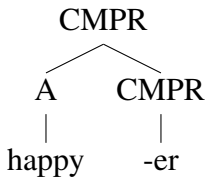


(130)



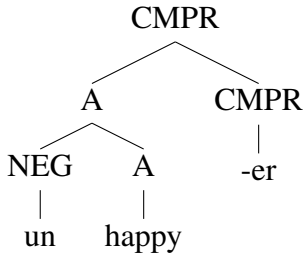
The main idea is that adjuncts (which preserve the category they added to as a projection of, and are not arguments) can 'tuck in' to the structure below the root of the tree. With this in mind, let us return to the morphological case at hand. In *unhappier*, NEG is a prefix that does not change the category of what it attaches to. According to Newell, therefore, it can tuck in countercyclically, not necessarily at the root of the tree. Specifically, assume that the *happy* and CMPR have merged, and that the structure as a whole undergoes Vocabulary Insertion, thereby determining the allomorphic shape of CMPR:

(131)



Subsequenly, the formative NEG late-adjoins to the adjective itself, below the comparative morpheme:

(132)



Under this model, late-adjunction of an adjunct will then add to the projection of A, and the structure as a whole will be eventually interpreted again at PF and LF. The item *-er* has already been inserted, so that cannot change (the bisyllabic-or-less requirement has already derivationally been met), but the fact that at LF the addition of *un-* has flipped the scale of comparison (from more' along a scale of happy' to more along a scale of unhappy') can take place, the same way that late adjunction of the relative clause in (130) adds to the interpretation of the object noun phrase, but does not retroactively change the status of a Binding Theory requirement that has already deriva-

tionally been done with.

Returning to the main thread of discussion, therefore, there are two ways to accomplish the transformational difference between (125) and (126). One involves moving CMPR later, so that it is higher than NEG, and the other involves late-sandwiching of NEG between the scope of CMPR and the adjective. Both draw on existing mechanisms employed in derivational phrasal syntax, and present solutions to the bracketing paradox whereby a series of ordered operations, in which requirements on formatives can be checked or evaluated early, and then the representation can subsequently change, so that the two 'sides' of the paradox in fact refer to distinct derivational moments.

20. Many-to-One Relations: Fusion

There are cases in which exponence does not line up with what we otherwise conclude the underlying syntax of a word to be, and by one-to-many and many-to-one relations, we refer to cases in which a single syntactic terminal apparently corresponds to two distinct positions-of-exponence within the same word, or when two syntactic terminals that we know are underlying distinct positions are realized by a single exponent. These two cases have been given the descriptive names of 'fission' and 'fusion', respectively, and we will start by illustrating fusion, which generates what are called 'portmanteau' morphemes that apparently seem to include exponence of two distinct syntactic positions. To see an example of this, consider the way that subject agreement and negation are expressed in Swahili. First, note that in (133) and (133), it can be seen that subject agreement (*tu* for the 1pl) is distinct from the negation marker, which is generally *ha* across all persons and numbers.

(133) tu- ta- pend-a kiswahili
we- will- love Swahili

(134) ha- tu- ta- pend-a kiswahili
neg- we- will- love Swahili

The first person singular marker, *ni-*, occurs in a position before tense that corresponds to where we expect it to.

(135) ni- tu- pend-a kiswahili
I- will- love Swahili

However, counter to expectations, the combination of *ha* 'negation' and *ni* '1sg.' is not employed:

(136) *ha- ni- ta- pend-a kiswahili
neg- I- will- love Swahili

Instead, the 'portmanteau' prefix *si* is used, in place of either the *ha* or the *ni*:

(137) si- ta- pend-a kiswahili
not.I- will- love Swahili

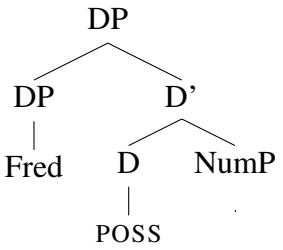
This case is superficially analyzable as 'fusion', therefore, because it looks like one can say that the *si* is a specialized exponent that somehow is realizing both subject agreement and negation at the same time. Indeed, such a case might be thought of as analogous to what happens in vowel coalescence sometimes, whereby when there are two adjacent vowels that cannot be realized, a repair involves a single vowel that is a mixture of the features of the two – say, one word ends with an /a/ and the next word begins with an /u/, but what is pronounced is a single [o] that represents a combination of the two underlying vowels. However, it is not altogether clear that (137) *must* be treated as a case of fusion. One could in fact say that negation has a contextually-specific allomorph *si* in the context of 1sg, and that 1sg has a contextually-specific allomorph $\emptyset$ in the context of negation. The appeal to 'portmanteau' morphemes often occurs in morphological theories which

attempt to deny the existence of zero morphemes or deny the existence of contextually-sensitive allomorphy – both of which are untenable assumptions, and hence the argument that elements like *si* correspond to two syntactic terminals – as opposed to corresponding to only one, with the other being realized by a ∅ – need to be formulated as logically or conceptually inescapable arguments in order to force the conclusion that fusion is needed as an operation. In other words, putative cases of fusion can often be treated as instead involving one null allomorph and a muturally conditioned allomorph.

Implicit in the decision to even think that (137) is a matter of postsyntax, rather than syntax, comes from the intuition that it is unlikely that there is a different syntax underlying, say, the combination of 1sg with negation than that of 1pl with negation. A similar kind of case can be raised with the 'portmanteau' forms of prepositions plus articles found in French, where the preposition *de* when combined with what would correspond to the masculine definite article *le* is instead realized as a single morpheme *du*. While it may appear initially tempting to say that this involves syntactic head movement to the preposition P by the D element in its complement, no such head movement occurs with the feminine definite article, where the preposition plus article are simply realized as the sequence *de la*. Treating the difference between the masculine and feminine definite articles in terms of whether the syntax involves head movement of the former but not the latter would seem to be a very parochial type of move for the syntax to make, and given the general heuristic that syntactic operations are more principled and universal while postsyntactic operations reflect specific morphotactic requirements and the irreducibly arbitrary differences in exponence that comprise crosslinguistic variation, the French cases also seem to be more gracefully handled in terms of mutually contextually-specific allomorphs for the P and D in question.

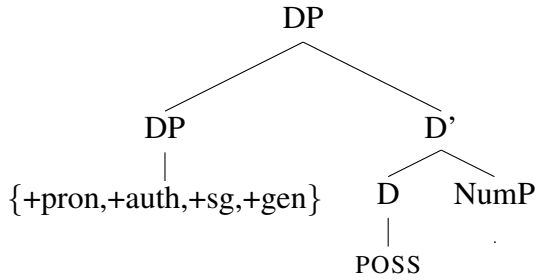
A final potential case of fusion involve the possessive pronouns in English, whereby the combination of a pronominal possessor with what is usually *'s* is instead realized by a single overt element *my*. Consider the syntax underlying possessives. In (138), the D element indicated as Poss will be realized by the exponent /-z/.

(138)



However, when the specifier is [+pronominal], POSS is realized as ∅:

(139)



In fact, a null realization of POSS is independently available in English – it is found in African-American English varieties where (138) is realized with no overt *'s* (Rickford 1999), and is found across a wide range of English varieties when the specifier is already pluralized, as in *The cats' food is over there*, which is pronounced [kaets] and not [kaetsəz]. Thus, to treat *my* as a portmanteau resulting from fusion would seem to ignore the fact that the POSS head itself may in fact have a null allomorph employed in these cases.

In sum, there are not many knockdown arguments for positing a special postsyntactic operation of fusion that literally takes two distinct syntactic terminals and turns them into only one, after the syntax is completed but before vocabulary insertion.[16] Many-to-one exponence can be dealt with by appeal to zero exponents, as long as the particular analyses are consistent and, as in the cases discussed above, ideally draw independent support from other cases of null allomorphs in the language. One-to-many exponence, however, clearly cannot be dealt with in terms of null allomorphs, and such cases require an implementation of the notion of fission. It is to these cases that we now turn.

21. One-to-Many Relations: Extended Exponence

In the case of fission, on the other hand, we see two exponents where we would potentially expect only one based on the syntax. Compare Modern Hebrew *ti-xtev-u* 'y'all'll write' and *nixtov* 'we'll write', both in future tense:

(140) *ti-* xtev *-u*
SubjAgr.2Pers- write -Subj.Agr.Plural

(141) *ni-* xtov
SubjAgr.1Pers.Plural- write

Fission occurs in (140) but not (141), even though both presumably have the exact same syntax. This kind of split exponence is thus a post-syntactic matter. In classic DM, this was dealt with by a mechanism called Fission (Noyer 1992, Halle 1997) in the following way. Fission occurs *during* Vocabulary Insertion, and certain vocabulary entries, like Hebrew 1pl *ni-*, match/use up all features on the terminal. However, other vocabulary entries, like Hebrew 2pl *ti-*, incompletely match them, and the 'leftover' feature(s) are realized by a second instance of Vocabulary Insertion. Under these models, fission emerges due to the particular inventory of vocabulary entries, and whether they match up the terminals and leave a leftover feature or not. Note that in this sense, fission is distinct from being, say, listed as a circumfix on the exponence list, because the two exponents in (140)

[16]The most convincing cases would potentially come from ones where the portmanteau in question involves not just two demostrably distinct syntactic terminals, but say three or more. In such cases, positing two adjacent null allomorphs might seem like it is missing a generalization. Although the argumentation that relates to such cases is limited, there are some proposals for how to handle them in the literature in terms of what is called 'Nonterminal' Realization – the idea that an exponent can correspond to a set of multiple, structurally adjacent heads. Many of these proposals, however, require major changes to the basic principles of Vocabulary Insertion and the Maximal Subset Principle, as they require comparison across the feature sets of multiple adjacent heads in order to choose the most specific item.

correspond to a distinct set of features on the terminal node (e.g. person features to the left, and number features to the right, both sets coming from the same terminal), whereas with a circumfix, there might be a prefixal and suffixal component to the exponent but these wouldn't necessarily correspond cleanly to distinct and isolable features from the terminal they are exponing.

In the present discussion, we will consider a new post-syntactic mechanism for fission, as an operation on Terminals and Features prior to Vocabulary Insertion (and hence logically prior to consultation of the exponent list). The idea, like that when we considered impoverishment and meta-syncretism, is that generalizations about the form of exponence that are regular, recurrent, and above and beyond the form of particular details of the exponent list should be the result of operations (e.g. rules and constraints) that occur earlier than the process of Vocabulary Insertion. Putting it differently, the more systematic something is, the more derivationally earlier it should happen.

To anticipate the solution, we propose a mechanism for fission inspired by *Cross-modular Structural Parallelism*, one which is rule based and capitalizes on the notion of splitting up antagonistic features, with 'Pied-Piping' of Orthogonal Features, argued to be empirically necessary for Basque, and which affords crosslinguistic predictions about recurrent patterns of splitting.

(142) The Cross-Modularity Hypothesis:
Operations on abstract morphological structures are the same as ones that operate on phonological representations.

The idea is same grammar, but different alphabet: *phonological* features in *segments* vs. abstract *inflectional* features in *morphemes*. Here, therefore, we present an analysis of person-number fission in Basque verbal morphology, which is a crossmodular parallelism analysis, as it is based on diphthongization in Southern Italian, and in fact provides an empirically and typologically more satisfying approach than the predecessors mentioned above.

First, we turn to some background on the phenomena and Basque finite auxiliaries. Most finite sentences have a verbal complex with a tensed auxiliary:

(143) (Gu-k) (seu-∅) ikus-i *s-aittu-gu*.
we-ERG you-ABS see-PRF AUX
'We have seen you.' Lekeitio (Biscayan, Hualde et al. 1994)

Within the AUX are many morphemes: *Abs – T/Agr – Dat – Erg – C*. Let's discuss a few of these in turn. *T/Agr* includes (present/past) tense and agreement. *C* (often null) includes clause-type, tense, and agreement. Finally, there are pronominal *clitics* doubling absolutive, dative, and ergative arguments. Focusing on the auxiliary above, s- is a 2Sg absolutive clitic, *-aittu-*: 2Sg agreement, present tense, and *-gu* is a 1Pl ergative clitic.

We now turn to Person-Number Fission, as it relates to the expression of Plurality in Basque pronominal clitics. Plural clitics are split into two exponents in 2nd and 3rd person. Specifically, the plural form of a pronominal clitic is typically identical to the singular form of the clitic, plus an additional suffix -e. This split exponence is found in all plural clitics, except in the 1st person. The data below are from the Lekeitio dialect (Hualde et al. 1994). Note that 3rd person absolutives simply lack a clitic form altogether (this results from a fact about their syntax, not their exponence,

and hence their syntactic nonexistence is distinct from that of being a zero exponent realizing something syntactically contentful):

(144)

	Absolutive			Dative			Ergative		
		Sg	Pl		Sg	Pl		Sg	Pl
	1sg	n-	g-	1st	-(s)t(a)	-ku	1st	-t/da	-gu
	2nd	s-	s-…-*e*	2nd	-tzu	-tzu-*e*	2nd	-su	-su-*e*
	3rd	—	—	3rd	-tz(a)	-tz-*e*	3rd	-∅	-∅-*e*

The problem about fission and how it should be implemented now comes into focus. Like agreement in Hebrew's future tense, the realization of clitics in Basque sometimes involves splitting one morpheme into positions-of-exponence. For example, the 1st person dative plural is realized as a dedicated single affix *-ku*, but 2nd person is *-tzu* + *-e*, and 3rd person is *-tz* + *-e*. Now, importantly, this fission of a 'plural clitic' *-e* occurs for 2nd/3rd *ergative, dative, and absolutive* clitics, and this recurrence across three different cases begs a generalization beyond individual vocabulary entries.

In search of a solution to how to implement fission as the systematic splitting of person and number features specifically when the person features are [−author] and the number features are [−singular], we turn briefly to a wholly parallel phenomenon within phonology, whereby certain phonological features that would be otherwise combinatorially created nonetheless must undergo splitting into two separate positions. The case at hand from phonology involves taking the result of the morphophonological process of 'metaphony' and enacting the fission of a dispreferred bundle of vowel features into two separate vocoids, yielding diphthongization.

Metaphony in Southern Italian (Calabrese 1998) involves the formation of the plural forms of nouns and adjectives, where in certain contexts, the only marker of plurality is a change in the quality of the stressed vowel, such that stressed mid vowels become high. (This process can be thought of as akin to English pairs like *goose* ∼ *geese*, though is orders-of-magnitude more productive in Italian metaphony. Now, consider the dialect of Arpino, which has the following 7 vowels – and in addition to noticing which vowels it does have, notice which vowels it *doesn't* have:

(145)

i	u	[+high, +ATR]
e	o	[−high, +ATR]
ɛ	ɔ	[−high, −ATR]
a		

The combination [+high,−ATR] is banned in Arpino: there are lax mid-vowels, but not lax high-vowels. Now, we turn to the process of metaphony. Mid-vowels undergo Metaphony in Arpino Plurals, according to the rule in (146), with examples shown in (147):

(146) [−*high*, +ATR] → [+*high*, +ATR]

(147)

Singular	Plural	
fjórə	fjúrə	'flower'
mésə	mísə	'table'

However, the lax mid-vowels show a different outcome, as simply raising them one step in height would create the 'forbidden' combination [+high,−ATR]. Instead, these vowels unexpectedly show diphthongization in the plural:

(148)

Singular	Plural		
fɔ́rtə	fw*ɔ́*rtə	'strong'	not *fʊrtə
vɛ́rmə	vj*ɛ́*rmə	'worm'	not *vɪrmə

The key notion from Calabrese's work on diphthongization, which we can transplant back to the study of morphological features, is that when certain feature combinations are created, they cannot survive 'as is', and must be split into two separate bundles. In the case of diphthongization, the antagonistic features are split from each other into two segments, which otherwise share all orthogonal features. Let's return to the expression of plurality in Basque pronominal clitics. Plural clitics are split into two exponents in 2nd and 3rd person. Fission can be modeled on a parallel with metaphony-driven diphthongization, however, this time as a morphotactic constraint on 'co-exponence', banning the use of an exponent that realizes the two features *[−author, −singular]. Clearly, the combination of these two features is fine as far as syntax and semantics are concerned. However, suppose that a language-particular morphotactic constraint (active, as it turns out, in Modern Hebrew as well as in Basque, and across three different case paradigms in the latter) bans 2nd/3rd person clitics being realized together with plural. Let's take the repair mechanism parallel seriously, and assume that, prior to Vocabulary Insertion, in languages where this constraint is active, one must split the antagonistic features into 2 separate clitics. Crucially, the fissioned clitics share all other features besides the antagonistic ones that had to be split. This is shown below:

(149) No Fission in 1st person:

$$\begin{bmatrix} \text{Absolutive} \\ +\text{participant} \\ +\text{author} \\ -\text{singular} \end{bmatrix} \quad \begin{bmatrix} \text{Dative} \\ +\text{participant} \\ +\text{author} \\ -\text{singular} \end{bmatrix} \quad \begin{bmatrix} \text{Ergative} \\ +\text{participant} \\ +\text{author} \\ -\text{singular} \end{bmatrix}$$

g- *-ku* *-gu*

(150) 2nd person requires fission of forbidden co-exponence:

$$\begin{bmatrix} \text{Abs/Dat/Erg} \\ +\text{participant} \\ -\text{author} \\ -\text{singular} \end{bmatrix} \rightarrow \begin{bmatrix} \text{Abs/Dat/Erg} \\ +\text{participant} \\ -\text{author} \end{bmatrix} \quad \begin{bmatrix} \text{Abs/Dat/Erg} \\ +\text{participant} \\ -\text{singular} \end{bmatrix}$$

s- (absolutive) *-e* (absolutive)
-tzu (dative) *-e* (dative)
-su (ergative) *-e* (ergative)

(151) 3rd person requires fission of forbidden co-exponence:

$$\begin{bmatrix} \text{Dat/Erg} \\ -\text{participant} \\ -\text{author} \\ -\text{singular} \end{bmatrix} \rightarrow \begin{bmatrix} \text{Dat/Erg} \\ -\text{participant} \\ -\text{author} \\ \text{-}tz \text{ (dative)} \\ \text{-}\emptyset \text{ (ergative)} \end{bmatrix} \begin{bmatrix} \text{Dat/Erg} \\ -\text{participant} \\ -\text{singular} \\ \text{-}e \text{ (dative)} \\ \text{-}e \text{ (ergative)} \end{bmatrix}$$

We may now compare this mechanism, built on a morphotactic constraint and its repair creating a new position of exponence, with orevious accounts of Fission in DM (e.g., Noyer 1992, Halle 1997), which were driven by entirely by what happened to be the feature specification of the exponents realizing these terminal nodes. Such accounts cannot express the fact that fissed morphemes share most features (e.g. ones orthogonal to the constraint), and in addition, fail to capture cross-categorial generalizations, e.g. 2nd/3rd, not 1st, undergo Fission in clitics/pronouns in *all cases* in Basque (what we might call a 'meta-fission', parallel to the notion of metasyncretism). Finally, if fission were purely a property of the particular inventory of exponents within a given language (and whether they happen to include incomplete 'leftover' feature or not), this would make no crosslinguistic predictions about recurrent patterns of fission (e.g. 1st plural vs. 2nd/3rd, as opposed to other logically possible splits).

By contrast, given exponence co-occurence constraints, one would expect constraints like *[−author, ±singular] to have crosslinguistic generality. Thus while there are languages in which no or all plural clitics undergo fission, there are to our knowledge no known languages in which 1st person plural undergoes fission but 2nd/3rd plural does not. Similarly, under this specific co-exponence constraint, there can be no expected pattern of fission that would hold, say, in the singular but not the plural. In fact, a pattern akin to Basque, where plural exponence of non-1st-person elements takes place, can be found in Semitic (as mentioned above for Modern Hebrew, but also in Egyptian Arabic), in Georgian object clitics (152), and in Kadiwéu object clitics, and of course none of these four language families are historically related.

(152) Georgian object clitics:
Fission in 2 but not 1 (3rd doesn't cliticize)

	sg	pl
1	m-xatav	*gv*-xatav
2	g-xatav	*g*-xatav-*t*
3	xatav	xatav

(153) Kadiwéu object clitics:
Fission in 2 but not 1 (3rd doesn't cliticize)

1sg	i-diki
1pl	Go-diki
2pl	Ga-dikil-i

In a more general sense, then, pursuing fission based on crossmodular parallelism leads to one formulate constraints on shared exponence of certain features. Like all constraints, these share

the properties of stating a generalization across vocabulary entries, enabling a division of labor between the constraint and its repair, and spawning clear crosslinguistic predictions. Future work can chart the extent to which other cases of 'extended exponence' (e.g. one-to-many relations between syntactic terminals and exponence) can be modeled specifically by appeal to morphotactic constraints on antagonistic features that cannot be exponed together and require the creation of fissed positions.

22. Displacement and Doubling on the Output of Syntax

We've talked a lot about how head-movement, the mirror principle, and the exponent list are responsible for the order of formatives inside M-words. But what about cases that seem to violate what the syntax would deliver, and therefore seem to implicate some postsyntactic rejigging of the basic order established by syntax? Continuing with the theme we've developed thus far, let's explore the nature of operations that change morpheme order, and the constraints that might motivate such operations. We may start by distinguishing two kinds of morphotactic constraints on morpheme order. *Absolute* positional morphotactics are of the kind that state that a particular morpheme cannot occur at an edge, say, the leftmost part of the word. Such NONINITIALITY constraints turn out to be somewhat akin to the same requirements that cause most Germanic languages to be V2.

Relative positional morphotactics are of the kind that would state that certain kinds of morphemes must precede or follow others, regardless of what the syntax generates. One such case, to which we will return below, is the so-called 'CARP' template of Bantu languages, in which the causative, applicative, reciprocal, and passive morphemes must occur in exactly that linear order, even if the semantic scope is one where, say, the applicative is above the reciprocal, and hence should follow it. If the mirror principle doesn't happen to deliver this result, then morphotactic well-formedness will be achieved by postsyntactic reordering. Another kind is the morphotactic recurrence of the ordering "derivation before inflection", whereby derivational morphemes (e.g. *-tion*) come closer to the root than inflection (e.g. *-s*). While this latter constraint is usually guaranteed anyway by the order of syntactic heads in the tree (i.e. NumP is inherently higher than N-, whether or not the two come to form a single M-Word or not), we will see that certain diachronic changes can introduce violations of it, which will then require postsyntactic reordering.

Consider what happens as a consequence of the attachment of the modifier *in-law* to a noun like *sister* to yield *sister in-law* (a 'type' of sister, or kinship relation at any rate), or formations like *secretary general* (which is a kind of secretary). When pluralized, this postnominal modifier itself is not the head, and therefore one might expect forms such as *sisters in-law* or *secretaries general*. However, such cases involve a 'trapped' internal inflection, and what a lot of work on morphotactics reveals is that inflectional affixes (in particular, because they reflect relations with other sentential elements) prefer to be as external as possible in the M-Word. As a result, the tendency for many speakers is to place the plural affix outside of the modifier, producing forms such as *my sister-in-laws are visiting* or *the long line of secretary generals*. We can view this restructuring of the placement of the inflectional affix as the result of a more general morphotactic that governs relative position, one in which derivation precedes inflection (for suffixes). This was formulated as one of Greenberg's (1963) Universals; he listed it as number 28 in his list that derivation should be linearly between the root and inflection.

In a paper on diachronic change called 'the externalization of inflection', Haspelmath (1993) documented a number of cases where, when inflection gets trapped inside of derivation, eventually it would like to move out. For example, Spanish plural agreement becomes trapped inside the verbal reflexive marker *se* (a formative we've seen before) in imperatives like 'sit down!' addressed to a plurality of addressees, and this verb (*sentar se*) is one of the verbs (like English *perjure*) that is obligatorily reflexive. As a result, the form of this imperative is as follows:

(154) siente -n -se
sit -2pl. -refl ('sit down! (pl. imperative)'

However, the reflexive marker *se*, originally a weak pronoun, comes to be 'univerbated' (in other words, turned into part of the same M-word) as the verb plus agreement. As a result, the plural inflection in (154) is, like in the example of *secretaries general*, 'trapped' inside the word, and would like to move the rightmost, external position. That is to say, the reflexive marker (a kind of derivational affix, in the sense that it changes the valency and meaning of verbs that it reflexivizes) follows the purely inflectional number suffix in (154), and according to DERIV $\succ$ INFL this should become

(155) siente -se -n
sit -refl -2pl.

However, Haspelmath asserts that "Language change must be gradual, otherwise innovating speakers would not be understood by conservative speakers" (p. 302), and therefore introduces the notion of intermediate hybrid forms, such as (156)

(156) siente -n -se -n
sit -2pl. -refl -2pl.

The idea is that as part of the transition from the original form (154) *siente -n -se* to the morphotactically preferred form (155) *siente -se -n*, "speakers have no choice but to create hybrid forms" like *siente -n -se -n*. As Haspemath puts it, "innovations can take only one step at a time, so hybrid forms are necessary". (p.302). However, he leaves open the question of the transition from hybrid forms such as (156) to the eventual (155): "How do speakers get rid of the residual, nonfunctional internal inflection? Some details of the final cleaning up remain to be accounted for" (p.303). We therefore wish to characterize formally the notion of hybrids, and the cleaning up mechanism. Put differently, the goal of a theoretical characterization of this change is to provide a mechanistic explanation that can account for the morphotactic violated by the old forms, the creation of hybrid forms as a response to the morphotactic, and the "one step" innovation that leads to eventual metathesis.

Unaware of these Haspelmathian desiderata, Harris & Halle (2005) developed a framework to represent partial reduplication. Interestingly, their formalism was developed for cases of partial reduplication in phonology, where only some of the segments in a single morpheme are copied, such as in the Kusaiean language in which *kulus* 'to peel' has a reduplicated form *kul-kulus* 'to peel bit by bit'. The way to represent the demarcation of the copied sequence in their formalism is by square brackets, and thus *[kul]us* would be interpreted as an instruction to copy the sequence delimited within the square brackets. Representing it schematically, partial reduplication can be interpreted as follows:

(157) a. [AB]CD $\rightarrow$ ABABCD
b. AB[CD] $\rightarrow$ ABCDCD

While the formalism above represents reduplication of a subsequence, there are cases of partial

reduplication in which the 'reduplicant' (the result of copying) is not contiguous to its original source. In the formalism, these can be characterized by angle brackets, which involve an instruction to delete part of the copied sequence. More specifically, let us consider an angled bracket like $>$ as indicating that the material to its right within the square brackets is deleted in the second copy that is made:

(158) [A $>$ B] $\rightarrow$ ABA~~B~~ $\rightarrow$ ABA

Interestingly, a mnemonic way to teleologically characterize the result of (158) is to think of the "trapped item", which must move out to the right, points towards its destination with an angled bracket, although mechanistically this involves copying plus deletion, as shown above. By parity and symmetry, a left angled bracket indicates that the material to its left must be deleted in the first copy that is made:

(159) [A $<$ B] $\rightarrow$ ~~A~~BAB $\rightarrow$ BAB

Now let's turn to the interesting, unexpected result of the formalism developed so far. What is the result of a sequence of brackets like [A><B]? Taking it step by step, it involves deletion of the material to the right of the right-pointing bracket the second time around, and deletion of the material to the left of the left bracket the first time around:

(160) [A $>$ $<$ B] $\rightarrow$ ~~A~~BA~~B~~ $\rightarrow$ BA

As you can see above, therefore, the combination of both brackets delivers the result of metathesis, the wholesale reordering of A and B. What is interesting and important about this formalism, therefore, is that is closely ties together partial reduplication with metathesis, where the latter becomes in a sense a very special case of the former. Let's now return to our examples from the Spanish imperative. Note that the trapped inflection would like to 'move rightwards' to the external position, and that this is accomplished by placing a set of copy and delete delimitations on the sequence of terminal nodes as follows:

(161) siente [-n $>$ -se] $\rightarrow$ siente -n -se **-n**

The result of (161) is the hybrid form that we have seen above. Moreover, the metathetic version involves the simple addition of a single angled bracket:

(162) siente [-n $>$ $<$ -se] $\rightarrow$ siente -se **-n**

While (162) respects the morphotactic DERIV $\succ$ INFL, the question arises as to how (161) in fact represents a morphotactic improvement over the version of it without any partial reduplication at all. To see how in fact the external copy of the *-n* in (161) somehow 'helps' to improve the well-formedness of the structure, suppose DERIV $\succ$ INFL has a weaker version:

(163) *UNIFORM INFL $\succ$ DERIV: An inflectional affix I cannot uniformly precede a derivational affix D, where *uniformly precedes* means that all tokens of I precede all tokens of D.

Given this definition, in *siente -n -se* ***-n***, inflectional *-n* no longer uniformly precedes derivational *se*. Now, we can understand the hybrid forms as being gradually eliminated by the addition of the extra >. Stepping back from this particular example, we observe an important prediction of the Harris-Halle Formalism (also known as the Generalized Reduplication (GR) formalism). Given the fact that [A ><B] is but a step away from more conservative [A > B], all cases of synchronic morpheme metathesis are likely to be accompanied by dialectal and diachronic **doubling**. In other words, wherever we see morphotactic reordering that involves displacement in the form of metathesis, we should expect that a similar form with partial reduplication should be nearby in time and space.

The existence of doubling cases, such as *siente -n -se -n*, provide an important argument for this type of operation being post-syntactic, as opposed to the result of syntactic movement. While syntactic theory includes 'the copy theory of movement', with the mechanism for pronouncing higher or lower copies of moved elements, it generally does not provide any way of pronouncing two copies of a moved element, except for very limited cases. Thus *Who will you vote for who?* is essentially never attested. Similarly, in *siente -n -se -n*, the doubled element *-n* is an agreement morpheme hosted on T, and under no mainstream theories of syntax would T itself be moving to above the reflexive marker and thereby setting up a structure with two copies to pronounce. In short, the existence of doubling (alongside displacement) does not seem to be a syntactic matter, and for this reason is handled as a morphotactic operation on the output of syntax, in response to M-word specific requirements on well-formedness, and with morphology-specific repairs employed to acheive it.[17]

23. Case Studies in Positional Morphotactics

In support of the idea that displacement on the output of syntax is closely linked to doubling, we can consider a few case studies. As we have already discussed, the comparative formative in English starts out hierarchically higher than the adjective that it takes as an argument, and thus we might expect it to precede it, if it were a freestanding word. However, this morpheme is an affix that cannot be pronounced in isolation. While it might be tempting to say that the Adjective moves to the Comparative head via Head Movement in English, this clearly wouldn't explain what's happening in the case of *more intelligent*, where no head movement is going on. We wouldn't want to say that head movement, a syntax-internal operation, is conditioned by the phonological size of the adjective. As a result, it is more sensible to assume that head movement of the adjective to CMPR never takes place in English, and instead the CMPR head is left alone by the time syntax is done. However, it is morphotactically weak and requires a host to its left – we may say that it has a NONINITIALITY requirement within its own M-Word. As it turns out, there are two ways to satisfy this constraint. One is an epenthetic morpheme, inserted to its left, which

[17] Interestingly enough, while the GR formalism was originally motivated based on partial reduplication in phonology of a subsequence of segments, its extension to morphology may in fact work better for the latter, as purely phonological synchronic metathesis is somewhat rare (the diachronic cases, such as Latin *milakro* → Spanish *milagro* do not involve an underlying form being synchronically changed) and as phonological metathesis is, as far as we can see, usually not accompanied by partial reduplication. It is an interesting consequence of this 'crossmodular' comparison that the formalism that closely links metathesis to doubling may be better suited to morphology than to phonology.

happens in the case of *more intelligent* or the superlative *most intelligent*. We can call this operation *mo*-support – a morphotactic repair with the purpose of ensuring that a dummy morpheme, *mo-*, inserted postsyntactically, guarantees that CMPR will be non-initial within its M-word. However, another possible repair – restricted to the case of adjectives that are bisyllabic or smaller – involves displacement:

(164) [est > <unkind]: *unkind -est*

(165) *mo-* support when host too big: *mo-st helpful*

(166) *mo-* support when host too far: **the amazingly kind-est person, the most amazingly kind person*

Recall the prediction of the GR formalism is that whenever metathesis is found, partial reduplication should not be far away in time or space. In fact, Shakespeare used the hybrid form, when speaking of *the most unkindest cut*:

(167) a. Shakespeare uses hybrid form: [est > unkind]: *-est unkind -est*
b. . . . plus *mo-* support: *most unkindest*

In sum, we have seen that positional morphotactics, whether they be relative (in the case of derivation and inflection, or the CARP template in Bantu) or positional (as in the non-initiality requirement on English CMPR and SPRL are constraints that may be repaired by a variety of 'tactics': epenthesis, metathesis, or doubling.

24. Variable Post-Syntactic Processes and Social Factors

Up until now we have developed tools that model the competence of speakers in assembling, molding, and realizing the form of words, and viewed a number of conditions and operations that operate along the way to yield a deterministic surface form. However, a ubiquitous phenomena across languages is that of intra-individual variation, the kind that is studied within sociolinguistic and variationist investigations of language. Generally speaking, sociolinguistic studies aim to study the amount of correlation between (and where possible, develop causal models of) a linguistic variable and a social variable. William Labov's research during his Masters' thesis, conducted on the island of Martha's Vineyard at a time when it was rapidly changing due to a conflict between increased tourism and traditional fishing town identity, included a correlation between a linguistic variable – the amount of diphthong centralization (pronunciation of /aɪ/ as [ʌɪ]) a given speaker exhibited – and a social variable. The social variable in this case was the quantitative result extracted for each individual from the rating on a questionnaire of subjective feelings of 'Island Identity' (e.g. rating from 1 to 5 how strongly one held attitudes such as 'the island shouldn't change for the sake of tourists'). Plotting the results for a group of individuals, therefore, allows one to put, say, diphthong centralization on the y axis and degree of island identity on the x axis, and then calculate the amount of correlation, and when, as was the case in this study, that the correlation between diphthong quality and a social variable is strong, it allows one to conclude that social identity may consciously or unconsciously mediate the rate of a variable phenomenon. In his PhD thesis, Labov studied the amount of nonrhoticity (*r*-dropping) in codas, by asking sales clerks in New York department stores where, say, men's belts could be found (knowing in advance that the answer would always be 'fourth floor', by asking questions he already knew the answer to). He then calculated the amount of nonrhoticity as a linguistic variable (e.g. 'fowth flooah'), and correlated this with the type of department store (using variables such as price of clothing as a quantitative reflection of whether the consumer demographic of that store was socioeconomic class A, B, or C). In general, variable phenomena (such as flapping in American English [sɪɾi] 'city', or glottalling in British English [sɪʔi] 'city'), which are optional rules often conditioned by one's age, gender, identity, socioeconomic class, social attitudes, and so forth, may be modelled in terms of coefficients that act as the 'weights' on a bunch of conditioning factors. Recall the equation for a line you may have learned in maths classes:

(168) $y = ax + bw + cz$

In such equations, the value of y is the result of the sum of weighted variables. For example, x, w, and z might be demographic factors, and a, b, and c are the coefficients that reflect the amount that these variables matter for a given phenomenon. In fieldwork by Silva & Nevins on the completely optional process of consonant weakening in the Brazilian language Maxakalí, for example, the authors found that consonants became weakened in a pattern predictable by speakers' age, and also by whether or not the consonant was in a stressed syllable. However, in that particular sociolinguistic study, it turned out that gender had no predictive power, and was therefore essentially 'silent' as to whether consonant weakening happened or not. In this case, then, x and y and z would be speaker age, speaker gender, and word-in-question stressed-or-not, respectively, and the

coefficient for this third one would simply be zero. Let's therefore formalize this as follows, where instead of calling the value y, we call it p, for the probability of the process happening, to range between 0 and 1. We can therefore, annotate the rule of glottaling in British English in a format such as the following:

(169) t → ʔ (with rate of application *p*) / v́ _ v

The formalization in (169) is known as a 'variable rule', but it is variable in a very restricted and particular sense of the word. The structural description of the rule is deterministic: in between a stressed vowel and an unstressed vowel. The structural change is also deterministic: /t/ changes into [ʔ]. There is no variability or nondeterministicness either in what the structural change is nor in what the structural description of. What *is* variable is whether, every time the structural description of the rule is met, the structural change actually fires or not. Thus, whether a given British English speaker, when he or she has a /t/ in between a stressed vowel and an unstressed vowel, actually applies the rule of glottaling turning the /t/ into a [ʔ], will apply with probability *p*, where p represents the amount of weights on a host of conditioning factors (such as age, geographic area of upbringing, amount of projected 'street cred' the speaker is unconsciously aiming for, and so forth). The theory of phonological rules does not change as the result of the fact that this particular rule is conditioned by social factors – the formulation of the rule remains with the traditional, deterministic structural description and structural change with clear input-output relations. What changes is the fact that the rule only fires with a probability *p*, ranging between 0 and 1, that is potentially conditioned by the weighted sum of these social factors. As such, the continued formalization of phonological rules can proceed in a 'modular' fashion, without integrating social factors into the *nature* of the rule, and only into a probabilistic annotation on that rule, the computation of which can be carried out in a modular, distinct part of the grammar focused on the relation between social factors and language use.

In a sense, this division of labor can be seen as parallel to the distinction usually made between semantics and pragmatics: the former refers to the constant, fixed, deterministic meaning of lexical items (such as the fact that *some* means nothing more than ∃ in logic), while the latter refers to the inferences made in real-time conversational context (such as the fact that when *some* is used, it can lead to the additional inference that the speaker did not choose to use the logically stronger quantifier *all* instead, and this choice must have been for a good reason). The theory of semantics does not need to necessarily change the denotational entry for *some* simply because in certain conversations, inferences are made in language usage that it implicates *not all* – the latter is the result of a separate, pragmatics module. The relationship between phonological rules and between social factors in language usage that condition the probability that such rules are applied reflects a similar modular distinction between a situation-invariant formalization and a situation-dependent influence of and knowledge about how speaker- and hearer- specific factors may add to the choice of and interpretation of certain combinations and operations.

When it comes to phonological phenomena such as glottaling above, therefore, the division of labor between phonological formalisms and weighted probabilities have enjoyed a certain currency in sociolinguistic research – indeed, the computer program Varbrul, the staple of sociolinguistic analyses, is based on a weighted-sum among conditioning factors, and the search for which fac-

tors matter more than others, in contributing to the value *p* for a given linguistic variable (such as diphthong quality, or glottaling). In the domain of morphosyntax, however, there has been somewhat less discussion about how the weighted-sum model might apply to variable phenomena, in part because of the question of how such phenomena themselves are modeled in rule-based terms. Luckily enough, as it turns out, the kind of model of morphosyntax we've been developing all along is well-suited to the variable rule approach as well, because a given morphotactic rule (say, impoverishment) can be annotated with the probability *p* in a manner entirely parallel to glottaling above. Let's take as a case study the syncretism found in English auxiliary systems: it has been found that in Monmouthshire English (as studied in Nevins & Parrott 2010), speakers might say either *I am* or *I be*, where the latter reflects the elsewhere item in the exponent list for this particular auxiliary system. This can be modeled as impoverishment:

(170) [+part, +auth] (with rate of application *p*) → ∅ / [+cop, −past, +sg]

The formalization of the environments where this impoverishment rule will apply is explicit, deterministic, and situation-independent, as is the result of the rule (feature-deletion). However, the rule has been annotated with a probability p that integrates the weighted sum of a number of conditioning factors that sociolinguistic studies may reveal (e.g. age, class, gender, and even other potentially influential factors of a more linguistic nature, such as the token frequency of the word following the copula, if that turns out to be relevant). Studies of this sort, where impoverishment rules are formulated in categorical terms but have a variable probability of firing, have been applied not only to English but to intra-individual variable phenomena in the verbal systems of Catalan and European Portuguese, and in principle could be applied to any of the kinds of morphotactic rules we have examined (e.g. bracket insertion that causes metathesis and/or doubling), and it is a matter of sociolinguistic and variationist research (in the form of naturalistic interviews, speaker attitude surveys, or a host of other methods) to uncover the relative weights of factors that matter most in conditioning whenever a rule does not apply with a 100% rate of firing. Even phenomena such as the use of *ain't*, when a given speaker has more exponents than just this one (e.g. sometimes says *I haven't been* and sometimes says *I ain't been*, with an overall p between 0 and 1), can be modeled in terms of an impoverishment rule that allows the less-specific exponent *ain't* (which is, after all, underspecified for number and person, as well as [± cop], differentiating *have* and *be*).

In pursuing the parallel, therefore, between glottaling as a kind of feature deletion (after all, it removes the supralaryngeal specification of the coronal stop), and insertion of *ain't* as feature-deletion, both of which are conditioned by a probability *p* – the nature of which is the object of study of variationist research using sophisticated tools such as Varbrul) – we reach the question of whether the transmitted 'social meaning' of, say, performing glottaling or realization by *ain't* have identical representations in terms of how this 'meaning' is arrived at. It is somewhat of an unfortunate consequence of working at the level of lexical items (like *ain't*) that some researchers have decided that 'what it means' for a speaker to use *ain't* instead of *haven't* somehow merits the annotation of its social status (e.g. stigma or street cred) within its core denotation, and such a move seems erroneous for many reasons to which we turn. To be clear: the stigma or street cred associated with the use of *ain't* in a given conversational situation is not part of its encyclopedic entry, any more than the stigma or street cred of glottaling is not part of its encyclopedic entry,

nor in turn any more than the fact that *not all* may be implicated in the usage of *some* is not part of its encyclopedic entry (as we know, indeed, from the fact that such pragmatic implicatures are *cancelable*, while assertions – part of the encyclopedic entry – aren't). The inverted Y-model we have adopted – in which syntax jointly feeds LF and PF, at which point the two modules go their separate ways, and the exponent list and the encyclopedic list are distinct entities with no access one to the other – indeed allows no kind of access within the encyclopedia of whether or not a given PF operation such as impoverishment (or for that matter, glottaling) has happened or not. There is simply no way in the current model for the encyclopedia to have access to which exponent was inserted by the Maximal Subset Principle, nor which impoverishment rules potentially fed or bled Vocabulary Insertion of specific exponenets. Any kind of social meaning associated with the rate of application p of specific morphotactic rules, therefore, cannot be part of the encyclopedic entry, and must be the result of a computation that takes place somewhere else, modularly speaking, during the usage and knowledge of usage of language in real-time conversations.

So where would this kind of knowledge lie? If a given speaker enacts glottaling (and therefore signals, consciously or not, that they are a UK, and not US speaker, to their interlocutor, and possibly further specific class- or regional- information), where does the interlocutor make these inferences from? They would have to come from the interlocutor's inferences about the relative coefficient weights on the speaker's p and the factors themselves. Thus, if I hear you perform glottaling, I can make the inference that, for you as a speaker, the value of the factor "from UK" is 1, and that its coefficient is nonzero; and similarly with other inferences I might have about the causal structure of your equation p for this rule as they pertain to class and regional information (and of course, part of these inferences might be missing for me, should I lack sufficient information about these conditioning factors). As such, therefore, any 'social meaning' associated with the use of *ain't* or the pronunciation of /t/ as [ʔ] that cause me to infer you are communicating 'I am a proud, edgy, twenty-something East Londoner' are cancelable, situation-dependent, and based on pragmatic-like inferences independent of the fixed, time-independent denotations of items on the encyclopedic list.

Having discussed the question of *how* to model variable phenomena in exponence (e.g. impoverishment), we still should spend a bit more time discussing why certain morphosyntactic phenomena show any variability at all to begin with, and indeed echo an undercurrent of work within generative linguistics that it is equally important to study where variability occurs as to delimit where it does *not* occur. In the case of English morphosyntax, there are two well-studied realms. The first is auxiliary leveling, as mentioned above for the case of Monmouthshire *I be*; there are also numerous cases in the presence of negation (e.g. Smith Island *I was* but *I weren't*; the variability with *ain't*, the **amn't I lucky* variable repairs mentioned earlier in the text). The second is the case realization of pronouns in coordination (e.g. *John and me went to the store*), which is particularly vexed in the environment of the Saxon genitive (e.g. *John and me's book, John's and my book, John and I's book*, etc.). In the auxiliary leveling cases, the variability in patterns most likely has to do with the effects of marked combinations (e.g. past tense, negation, etc.) and variable repair mechanisms to delete one of the offending features, as well as a certain amount of ambiguity in the analysis of the English auxiliary paradigm itself (as it involves syncretism across plurals, but also involving the number-neutral *you*). In the coordination cases, it likely has to do

with ambiguity in the way that case assignment (or case allomorphy) rules, formulated on the basis of adjacency and hierarchy in non-coordinated cases, interact when the structure is suddenly one headed by a coordinate phrase, interrupting the direct sisterhood relations between specifier and head, as well as a certain amount of prescriptive intrusion on the ordering of pronouns within the coordination (e.g. 'avoid putting me first'), and a certain amount of hierarchical decision-making among what may be competing constraints on realization of case on each coordinate. As such, certain patterns of meta-variability might be predicted and explored, such as the lack of a dialect of English with variable *was/were* leveling in the affirmative, but not the negative, or in the present, but not the past, or in coordinated pronouns in nominative / spec,T position, but not in genitive / spec,D position.

25. Speech Errors and Late Realization

Psycholinguists have long been delighted at the possibilities for exploring 'how the system really works by seeing where it breaks down' in the study of speech errors, slips of the tongue, and spoonerisms. In order to kick off the discussion, consider the following list of speech errors as collected and/or organized by Roland Pfau, all of which involve *transposition* at some linguistic level or of some specific linguistic unit. The arrows in the examples below always point from the intended utterance to the one that was erroneously produced in its place.

(171) examine the horse of the eyes ← eyes of the horse

(172) He taught courses → he caught torses

(173) a weekend for maniacs ← a maniac for weekends

(174) I saw him digging it up → I see him dugging it up

(175) I regard this as imprecise → I disregard this as precise

(176) the bonsai died because I didn't water it → The bonsai didn't die because I watered it

(177) the gardener has to pull up the dead flowers → has to die the pulled up flowers

(178) I think it's careful to measure with reason ← I think it's reasonable to measure with care

(179) Rosa only dated shrinks → only date shranks