

1 Overview

We'll have one reading per day, and one exercise per day, which is highly optional but highly helpful.

For the next four days, we'll focus on the morphological arrangements and operations in the following systems of agreement:

- Day 2: Semitic circumfixes (Arabic, Hebrew) [when phis are in two places]
- Day 3: Algonquinian person (Cree, Ojibwa) [when phis are realized with unexpected values]
- Day 4: Spanish themes (gender and clitic structure) [when phis don't match the 'real world']
- Day 5: Hindi ergativity (long distance and default agreement) [when phis cross paths with grammatical relations]

But none of that can start before I introduce the lenses through which we'll look at these, so today we discuss:

- Geometry: The *Organized* Structure of Phi Features
- Underspecification: The *Mapping of Phi Features* to a Set of Affixes
- Cyclicity: The *Realization of Phonological Form* in a stepwise derivation provided by the syntactic tree

2 Why a Geometry?

I'll start with a model of phi-features that is not a bundle of values such as $\{+2, -sg, +masc\}$, but rather one with a hierarchical structure. I will draw it on the board; see Harley & Ritter's paper.

- A geometry makes predictions about the relative markedness of featural complexes cross linguistically (e.g. that duals and inclusive 'we' are not very common)
- A geometry provides a familiar formalism for operations such as delinking and reassociation
- A hierarchical geometry captures asymmetries that occur in syncretic phenomena (e.g. that when a contrast is neutralized it is done so in favor of the less marked form)

- A hierarchical geometry arbitrates during the realization of one of two affixes neither of which are fully specified (think of having two pegs in two holes, and having to map them to one slot; which do you choose?); eliminating the need for extrinsic ordering

3 Underspecification

Let's start right away with facts about the world: The nominative and accusative are identical in the neuter in German

All forms of the present tense are identical except for 3ps in English

Can this be accidental homophony? Would an efficient brain list:

like = 1ps

like = 1pp

like = 2ps

like = 2pp

like = 3pp

likes = 3ps

Of course not. An efficient brain would list:

/like/ ↔ √*LIKE* (in the locus of being immediately dominated by verbal, aspectual, and tense nodes)

/-s/ ↔ [non-participant in the conversation] [group cardinality = 1]

(under the node reflecting values in the variable space characterizing dimensions of the reference set of the external argument of the verb)

/0/ ↔ anything else under the node reflecting values in the variable space characterizing dimensions of the reference set of the external argument of the verb

While the Vocabulary Items (that is, the list of affixes pairing phonological content with a syntactic terminal) above were given rather wordy definitions, we will a) adopt shorthand and b) realize that the loci definitions above are those and only those determined by principles of syntactic composition that we must independently model.

Affixes in a list like the one above won't be very efficient if they are to be the atoms that combine in Merge and Move, precisely because they're underspecified:

- (1) **Anytime there are two levels of distinct representation** (and any linguistic model needs at least two here: phonological form and syntactic terminals) **and one representation carries less information than the other** (our hypothesis about the featural content of elsewhere affixes), **and a deterministic mapping is needed between these levels** (because the sounds of the utterance "the drums play softly" reflect the same syntactic structure every time) **then the level of representation with less information must be the target of the mapping** (in derivational terms, it must be determined "after" the first).

We will examine many context-sensitive processes, such as different realization next to different syntactic neighbors (contextual allomorphy) and modification of morphosyntactic features when in a given neighborhood. If morphology is to precede syntax, how will the system "know" that the affix that goes "in the numeration" will end up with a particular derivational sister that's the right one? To give a very coarse metaphor: would you put on all of your clothes before you had absolutely any idea what the weather was like outside?

Underspecification: In order for a VI to be inserted at a terminal node, the features that the VI realizes (is the exponent of) must be a subset of the features at the terminal node. **Insertion may not take place if the item has identifying features that are not at the node.** VIs are characteristically underspecified, leading to *syncretism* in paradigms – cases in which the same VI is inserted for different feature complexes. The familiar example is the zero affix in "eat+0", which realizes the terminals made up of [1 sg], [3 pl], etc.

But underspecified VIs *need not be zeros* – take the example of /-e/ which realizes strong adjectival themes in Dutch for neuter singular and both plurals, and /0/ in the singular neuter – in this case the VI whose phonological form is zero is *more* specified. The underspecified item is often called the Elsewhere item, following a tradition in phonology going back to the Indian linguist Panini: "More specific first, then, if nothing matches, go to the general case". When doing a DM analysis (as a linguist or a child, presumably) the elsewhere item is the one that appears in the most heterogeneous environments.

(2) **Exercise: What is the elsewhere: Icelandic weak adjectival inflection**

	masc-sg	fem-sg	neut-sg	masc-pl	fem-pl	neut-pl
nom	i	a	a	u	u	u
acc	a	u	a	u	u	u
dat	a	u	a	u	u	u
gen	a	u	a	u	u	u

Answer: u.

4 Cyclicity

A derivational implementation of the locality restrictions displayed by *contextual allomorphy* is easily modeled as a cyclic, one-by-one transduction from abstract syntactic terminals to syntactic content. The bottom-up orientation is probably one with which we are familiar, so we will adopt that for current purposes, as it seems an efficient model of the actually observed "bottom-up" as opposed to top-down asymmetries in contextual realization.

Vocabulary Items must inevitably specify the environment in which they realize certain features. The feature [comparative] is realized by /-er/ in the environment of a bisyllabic adjective and /more/ elsewhere. Note that phono-

logical content can be realized by a null affix, as in the case of the past tense morpheme in "hit+0". Another famous example: \sqrt{GOOD} (a root expressing a constellation of affairs/properties the speaker likes) is realized as /bett/ in the environment of a [comparative] feature- head and /good/ elsewhere.¹

A more intricate example comes from Itelmen, a Chukotko-Kamchatkan language spoken on Russia's Bering sea coast.

- (3) t'-alcqu-γin
1s:su-see-2s:ob
"I see you"
- (4) t'-k'ol-k(icen)
1s:su-came-1s:su "I came"

In Itelmen, prefixes always express features of the subject.

When the verb is transitive, suffixes express features of the object.

When it is intransitive, suffixes express features of the subject. Moreover, when the object is 3rd person, the choice of the suffix is determined by the person features of the subject and the number features of the object². The object marker is conditioned by the subject marker, and not vice versa.

Now let's inspect the class markers; there are two conjugation classes in Itelmen. The class marker is highly allomorphic depending on the features of both the subject and object.

- (5) t'-ankzu-s-ce?n (lelku-?nl-a?n)
1s:su-help-pres-3p:ob/1 (mouse-pejor-pl)
"I'm helping the mice"
- (6) t'-tO-s-ki-ce?n (c'eβuzlaχ-a?n klcl-e?n)
1s:su-bring-pres-class2-3p:ob/1 (tasty-pl rotten.heads-pl.)
"I'm bringing tasty rotten (mouse) heads"

This leads to a structure [subj-agr [obj-agr [class [verb]]]]. A flat structure could not account for the range of attested and non-attested dependencies. A hierarchical structure that allowed inward or outward allomorphy could not predict the asymmetries of Itelmen (e.g. that conjugation class never affects subject agreement or object agreement, and that object features never affect subject agreement).

The cyclic procedure of spell-out described thus far is one that trades each abstract syntactic terminal for phonological content. Call this *Vocabulary Insertion*. Now, suppose that a VI (Vocabulary Item) is listed in the brain as [A] ↔ /ga/, meaning "anytime you are spelling out a syntactic tree, and you see the feature [A], supply the phonological content /ga/". Suppose we assume iterative rule application (or anything that will allow me to provide identical

¹A strong conjecture is that this type of suppletion never happens for roots, which requires saying that forms that typically undergo suppletion (e.g. good, woman, numbers) are universal and built of primitive universal non-root features.

²An analysis of 3rd person as no person dovetails well with that.

phonological exponence for a given abstract feature complex more than once, since I can say "I like what I like"); what is to prevent /ga ga ga ga/ at a single terminal? Clearly we have to assume either a nonredundancy constraint, or that spellout of a feature "consumes" that feature; in other words, yields it invisible to further operations in the morphology component.

- (7) If spell-out of a feature is a literal trade of phonology for syntactic features, then once a terminal is spelled out, it will no longer be present as the environment for contextual allomorphy for another terminal.

5 Emergence of the Underspecified

We have discussed pretty much faithful mappings from phonology-free syntax to syntax-free phonology. But sometimes what the phonology reflects seems at odds with the syntax and how it is understood. In some of these cases, an operation on a terminal will make the expected VI ineligible, and a relativized elsewhere will surface. First let's examine the paradigm of Spanish clitics:

(8)

		3m	3f	2	1
ACC	sg	l-o	l-a	t-e	m-e
	Pl	l-o-s	l-a-s	o-s	n-o-s
DAT	sg	l-e	l-e	t-e	m-e
	pl	l-e-s	l-e-s	o-s	n-o-s
REF	sg	s-e	s-e	t-e	m-e
	pl	s-e	s-e	o-s	n-o-s

Why the dashes above? Harris, etc analyze clitics as possessing the same internal structure as every other Spanish noun, verb, and adjective: having a theme vowel³. When we look at the difference between accusative and dative 3rd person clitics: they differ only in theme vowels. While an atomic analysis of "les" vs. "los" would make no predictions as to their similarity, the decomposition allows a simple condition:

- (9) Spell out the theme as class 3 (-e-) in the environment of Dative case. A reasonable list of the remaining VIs is as follows⁴

- (10) Under the terminal node [DET]:
- /n/ [class1] ↔ 1 in the env pl
 - /m/ [class 3] ↔ 1
 - /0/ [class1] ↔ 2 in the env pl
 - /t/ [class 3] ↔ 2
 - /l/ ↔ [] / Case

³The rules for insertion of theme vowels are conditioned by syntactic gender and conjugation class; we will go into it on Day 4.

⁴Note that some VIs carry an inherent specification for conjugation class and some don't this is an obvious fact of Spanish based on *gato/gata* vs *poeta/poeta* vs *comandante/comandante*, with the first root underspecified for class, the second root specified for class 2, and the third for class 3. More on this on Day 4.

/se/ [class 3] ↔ []

(Note that neither /n/ or /0/ *realize* plural, they are just more specific allomorphs than their counterparts with the identical features.) Now: Latin American dialects lack "os", and it is replaced by "les" when Dative and "los" when Accusative. Let's assume this is the result of the following rule:

- (11) Remove the 2nd person feature from a DET in the environment of plural.

Notice that this rule has a different effect from an inventory gap:

- (12) There is no VI for "[2] in the env pl" in Latin American Spanish (i.e. remove it from (11))

The effect of (12) would be to replace "os" with "te". This simply doesn't happen; (11) predicts this wouldn't be the case, as [2] is deleted and an unspecified VI is inserted. Moreover, the fact that the 2p clitic in Latin American Spanish shows Case distinctions ("les" vs "los") falls out naturally as the VI, now unspecified for inherent class, is subject to the class 3 / dative condition.

This is a result that shows that impoverishment cannot be replaced by positing inventory gaps.

Let's turn to the Spurious "se" in Spanish, which appears instead of "le" or "les" for a 3rd person dative when it is adjacent⁵ to an accusative clitic. Suppose this is the result of the following impoverishment rule⁶

- (13) Remove the Dative specification on a clitic in the environment of an accusative clitic.

The fact that "se" appears instead of "le" is now expected, as it is the least specified VI. But what is interesting is the comparison between Continental and Latin American Spanish. In the former, "I gave it to you" is *os lo di*, as the removal of the Case feature on the dative terminal will still leave *os* as the most specified VI; hence there is no appearance of the impoverishment effect in the surface for 2nd person datives.

However, in Latin American Spanish, the impoverishment of [2] in the plural (from (11)) excludes "os", and the impoverishment of Case in the environment of a neighboring accusative (from (13)) excludes "les"; hence LAS expresses this

⁵The ordering of clitic sequences is another matter of open interest, subject to wide variation. Bonet proposes a template, but one would want to find some principles underlying allowable templates. Note also that "se se" is ungrammatical (impersonal - reflexive) but while in Spanish the impersonal has to go (and be expressed periphrastically) in Barceloní it is the reflexive that is expressed periphrastically.

⁶I am fairly confident that should one wish, many of these impoverishment "rules" could be replaced by licensing conditions a la GP. As a constraint-based approach, Grimshaw (1997), in an OT analysis, attributes the *le lo to a morphological OCP (in particular, these are both -reflexive) combined with the universal markedness hierarchy *Dat >> *Acc. See Cuervo (2002) for qualms, however, with the -reflexive part of this, involving sequences of 3 clitics.

phrases as *se lo di*⁷.

One would be extremely hard pressed to find a syntactic motivation for this rule⁸; it is quirk of morphology that, should we want, we can find a functionalist explanation for (e.g. simplification, parsing cues) much in the same way that we can for final devoicing, which is essentially an operation of the same nature: In a Given Environment, Remove a Feature from the Representation. It is an accepted fact that German speakers perform an "removal" operation on the voicing feature at the end of a word every time they speak; the proposal of impoverishment is one in which such removal can take place on syntactic features as well.

Note that from an early insertion point of view, the spurious *se* is bizarre, esp. for LAS – the "reflexive" clitic *se* is introduced to realize 2 dative plural when it normally realizes none of those features. Moreover, a pre-syntactic paradigm optimization would have to include a third "dimension" of the cells, with "precedes accusative clitic" in the new dimension, and "se" in a completely heterogeneous range of places⁹. Moreover, the existence of such a paradigm is an undesirable move to begin with, as the analysis of these clitics into the constituents of separate Determiner, Theme, and Number nodes is more insightful. VI 1: [f1 f2] ↔ /kuhu/ VI 2: [f1] ↔ /nyaali/

6 Enriching Impoverishment

Stump argues that the Macedonian paradigms below suggest that an item-based approach misses a generalization about the 2sg and 3sg syncretism:

(14)	padn- ('fall')	present	imperfect	aorist
	1sg	padn-0-am	padn-e-v	padn-a-v
	2sg	padn-e-š	padn-e-0-še	padn-a-0
	3sg	padn-e-0	padn-e-0-še	padn-a-0
	1pl	padn-e-me	padn-e-v-me	padn-a-v-me
	2pl	padn-e-te	padn-e-v-te	padn-a-v-te
	3pl	padn-0-at	padn-e-0-a	padn-a-0-a

Stump proposes a rule stating that 2nd person becomes 3rd person in the environment of [sg, past]. Of course, this rule could go the opposite way, too; there is no theory of markedness in his "Rules of Referral". A feature-geometric delinking, on the other hand, could only delete the [addressee] node, leaving a bare referent node, whose interpretation is by default 3rd person; the opposite

⁷Note the lack of plural on "se" is an allomorphy present in all of its uses; but its expression will resurface in an interesting way on Day 3's exercise.

⁸Unlike, say, the person case constraint, which is extremely widespread, effects like the spurious *se* are missing from even closely related languages.

⁹Moreover, for LAS, something additional would have to be said about why, given an input like [2-pl-DAT ACC], a constraint like *[les ACC] could not be obviated by insertion of "te", which at least realizes [2].

could not be expressed by an impoverishment rule¹⁰. The rule would be to remove [2] in the environment of [sg] and [past].

7 Exercise: Tamazight Berber

Imperfect for "dawa" (cure)

As a first try, we write down all the VIs (with any accidental homophony) for:

	singular	plural
1	dawa-γ	n-dawa
2m	t-dawa-d	t-dawa-m
2f	t-dawa-d	t-dawa-m-t
3m	i-dawa	dawa-n
3f	t-dawa	dawa-n-t

If possible: notice the syncretism that occurs – where? What contrast is being neutralized?

What featural distinctions are not made in a certain person?

Can you think of any way to eliminate the accidental homophony via impoverishment?

What might be a principled reason there is only one suffix for 1plural?

¹⁰This is not to say that such things can never happen .e.g a 3→2 change, but it would be predicted to be much less common on familiar grounds of analytic simplicity.