Hug a Tree: Deriving the Morphosyntactic Feature Hierarchy.*

Heidi Harley, Massachusetts Institute of Technology

1.

2. 1 Introduction

Any theory attempting to account for morphological phenomena will, of necessity, make reference to morphological features. Indeed, purely descriptive accounts employ morphological features, as the very notion of a "paradigm" is defined by cross-classifying different varieties of features, e.g. number and person. In a theory where morphological operations are an autonomous subpart of the derivation, such features acquire status beyond that of a descriptive convenience. They become linguistic primitives, manipulated by the rules of word-formation.

Different approaches have treated these primitives in various ways. Most often, they are regarded as an undifferentiated bundle, broadly divided into nominal vs. verbal elements, but with little or no organization internal to the broader categories. Noyer (1992) motivates some organization: he posits a *Universal Feature Hierarchy*, that interacts with the feature bundle and his proposed morphological filters to account for an observed cross-linguistic tendency to favor certain morphological contrasts over others. In this paper, I hope to take this hierarchy one step further, and motivate the idea that these features, like other linguistic primitives, are subject to a specific sort of hierarchical organization, namely a morphological feature geometry.

Feature geometries have long been an accepted fact of phonology (Clements (1985), Sagey (1986)), and a morphological feature geometry would presumably provide at least some of the same theoretical advantages as its phonological counterpart. The primary goals of a phonological feature geometry, as described by Noyer (1992), are as follows:

a. Subtrees define natural classes of features for phonological rules
b. (Some) nodes in a phonological feature tree correspond to the anatomy of the vocal tract.
c. The dependency relation encodes contrastiveness

(b) is not necessarily true - many nodes in a phonological feature tree do not correspond to anatomical structures (and phonology does not necessarily reduce to phonetics). An ideal morphological feature geometry would accomplish (a) and (c) for the level of morphological structure.

* I am indebted to the following people for their help and encouragement: Andrew Carnie, Chris Collins, Morris Halle, Jim Harris, Alec Marantz, Rolf Noyer, and Colin Phillips. Inconsistencies, incompletenesses, and general wrongnesses are totally and utterly my fault.

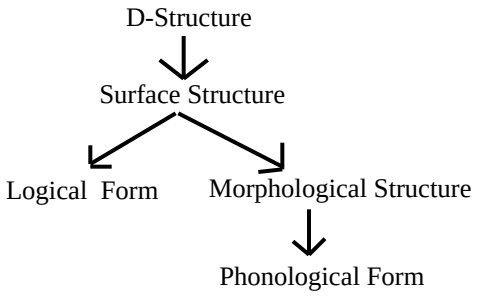
I show that the feature-geometrical approach to morphology defines a notion of complexity from which the rule-ordering effects of Noyer's hierarchy can be derived, as well as determining which features in a given filter will be Impoverished. These properties then fall out from the representation, rather than a completely separate metric for determining relative "weight", or markedness. Further, a feature-geometrical approach is suggested by certain cross-linguistic groupings of morphological features. Such groupings can never be more than coincidence in a framework that represents features as unorganized bundles rather than geometrically.

3. 2 Framework and Assumptions

3.1.

3.2. 2.1 Morphological Structure

I will take as a starting point the model of morphological analysis adopted in Noyer (1992). This model situates a level of Morphological Structure[1] within the familiar derivation schema of government-binding theory, between SS and PF, after the branching to LF, as follows:



MS consists of the rules and principles that supply the component of the phonological form of words which is sensitive to syntactic processes, for example agreement, case-marking or incorporation. Syntactic heads (X^{o}s) are at least partially made up of morphosyntactic features, all of which are fully specified by the end of syntax. MS rules function to first simplify, and then phonologically spell out, these features. Full phonological specification is then determined at the next level of representation, PF.

MS processes come in several flavors. The initial simplification (called *Impoverishment* after Bonet (1991)) is driven by language-specific limits on the complexity of inflectional categories, using a mechanism first developed for phonology by Calabrese (1988). Linear relationships between X^{o}s are also specified at this level; such relationships do not hold in the syntax, but must hold in phonology (Halle and Marantz 1993); the process is called *Linearization*. Further, two X^{o}s can undergo *Merger* or *Fusion* operations, producing a single morphological head (analogous to head-movement in the syntax)[2]. The linearized morphological heads that result from the combined effects of these processes are termed M^{o}s.

---

[1]Noyer in fact terms this level "Morphology", but to retain the nomenclatural parallel in DS and SS I adopt Halle and Marantz's (1993) sobriquet "MS".

[2]These two types of X^o melding have different characteristics and consequences, see Halle and Marantz (1993) for discussion.

These strings of Mºs then undergo two processes which map them onto phonological strings. One of these processes is a set of well-formedness conditions on morphological words, for example the requirement in Latin that nouns be supplied with a "thematic vowel" or stem formative. Such conditions convey no syntactic information whatsoever, and can be completely devoid of any sort of morphosyntactic information as well, as demonstrated in Harris (1991) for Spanish "word-markers". The other process is the set of rules mapping Mºs to phonological spell-outs, divides into two blocks: one, termed "readjustment rules", change phonological information in certain environments, and the other supplies phonological material, in the process "discharging" the morphosyntactic features of the Mºs. This division separates the actual realization of morphosyntactic features (affixation) from phonological changes which are merely conditioned by morphosyntactic features: for discussion of this division, see Noyer (1992) and Halle and Marantz (1993).

The processes of MS can be schematized as follows:

Syntax
↓
Morphosyntactic well-formedness:
Linearization, Merger/Fusion, Impoverishment
↓
Mº String
↓
well-formedness conditions on
morphological words; readjustment
rules, affixation/discharging
↓
Phonological Form

Of these processes, the two that operate directly on morphosyntactic features are Impoverishment and affixation, and these are therefore the processes with which Noyer's Universal Feature Hierarchy is most concerned.

1.3. 2.2 Affixation

1.3.1.

1.3.2. ***2.2.1 Discharging***

Morphosyntactic representations in Noyer's framework contain both morphosyntactic features and "positions-of-exponence", which are slots in the morphosyntactic representation that can accept a phonological string. Affixation in Noyer's framework results in the "discharging" of both morphosyntactic features and these positions-of-exponence. Discharging a feature or position renders it inert to future affixation rules: a discharged element cannot be the principal exponent of (i.e. spelled out by) any later rule, rendering "double" realizations impossible - e.g. a true double plural marking, where the same feature [pl] is realized twice (as is claimed for Breton double plurals by Stump (1989)), is impossible in this framework.. Similarly, one position cannot be filled twice, realizing two Mºs (although two Xºs can undergo Fusion to create a single Mº, which can then fill a single position, e.g. the fused number-case

affixes of Latin). Discharged features can, however, still condition morphosyntactic allomorphy (readjustment rules). Discharging features thus gives the "bleeding" effect necessary to ensure that later rules do not apply after earlier rules have already realized a given feature, for instance, it blocks the application of the English rule *[pl]* →*/-z/* when the more specific rule *[pl]* → */-en/ /OX___* has already applied, preventing the formation of **oxens*.

1.1.3. ***2.2.2*** ***Affixation and the Universal Feature Hierarchy***

Rule ordering for affixation is determined in two ways Firstly, according to the Paninian principle, when a rule refers to a subset of the features mentioned in another rule, the latter applies before the former. Where this does not suffice to decide the ordering of two rules of totally or partially disjoint content, Noyer's Feature Hierarchy determines which rule will apply first - the rule making reference to features higher on the hierarchy takes precedence over the rule making reference to lower features. This is required, for instance, in the rules he postulates to account for the Classical Arabic prefix-conjugation paradigm.

Noyer also invokes the feature hierarchy to order blocks of discharge rules, (as proposed in Anderson (1992). In particular, the Nunggubuyu transitive agreement rules for realizing the Subject and Object affixes can occur in either order: subject affix realized first, then object affix, or object affix, then subject affix. Two principles govern this ordering, one syntactic (subjects > objects), derived from the degree of syntactic embedding, and one morphosyntactic: (person features > number features > gender features > class features), derived from the Feature Hierarchy. The Feature Hierarchy, therefore, can determine the ordering of blocks of affixation rules (i.e. affixes), not only the ordering of rules within blocks.

1.4. 2.3 Impoverishment

Xºs come out of the syntax fully specified for all morphosyntactic features. As is readily observed, however, most paradigms do not fully cross-classify - not every block in a grid of intersecting person and gender features will be filled with a distinct phonological realization. For instance, French first person singular pronouns have gender in the syntax, as evidenced by the agreement on the adjective in sentences like *Je suis heureux/heureuse* -- however, the first person singular pronoun *je* is the same for both masculine and feminine speakers. Further, many languages make morphosyntactic distinctions others fail to reflect entirely - Zarma, a Songhang language of Africa, makes no gender distinctions whatsoever, unlike virtually all Indo-European languages.

Noyer proposes a theory of morphosyntactic filters to account for both the failure of paradigms within languages to fully cross-classify, and the apparent absence of certain morphosyntactic features from the paradigms of most/many languages that do in fact appear in others. These filters are modeled on those proposed by Calabrese (1988) for phonology.

1.4.1. ***2.3.1*** ***Phonological filters***

Calabrese's filters are central to his theory of acquisition of phonology. A child learner is provided by Universal Grammar with a phonological feature geometry capable of representing all the phonological

contrasts possible in human language. In the given language the learner is exposed to, it is highly likely that many of these contrasts will not be employed. Calabrese posits a series of filters: negative cooccurence restrictions on features of the geometry, such that an acquired filter will prevent representations containing occurrences of both the phonological nodes contained therein. Such representations are subject to "clean-up rules", also provided by UG -- for instance, a rule of Delinking that will remove one of the offending features/nodes from the representation.

An example of such a filter can be seen in example (1) below:

1. *[+low, -back]

This filter is active in Italian, representing the fact that the low front vowel /æ/ is absent from the Italian phonological alphabet.

These filters are presented to the learner by UG at predetermined stages of language acquisition in inverse order of complexity. The more complex a phonological segment a filter excludes from the representation, the later the stage at which it will be presented to the learner. Further, if a learner, hearing no positive evidence that the filter currently being presented is not in effect, acquires that filter, then the learner will acquire all of the subsequent filters excluding more complex segments without need for further filter presentation.

1.4.2. ***2.3.2*** ***Morphological filters***

Noyer proposes a set of morphological filters that prevent the occurrence or concurrence of certain of the morphosyntactic features. These filters account for features that are completely inactive in a given language, as well as the gaps in paradigms mentioned above. He gives as an example the Arabic Imperfect Indicative Prefix Conjugation, example 2 below[3]:

2.

singular	dual	plural	
y-aktub-u	y-aktub-aani	y-aktub-uuna	3m
t-aktub-u	t-akttub-aani	y-aktub-na	3f
t-aktub-u	t-aktub-aani	t-aktub-uuna	2m
t-aktub-iina	*	t-aktub-na	2f
÷-aktub-u	*	n-aktub-u	1

[3]Abbreviations are as follows:

1	first person
2	second person
3	third person
pl	plural
m	masculine
f	feminine
spkr	speaker
prt	participant
inc	inclusive

The gaps in the paradigm can be characterized as in (3) below:

3. a. There is no gender distinction in the first person
   b. There is no dual number in the first person
   c. There is no gender distinction in second person dual
   d. There is no first person inclusive v. exclusive distinction.

Noyer accounts for these with the filters in (4):

4. a. *[1 f]
   b. *[1 dual]
   c. *[2 dual f]
   d. *[1 2]

These restrictions are active in most Indo-European languages, which do not distinguish gender for (most) verb forms, have no dual number, and no inclusive-exclusive distinction. The threshold filter that is responsible for activating these filters is lower in Indo-European than in Arabic - in Arabic, all earlier filters have been turned off, as the filters in (4) characterize exactly the Arabic paradigm, while in Indo-European, these filters are true of the verbal paradigm but not characteristic of it - the filter *[dual] has applied, for example, entailing the application of the later, more complex filter *[1 dual]. [dual] is not syntactically active as a feature in Indo-European[4], while it is in Arabic. The distinction between filters that prevent a given feature from being at all active in a language and filters that merely produce "gaps" in a paradigm is not formally treated by Noyer, although he implies that such a distinction should be made. The effect produced by the latter type of filters is what he refers to as Impoverishment (following Bonet (1991)).

1.4.3. ***2.3.3*** ***Filters and the Feature Hierarchy***

The filter *[1 f], in 4a. above, expresses the notion that the features [1] and [f] cannot occur as part of the same M^{o}. The effect is to delete the feature [f], leaving first person with no gender distinction. Noyer points out that the same effect could be produced with a rule, as below:

5. $f \rightarrow 0 / 1$

This mechanism is too powerful, however, as it could just as easily express the opposite deletion:

6. $1 \rightarrow 0 / f$

The effect of (6) would be to make first person feminine homophonous with third person feminine. This rule is apparently not operative in any language, and therefore a properly constrained theory should not be able to encode it.

---

[4]This is not to say that words cannot be lexically marked as referring only to groups of two - "both" is an example of such lexical definition.

To ensure the correct application of only the former filter and the impossibility of the latter, Noyer again invokes the Universal Feature Hierarchy. In filter (4a), above, the feature lower in the hierarchy will delete, leaving the higher feature - in this case, [f], being a gender feature, is lower than [1], a person feature, and therefore deletes. Noyer's mechanism thus allows only the necessary operations.

4.

5. 3 Feature Geometry

5.1.

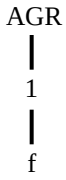
5.2. 3.1 Why not?

5.2.1. ***3.1.1*** ***Filters and repair strategies***

Noyer points out that if the hierarchy is represented graphically, as a feature tree, the deletion of [f] above could be represented as *Delinking*, one of the "repair strategies" employed by Calabrese to rescue filter-violating representations, and also by Bonet, to express Impoverishment in her morphological feature geometries.

Given that a geometry should allow the required delinkings and no others, Noyer postulates that a partial geometry representing (4a) would look like (7) below:

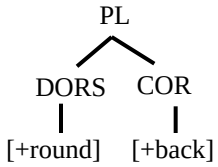
7.



This geometry expresses the notion that [1] cannot delink unless [f] also delinks. It follows that the smallest operation to repair the structure in accordance with 4a. will be the delinking of [f] only, and the opposite delinking, that of [1] while [f] remains, becomes impossible.

Noyer then notes that in Calabrese's system, while the phonological feature geometry expresses which delinkings are possible, it doesn't express which delinkings are necessary. So for example, to repair the filter violation of *[+round, -back] that occurs when the front rounded vowel /y/ is part of a borrowed word, an Italian speaker can delink either of the two features, given the following articulator-based feature geometry (Sagey (1986)):

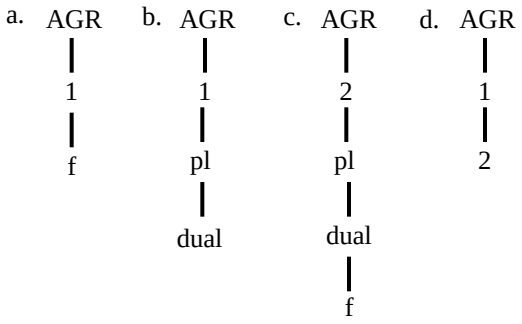
8.



Since neither feature dominates the other, the filter violation can be repaired by delinking either of the features. This prediction is borne out by the evidence, as Italian speakers can represent /y/ as any of /u/, /i/ or /iu/ (the latter represents the repair strategy *Breaking*, which will be discussed later). Noyer then concludes that the phonological feature geometry fails to determine which delinking is necessary in this case. In the morphological case, optionality of repair is not possible, and only one delinking is allowed for any one of Noyer's proposed filters. A feature geometry would have to express which repair strategies were necessary - that is to say, dominance relations would have to exist between features higher in the hierarchy and features lower in the hierarchy. Otherwise, delinking in order to repair filter violations should be optional, which at least in these cases, it isn't.

This, then, implies that partial feature geometries for all the filters in (4) above would have to look like those below::

9.



5.2.2. **(Note that in each of the above geometries, the hierarchy 1 > 2 > pl > dual > f holds as dominace relations.. A feature geometry should be able to capture the effects such a hierarchy (as first proposed by Silverstein (1976) and Lumsden (1987, 1992), and adopted and elaborated by Noyer (1992).**

5.2.3. ***3.1.2 Resulting problems: affixes as subtrees***

Phonological feature geometries can be seen to be motivated by the criteria listed in the introduction, above, repeated as (10) below:

10. a. Subtrees define natural classes of features for phonological rules

b. (Some) nodes in a phonological feature tree correspond to the anatomy of the vocal tract.

c. The dependency relation encodes contrastiveness

(10b) above, has no real morphological correlate. (10a) could have a morphological correlate: affixes (spelled out by MS rules) may only represent subtrees of the feature geometry. Noyer claims that given the conclusion above, i.e. that hierarchical relations between features are necessarily expressed by dominance in a feature geometry, the morphological analog of 10a. has immediate false consequences:

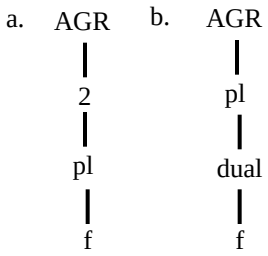
11. a. t-aktub-na `you (f, pl) write'

2-write-fpl

b. t-aktub-aani `they (f, dual) write'
f-write-dual

The features expressed in these forms give the geometries of 12, below (again, by the conclusion that hierarchical relations are expressed as dominance).

12.



In (11a), the prefix expresses [2], and the suffix [f pl]. If /t-/ is limited to being a subtree of (12a) (by (10a)) then it should express not only [2] but everything [2] dominates. In (11b), [f] is realized as a separate affix, although it is dominated by [dual]. If these feature geometries are correct, (10a) has no morphological counterpart, rendering a major part of the motivation for the additional structure of a morphological feature geometry moot.

5.2.4. *3.1.3 Subtrees and "enhancing" features*

Noyer's treatment of "enhancing features" poses problems for a morphological feature geometry as well. In phonology, one feature can be primary while another feature enhances the first (from Stevens and Keyser (1989)), as is the case in many languages with the features [+back] and [+round] - rounding back vowels enhances the phonetic contrast between them, but it isn't the case that all back vowels in all languages are [+round].

An enhancing morphological feature would have a similar function - it is potentially present, indeed, often present, but not necessarily so. The feature [dual] can be, and normally is, enhanced by the feature [pl][5]. However, it can be the case that [pl] is not present in some cases when [dual] is present. Take, for example, the Arabic perfect suffix-conjugation:

13.

a.	**katab-at**	*she wrote*
b.	**katab-at**-aa	*they (f.dual) wrote*
c.	katab-na	*they (f pl) wrote*
d.	katab-ti	*you (f) wrote*
e.	**katab-tum**-aa	*you (dual) wrote*
f.	**katab-tum**	*you (pl) wrote*

[5]in Noyer's terms, [-sg]

The dual form is created by the addition of the suffix /-aa/. In (13b), the third person feminine dual is formed from a base of third feminine singular (13a) - the dual suffix is added to the singular form. In (13e), however, the second person dual is formed from the second person plural (13f). A form explicitly marked as meaning “two” is inherently plural, and presumably overt plural marking is therefore unecessary. The feature [pl] is "enhancing" the second person dual in the case of the second person, while it is not present in the third person feminine dual.

Noyer's enhancing features are limited to those that are logically predictable, given the primary feature. When a form is marked [dual], [pl] is logically predicable - if a form is referring to a group of two, it is necessarily plural. At LF, [pl] will be otiose in the presence of [dual]. Thus the possibility of a filter deleting it from the MS representation, as follows:

14. *[pl dual] / third person

In a feature-geometrical theory, this approach does not readily suggest itself. If some morphological feature is logically implied by another feature, that is, if feature [F] entails the appearance of feature [G], the natural way to capture this relation in feature-geometry is through dependence. We can say that feature [F] is dependent on feature [G], as in the feature-geometrical formalism below:

15.

This would suggest that the most felicitous way to represent [dual] would be as dependent on [pl], corresponding to [F] and [G] respectively in (15) above. Such a formalism, however, leaves no possibility of the dependent feature occurring in the absence of the dominant feature, rendering an enhancement analysis of the Arabic paradigm above problematic.[6]

5.2.5. *3.1.4 Contrastiveness and dependency*

Consequence (10c) makes wrong predictions for this feature geometry as well. Noyer points out that (10c) expresses the notion that phonological features are dependent on articulators, and thus only contrastive when the node representing that articulator is present. No analogy with morphology therefore exists (as for (10b)), and, he claims, problems arise when (10c) is applied to the feature geometry suggested above. For example, if [2] is dependent on [1], the feature [2] should never be able to be expressed unless

---

[6]In subsequent treatments of number systems in Mam and Susurunga, Noyer captures the notion of the monovalent feature [dual] using a combination of features [-sg] and [-aug]. Such a treatment also leaves no room for [-sg] as an enhancing feature for [dual], rendering treatment of the Arabic paradigm above difficult.

A possible solution in the feature-geometrical approach might arise if the notion of "Breaking" as a filter-repair strategy is introduced, in parallel with Calabrese's Breaking strategy. Such an approach would split the subtree containing [dual] dominated by [pl] off from the tree in the third person and cause it to be realized separately as the /-aa/ suffix, leaving an unmarked NUMBER node in the original tree, causing it to be represented as singular. A precise formalization of such a notion would involve redefinition of several concepts in Noyer's framework, such as Splitting and Positions-of Exponence, and is beyond the scope of this paper.

the feature [1] is also expressed - that is, [2] would always contain the notion [1], giving a first person inclusive reading, and a simple category [2], corresponding to `you', should be impossible.

Based on these conclusions, Noyer remains noncommittal about the idea that a feature geometry would have anything to offer morphology.

5.3. 3.2 Why?

A morphological feature geometry can primarily be motivated to accomplish (10a) and (10c) above, that is, to define natural classes of features, and to encode contrastiveness/dependency. If a morphological feature geometry could be found that would encode these types of relations and in addition derive the effects of Noyer's feature hierarchy, it would seem natural to assume that some such organization of morphological features is indeed part of UG, especially because such organizing trees are assumed to be part of the phonological component, and it would seem likely that the closely related levels of MS and PF would make use of similar devices.

It then is necessary to examine the points Noyer raises as problematic for the sort of dependency/dominance relations entailed by a feature geometry.

5.3.1. ***3.2.1*** ***Encoding the feature hierarchy via dominance***

If dominance relations are necessary to capture the effects of the filter hierarchy, as Noyer maintains above, then it appears that affixes are not constrained by a morphological feature geometry, since the Arabic paradigm in (11) contains affixes that are not subtrees of such a geometry. Two moves are possible here. One is to claim that the feature geometry doesn't constrain the content of affixes, but only that of Mºs. The feature geometry is a representation which is subject only to the operation of filters, but not to the operation of discharging processes. This is supported in that filters constrain the content of Mºs, rather than the content of affixes. Noyer demonstrates that in Arabic, the formation of a first person dual is prohibited, even though there are affixes that could attach to the word without violating any well-formedness conditions to create such a form:

16. a. ÷-aktub-u *I write*
1-write-IMPF
b. t-aktub-u *you* (m) *write*
2-write-IMPF
c. t-aktub-aani *you* (dual) *write*
2-write-dual
d. *÷-aktub-aani *we* (dual) *write*
1-write-dual

Although (16d) fulfills Noyer's condition on the well-formedness of Arabic words which motivates the splitting of Mºs into prefixes and affixes ((16d) has both) it is nonetheless ungrammatical. This can only be because the Mº, prior to splitting, violated the filter *[1 dual]. Impoverishment could

therefore still be accomplished by delinking features from a morphological feature geometry, which would then have fulfilled its function and disappear, thus not constraining the content of affixes.

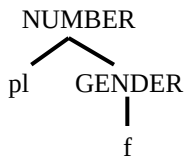
This solution is less than satisfactory, however. If a feature geometry is postulated for MS, presumably it should accomplish all the functions specified above - that is, take the place of the feature hierarchy in determining such things as rule ordering and affix ordering, and also constrain the content of affixes. If the feature geometry dissolves after the Xºs are transformed into Mºs, independent appeal to a feature hierarchy is still needed to determine rule ordering and affix ordering. Needless to say, having a feature geometry would then become fairly pointless, as the feature hierarchy could also accomplish the constraining of Impoverishment rules exactly as proposed by Noyer, and the extra theoretical baggage would serve no purpose.

Even if dominance relations encode relative status in the feature hierarchy and Noyer's proposed geometries are correct as far as they go, they could still function to constrain the content of affixes in the paradigm he outlines above. All that is needed is a further notion of discharging as a cyclic operation. The Mº for (11a) can be represented as (12a), and the suffix representing [f pl], /-na/, is attached, "discharging" these features. If discharging renders these features inert to further rules, the operation can be pictured as delinking the [pl] node, dominating the [f] node, from the tree and "attaching" it to the position-of-exponence. This would then leave the feature geometry of the Mº with only the [2] dangling down, dominating none of the already discharged features, enabling the prefix [t-] to attach, discharging the [2] but nothing else, as the earlier discharging of [pl] and [f]has rendered [2] a subtree of the Mº. A similar process could account for the example in (11b), with the opposite ordering - prefix attaches first, then the suffix.

This solution is also problematic, particularly because of the ordering problems. It would entail just the opposite result from that implied by the feature hierarchy — rules making reference to features lower in the hierarchy apply earlier than rules making reference to features higher in the hierarchy. Wherever Noyer's treatment of a paradigm involves ordering rules based on the feature hierarchy, with higher features discharging before lower features, the above analysis would predict the opposite ordering, with presumably dire consequences. Either all Noyer's treatments of the paradigms would have to be reworked, or some way of encoding the feature hierarchy in a feature geometry that does not depend on dominance relations needs to be found.

5.3.2. 3.2.2 Proposing a feature geometry for agreement features

To a certain extent, I will assume that Noyer's feature hierarchy is encoded via dominance relations, although only incidentally. It will in fact be encoded via a notion of embeddedness/complexity - the more nodes it takes to represent a feature, i.e. the more embedded it is (the more marked it is), the lower it is on the hierarchy. This entails that features higher on the hierarchy are less embedded than features lower in the hierarchy, but crucially, it does not entail that higher features actually dominate lower features. Thus, to capture the fact that [pl] is higher on the hierarchy than [f], a partial geometry might look like this:

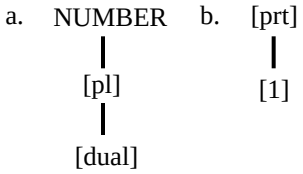


Any feature geometry of which (17) is a subtree will thus encode the notion that [f] is lower on hierarchy than [pl], as [f] will always be dominated by one more node than [pl].

The dependence of the GENDER node on the NUMBER node is suggested by facts about Arabic Concord relations, as noted by Noyer: Agreement between adjectives and head nouns in Arabic copies number and gender features together, but not person features, suggesting that in some respect number and gender form a subtree. Further, Noyer quotes Greenberg: "Whenever the verb agrees with a nominal subject or nominal object in gender, it also agrees in number... if a language has the category of gender it always has the category of number". This is suggestive of an organization whereby a node NUMBER dominates a node GENDER, according to the argument laid out in (15) (see also below). Note that a feature-hierarchical approach cannot readily capture these facts.

Further, I will assume that Noyer's "enhancing" features are in fact represented feature-geometrically as posited in (15) Any feature [F] that entails the appearance of another feature [G] is dominated by that feature [G]. Thus, [dual] is dominated by [pl], and [1] is dominated by [participant] (henceforth [prt]), giving the subtrees in (18) below[7]:

18.

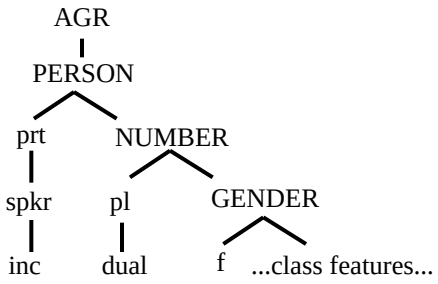


(This raises the predicted problem for the analysis of Arabic in (13) above.)

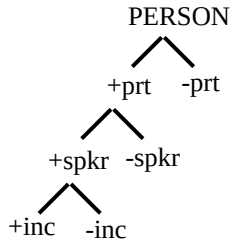
Finally, the broader hierarchy determining affixal ordering in Nunggubuyu above (person features > number features > gender features > class features) referred to above is represented via degrees of embedding, as outlined above. These nodes, analogous to the organizing nodes of phonology, are in a dominance relation with one another, although the features they themselves dominate are not. This gives the following proposed feature geometry:

19.

[7] A third necessary implication is that between the feature [inc] and the feature [pl] - inclusives are of necessity plural, but nowhere is this encoded in the feature hierarchy in (19). At first glance, it seems likely that no feature geometry will be able to accomplish this without giving up either the notion that affixes realize subtrees of the geometry (see sections 3.1.3 and 3.2.2) or the equally necessary implicational relationship between [inc] and [spkr]. I leave the question of whether this is a serious problem to further study.



This feature geometry is of the type proposed for phonology by Rice and Avery (1991), one in which only marked features can have dependents[8]. The person features could be represented as follows:
20.



For convenience, only positive values will be included in feature trees below; any node that could conceivably have another feature dependent on it and in fact does not have such a feature should be assumed to have the negative value for that dependent feature present. Whether negative values can be active in morphological structure (i.e. are visible for rule application) or are (radically) underspecified will be discussed in Section 3.4 below.

Note that a feature [2] is not present, rather it is represented by a tree that is [+prt] and [-spkr], and in which [inclusive] is a separate feature hanging off [spkr]. This obviates the problem noted by Noyer and discussed in 3.1.4 above, that a feature geometry in which inclusiveness is encoded via features marking both [2] and [1], and in which dominance encodes the feature hierarchy, the tree for simple second person should be impossible, since by the hierarchy [2], being lower than [1], should be dependent upon it, and thus any tree expressing [1] should express [2] as well. In this system, the dependency is reversed, making [spkr] lower in the hierarchy than [prt], but including a feature [inc] that represents first person inclusive as lower on the hierarchy than either. This geometry could be seen as one where increasingly specific sets are picked out by increasingly complex feature trees. The instance Noyer uses to motivate the portion of his hierarchy [1] > [2] is the application of a filter preventing first person inclusive, i.e. *[1 2]. In order to prevent [1] from deleting, making first person inclusive homophonous with second person, he posits that [1] is higher on the hierarchy than [2], thus ensuring that [2] will delete. In the tree above, that filter could be represented as

---

[8]Avery and Rice's geometries have monovalent features that come in marked/unmarked pairs; e.g. of the pair [oral] and [nasal], [oral] is marked and can have dependents (e.g. [lateral]), while nasal is unmarked and cannot have dependents.
The mathematical possibilities for this type of tree are the same as for binary features in which only positive values can have dependents - the value [nasal] could be represented as [-oral] with no difference in representational power.

21.

As [inc] is dependent on [spkr], it will necessarily delink to repair the representation. [spkr] cannot disappear and leave [inc] present. Thus, the reversal of [prt] and [spkr] in the hierarchy doesn't cause problems for that portion of the analysis.

5.4.

5.5. 3.3 Consequences of a feature geometry

5.5.1.

5.5.2. ***3.3.1*** ***Complexity of representations and the feature hierarchy***

Calabrese's phonological filters and Bonet's Impoverishments were motivated to reduce the complexity of representations. By delinking a node in a feature-geometrical representation, the complexity of the representation was reduced. Which delinkings were possible was an inherent property of the feature geometry itself.

Noyer's system, in which features occur in an undifferentiated bundle and this undifferentiated bundle is subject to deletion of features in response to filters, needs additional mechanisms to determine which delinkings are possible. The feature hierarchy is this additional mechanism. A filter such as *[f pl], active in English in the third person, works to simplify a representation. The undifferentiated bundle of features will be simpler, prima facie, no matter which of the features in the filter are delinked - one feature fewer in the bundle. The feature hierarchy imposes a restriction, stating that [pl] is higher in the hierarchy and therefore [f] should delete. No obvious correlation between the complexity of the given feature and its position in the feature hierarchy exits - it is merely stipulative to connect position in the feature hierarchy with the complexity or markedness of the feature. Such a connection would have explanatory power, however - it should be the more complex/marked element that is delinked, given the fact that filters work to simplify representations.

The connection between which element is delinked and which element is more complex/marked can be naturally represented if a feature-geometrical formalism is adopted. The more nodes necessary to represent a given feature, the more complex it is[9]. In filters like *[f pl] above, the most embedded feature - the one that is dominated by the most nodes - will delink.

As demonstrated in (17) above, [f] is more embedded than [pl], and therefore [f] will delink, rather than [pl]. Not only is the hierarchy encoded in such a formalism, an explanation for why the hierarchy is ordered as it is provided by the formalism.

[9]For an account of a phonological process involving node-counting, see Carnie (1994) (this volume), where sonority is determined via a counting mechanism; or also, Rice (1992).

5.5.3. 3.3.2 Complexity and rule ordering

As noted above, Noyer's feature hierarchy also serves to determine rule ordering, where such ordering isn't determined on the basis of the Paninian principle (otherwise known as the “Elsewhere Principle”). Roughly speaking, the principle states that where two rules X and Y in the same rule block compete for application, and the structural description of X is contained in the structural description of Y, Y will apply before X. For some rule blocks, however, this principle is not enough to ensure correct application of rules. Where X and Y have disjoint or overlapping structural descriptions, an additional principle is needed to determine ordering. Noyer uses the Feature Hierarchy - rules that refer to features higher on the hierarchy apply before rules that refer to features lower on the hierarchy. This process is illustrated below for the Arabic imperfect conjugation.

The paradigm appears as example (2), repeated here as (22) for convenience:

22.

singular	dual	plural	
y-aktub-u	y-aktub-aani	y-aktub-uuna	3m
t-aktub-u	t-akttub-aani	y-aktub-na	3f
t-aktub-u	t-aktub-aani	t-aktub-uuna	2m
t-aktub-iina	*	t-aktub-na	2f
÷-aktub-u	*	n-aktub-u	1

In order to capture discontinuous bleeding and violations of the Adjacency Constraint, Noyer proposes that the single person-number-gender Mº produced by the first half of the rules of MS is split into two positions-of-exponence, a prefix and an affix, to satisfy a well-formedness condition on Arabic nonperfect verbs. The rules discharging the features of the Mº will all be part of the same rule block, feeding and bleeding each other, although they are realizing both prefixes and affixes. The features occupy their position, and are discharged when a spell-out rule attaches a phonological string to the position. The rules assumed by Noyer are the following:

23.

a.	n-	1pl
b.	÷	1
c.	t-	2
d.	-aani	dual pl
e.	-na	pl f
f	-uuna	pl
g.	-iina	f (2)
h.	t-	f
i.	y-	Elsewhere
j.	-u	Elsewhere

The ordering of the rules that are in a bleeding relationship (application of one precludes the application of another) can all save two be determined by the Paninian principle. Rule (a) applies before rule (f) because the features in (f) are a subset of the features in (a). Rule (e) applies before rule (h) because the features in (h)are a subset of those in e. Similarly for rules (a) and (b), (e) and (f), (e) and (g), and (e) and (h).

Two crucial orderings are not covered by this principle, however. Rule (c) must apply before rule (h), or an ill-formed word would result, as Noyer shows in the following two derivations:

24. Correct ordering:

stem:	___ aktub ___	input [2 f sg]
c.	t-aktub ___	[prefix filled and [2] discharged]
g.	t-aktub-iina	[suffix filled and [f] discharged
h.	cannot apply since prefix position is filled	

25. Incorrect ordering:

stem:	___ aktub ___	input [2 f sg]
h.	t-aktub ___	[prefix filled and [f] discharged]
c.	cannot apply since prefix filled	
g.	cannot apply since 2 was not discharged	
j.	*t-aktub-u	[suffix filled by elsewhere affix]

The ordering of rules (c) and (h) is not predicted by the Paninian principle, since they discharge disjunctive features. However, Noyer notes, if an appeal is made to the feature hierarchy, the correct ordering can be derived. In a competition between rules referring to two different features, the rule making reference to the feature higher on the feature hierarchy will apply.

The second crucial ordering that is not predicted by the Paninian principle is that between rules (d) and (e) The contrasting derivations for these two rules are shown below:

26. Correct ordering:

		input [3 f pl dual]
d.	___ aktub-aani	[suffix filled and [dual pl] discharged]
e.	cannot apply since suffix filled	
h.	t-aktub-aani	[prefix filled and [f] discharged]

27. Incorrect ordering:

		input [3 f pl dual]
e.	___ aktub-na	[suffix filled and [pl f] discharged]
d.	cannot apply since suffix filled	
i.	*y-aktub-na	[prefix filled with Elsewhere affix]

Again, the correct ordering can be assured by appealing to the Feature Hierarchy - number features are higher than gender features, and therefore rule (d), which refers to number features, will apply before (e), which refers to gender features.

The two functions of the feature hierarchy, then, are summed up by Noyer in the Feature Hierarchy Hypothesis:

28. There is a universal hierarchy of morphosyntactic features. If [F] and [G] are morphosyntactic features and F is higher on the hierarchy than [G], then:
    a) if *[F G] is active at MS, then [F G] is impoverished to [F]

b) if two spell-out rules, one referring to [F], the other to [G] and not to [F], have disjoint or overlapping structural descriptions, then the rule referring to [F] applies first.

This is similar to an observation Noyer notes has been made by Kurylowicz, as summarized in Arlotto (1972):

29. *Kurylowicz's Fifth Law of Analogy*:
In order to reestablish a central grammatical distinction, a language will abandon one that is more marginal.

In Noyer's treatment, "more marginal" translates to "lower on the hierarchy". Again, a natural correlate to this notion is available if morphosyntactic features are organized in a feature geometry. More complex representations are more marginal ones. They are less common cross-linguistically, and are subject to simplification processes at MS. Thus, this notion as well can be correlated with degree of embedding, or complexity (in terms of number of nodes), or markedness, of the representation. The ordering effect illustrated in (24) above is captured if simpler representations have priority over more complex ones - if [prt], dominated by only one node, has priority over [f], dominated by three nodes. The ordering effect in (26) is less straightforward, as the competing features [dual] and [f] are dominated by the same number of nodes. However, these rules do not discharge only one feature, but two. If the number of nodes of necessity present in the Mºs to trigger application of these two rules is compared it can easily be seen that the representation of [pl f] contains two more nodes than [pl dual] does. [Dual] is dependent on [pl], and so the representation will necessarily contain four nodes of the feature geometry. However, in order to discharge [pl f] simultaneously, as in rule (e), at least five nodes of the feature geometry need to be present in the representation. Recall that affixes, in an ideal feature-geometrical theory, are subtrees of the geometry. The only subtree containing both [pl] and [f] is the one dominated by the node NUMBER, which dominates GENDER, whose daughter is [f]. Thus, this rule refers to a representation containing at least five nodes, while rule (d) refers to one containing at least four. Rule (e) makes use of a more complex and thus more marginal representation, and it will be ordered after rule (d).[10]

It is very important to note that this notion is only called into play when the Paninian principle is not enough to determine an ordering for two rules - when the structural description of one rule is *not* a subset of the structural description for another. When it is the case that one rule's structural description is contained in another's, the more complex rule applies first, according to the Paninian principle. So the ordering between (23a) and (23b) follows because the structural description for [spkr] is contained in the structural description for [spkr pl]. A more interesting example of this effect can be seen in the ordering between (23b) and (23c). In a feature system that makes reference to separate feature [1] and [2] this ordering is not a crucial one, as a representation containing [1] will not contain [2] and vice-versa (at least in Arabic, where the filter *[1 2] has applied) and therefore the ordering of (23b) and (23c) is not crucial, as neither could incorrectly apply to discharge the features of the other. However, in the geometry as I have represented it in (19), rule (23c) will refer to the [prt] node, and unless it is ordered after (23b), will

[10]Another option for dealing with this situation is to claim that rule (e) could not apply in these circumstances anyway, given a feature geometry where [dual] was dependent on [pl]. In Noyer's system, [pl] can be discharged independently of [dual], as there is no dependency relationship between them, but in a feature geometric system, rule e. could not apply as it does not discharge a subtree of any Mº with a dual feature - if it realizes [pl] in this situation, it must needs also realize [dual], but the rule only discharges [pl]. Thus it could not apply, and no ordering relationship with rule (d) need be specified.

incorrectly apply to forms containing a [prt] node and a [spkr] node. The ordering of (c) after (b), however, falls out from the Paninian principle on this account, as the structural description of [prt] is contained within the structural description of [spkr]. For this account to work, rules therefore must contain trees representing the features that trigger their application.

The two effects derived from the Feature Hierarchy Hypothesis can then be described as follows:

30. *Filter Implementation*:
When a filter of the form [*F G] works to simplify a representation containing the features [F] and [G], the delinked feature will be the feature dominated by more nodes in the feature geometry.

31. *Rule Ordering*:
When two rules whose structural descriptions are disjoint or overlapping are contained within a single rule block, the rule whose structural description contains the fewest nodes of the feature geometry will apply first.

These two definitions can be unified if they are both made to appeal to a notion of complexity, or marginality, defined with respect to number of nodes:

32. A feature-geometrical representation X is more marginal than a feature geometrical representation Y if X contains more nodes than Y.

The two definitions above can then be reworded as follows:

30'. *Filter Implementation (revised):*
When a filter of the form [*F G] applies, the delinked feature will be the feature requiring the more marginal structural description.

31'. *Rule Ordering (revised)*:
When two rules whose structural descriptions are disjoint or overlapping are contained within a single rule block, the rule with the most marginal structural description will apply after the rule with the less marginal structural description.

5.5.4. ***3.3.3*** ***Alpha-notation***

Based on the morphology of Mam, a Mayan language of Guatemala (and also on data from Sursurunga and Cree) Noyer claims that negative values for morphosyntactic features are available for rule reference, and further, that the device of α-notation is used in MS rules. What follows is a summary of his analysis of Mam and Cree.

Mam marks its verbs with the following set of person affixes:

33.

	Erg Set		Abs Set	Enclitic
1sg	n- / w-	chin-		-(y)a

2sg	t-	Ø/tz-/tz'-/k-	-(y)a
3sg	t-	Ø/tz-/tz'-/k-	*
1excl	q-	qo-	-(y)a
1incl	q-	qo-	*
2pl	ky-	chi-	-(y)a
3pl	ky-	chi-	*

The ergative set is also used to mark nominals to indicate possession:

34. a. n- wi:xh -a *my cat*
    b. t- wi:xh -a *your* (sg) *cat*
    c. t- wi:xh *his/her cat*
    d. q- wi:xh -a *our* (excl) *cat*
    e. q- wi:xh *our* (incl) *cat*
    f. ky- wi:xh -a *your* (pl) *cat*
    g. ky- wi:xh *their cat*

The possessive nominal marking paradigm is seen below:

35.

	singular		non-singular		
[+I -you]	n-/w-	-a	q-	-a	1 excl
[+I +you]	*		q-		1 incl
[-I +you]	t-	-a	ky-	-a	2
[-I -you]	t-		ky-		3

Note that there are two patterns of distribution. The form of the prefix is conditioned by the values of the features for first person and number. In Noyer's feature system[11], they can be captured by the following rules (The various sets of affixes all fall into one pattern; any rules that can deal with one can deal with all):

36.

[+I -sg]	q-
[+I]	n-/w-
[-sg]	ky-
Elsewhere	t-

The enclitic distribution pattern is more difficult. first person exclusive and second person receive the enclitic /-a/, and first person inclusive and third person do not. This pattern could be captured with the following rules:

37.

a. [+I +you]	-Ø
b. [+I]	-a
c. [+you]	-a
d. Elsewhere	-Ø

[11]Noyer uses the feature [-sg] for marking [pl], and [+sg] for [-pl].

or, since Noyer's system includes the feature "participant":

	e. [+I +you]	-Ø
	f. [+part]	-a
	g. Elsewhere	-Ø

However, Noyer notes, these approaches have the drawback that they treat the various instances of -Ø as homophonies, not as a natural class. Further, the complement to the /-Ø/ class - the /-a/ class - behaves as a natural class in another way: the emphatic may only be formed from the forms that took an enclitic - i.e. first person exclusive and second person:

38.	a.	n- wi:xh -a	*my cat*	w-n-wi:xh	*MY cat*
	b.	t- wi:xh -a	*your*(sg) *cat*	t-t-wi:xh	*YOUR cat*
	c.	t- wi:xh -Ø	*his/her cat*	*	
	d.	q- wi:xh -a	*our*(ex) *cat*	q-q-wi:xh	*OUR cat*
	e.	q- wi:xh -Ø	*our*(inc) *cat*	*	
	f.	ky- wi:xh -a	*your*(pl) *cat*	ky-ky-wi:xh	*YOUR cat*
	g.	ky- wi:xh -Ø	*their cat*	*	

The emphatic facts in (38) mitigate against any treatment that divides the rules governing enclitic distribution into two separate classes, as the first set of rules in (37) does. If it is possible to make reference to negative values of features at MS, and use α-notation, then the problem[12] of treating the two instances of /-Ø/ as homophonous can be dealt with, resulting in the following elegant rules governing enclitic distribution:

39.	Enclitic distribution rules		
	a. [α I] [-α you]	-a	
	b. Elsewhere	-Ø	

In Cree, pronominal clitics have two forms, one for first person exclusive and another for first person inclusive and second person, as can be seen below:

40.	a.	ni-miciso-nan	
		1CL-eat-1pl indic	`We (excl) eat'
	b.	ki-miciso-nawaw	
		2CL-eat-2pl indic	`You (pl) eat'
	c.	ki-miciso-naw	
		2CL-eat-2p1 indic	`We (incl) eat'

Again, in Noyer's feature system, these facts can be captured naturally if the value [-you] is visible at MS:

41.	a. [+I -you]	ni(t)-
	b. [+you]	ki(t)-

[12]If it is a problem - see section 4.1.

3.3.4 Problems with α-notation:

The α-notation, although it enables Noyer to capture the Mam facts elegantly, forces the costly move of making negative values visible in the system, rather than assuming an underspecification-type approach to negative values. And as he notes above, the natural classes picked out by allowing such features to be visible do not in fact seem to be exploited by the morphology of the languages in which these features exist.

Further, if α-notation is available for person features, it should be available for the whole system of features. Take, for example, Noyer's number feature [+/- sg]. [-sg] is visible in most linguistic systems. In systems in which, say, [-you] or [-I] was visible like Cree or Mam, rules referring to the natural classes picked out by the following should be possible:

42. a. [α you, -α sg]
    b. [α I, -α sg]

That is, one would expect to see paradigms in which second person and first person inclusive plurals patterned together with first person exclusive and third person singulars ([+you, -sg] and [-you, +sg]), or paradigms in which first person singulars patterned together with second and third person plurals ([+I, -sg] and [-I, +sg]). The existence of such paradigms seems extremely unlikely. Positing α-notation introduces a powerful device into the system that allows reference to many classes that don't seem to exist. The alternative, that is, not allowing α-notation, results merely in analyses that require a couple of homophonous affixes. [13]

Note that the feature geometry as outlined above doesn't allow α-notation, at least for the Mam (and Susurunga) cases. Since dependency encodes contrastiveness, the person features each require that their mother have a positive, marked value. This prevents reference to [+spkr, -prt] combinations, since [spkr] is dependent on [prt] having a positive value. Thus in this system, Mam must be treated as examples of homophony (but see section 4.2 below). Alpha-notation would still be possible in this system, of course, allowing for visible negative values, but it would have to be between person and number features, or person and gender features, or number and gender features (none of which depend upon each other). As pointed out above, however, the classes picked out by such notation seem extremely esoteric and counterintuitive, and what's more important, non-existent in natural language.

1.1.5. ***3.3.4 Cross-classification of features***

Noyer notes that in order to capture a four-way distinction such as that common in person features (third, second, first excl, first incl) a maximum of two bivalent features are required, for example, [+/- I, +/- you]. This could capture the person paradigm as follows:

---

[13]Halle makes extensive use of alpha-notation in his 1993 (ref?) treatment of the Russian number/case system. Eliminating alpha-notation would therefore pose a problem for such an analysis, and further study is required to explore alternative possibilities.

43.

	-/+you	+/- I
3	-	-
2	+	-
1 ex	-	+
1 inc	+	+

However, it provides no way of characterizing the group [-third person], a morphologically salient class as evidenced by paradigms in many languages, for example, Winnebago, Lummi, and Navajo. Noyer therefore assumes a third feature, [+/- participant]. In thus introducing a third feature into the paradigm, he doubles the number of distinctions the system can represent. Since it is evident that there is never eight person distinctions in natural language, and indeed, only the person distinctions available with the four-way system outlined above and one additional distinction - that of [-third person] - are exploited in natural language, some move must be made to limit the application of the value [participant] to only situations where either the value for [I] or [you] (or both) is positive. To accomplish this, he introduces the notion of an "enhancing" feature. The analogy is to the value [+round] in phonology, in the case of back vowels: the height distinctions between different back vowels can be made without rounding, but it is cross-linguistically common for back vowels to be "enhanced" by a feature [+round].

In a feature geometry like that outlined above, three features are required to make the four-way person feature distinction, but no appeal to the notion of an "enhancing" feature need be made. By adopting a tree in which only marked values may have dependents, each additional feature introduced creates only one additional distinction. The category of [-third person] is still marked by [participant], and any Mº representing "you" or "I (inc)" or "I (excl)" of necessity has a node [+prt] to which morphological rules can refer.

1.6. 3.3.5 Visibility of negative values

Above, α-notation is dismissed as too powerful a mechanism to make available to the rules of MS. It predicts the existence of morphological classes not attested in natural language. Further, it necessitates that negative values for features be available for rule reference in some languages. If α-notation is in fact not available, as I suggest, is it still necessary for negative values to be visible at MS?

As discussed above in footnote 10, there is an alternative argument for ordering the Arabic verbal inflection rules. Rule (23e), I argue, discharges a subtree of the Mº that is marked only for [pl] and [f], not for [dual], and so could not apply to an Mº marked for [dual], which is the case in which an improper derivation could arise. This argument requires one of two assumptions. Either rules only apply to Mºs whose feature content is not merely compatible with the feature content of the rules but is in fact identical - a patently impossible move in a feature-geometrical approach, where reference to organizing nodes should automatically also refer to dependents of such organizing nodes. Or, in the rule referring to [pl f], the value [-dual] is dependent on the [pl] feature, which thus prevents it from applying to any Mº specified as [dual]. For this to be possible, of course, negative values for features must be visible to MS rules.

Noyer's analyses of Mam and Sursurunga require reference to negative values, but only in the context of an α-notation approach. Further, in his treatment of Cree, in which he invokes the value [-you],

it is only a subpart of a rule that also refers to [+I]. The negative values in these cases are never referred to independently as natural classes. Noyer notes this on p. 175:

> "Certain difficulties remain with the theory, however. Observe that in the languages which require both [-I] and [-you] to be visible (Mam and Sursurunga, for example), the rules are in both cases rules referring to both values of [I] and [you] at once. Whether the values for person features agree or not is apparently a morphologically relevant piece of information. Similarly, in the Algonquian case, reference to [-you] needs to be invoked only in a rule which also refers to [+I]. In no case are the classes [-I] or [-you] picked out as independent classes. This suggests the theory of markedness I have proposed does not fully articulate the cost of morphological representations.[14]

Given the lack of reference to negative person values as natural classes in and of themselves, it seems more likely that either morphological features are underspecified, or that they are privative.

6. 4 Residual Issues, Solutions and Problems

6.1. 4.1 Homophony

In Mam, the morphology of the person system groups [first incl] and [third] as a class, and [first excl] and [second] as a class. There are many instances in natural language where morphology seems to pick out classes that no feature system would predict to be "natural". The usual approach is to treat such morphology as instances of homophony. Take, for example, the paradigm of jussive verbal affixes from Arabic, similar to those discussed earlier:

44. Jussive singular:

y-aktub	3m
t-aktub	3f
t-aktub	2m
t-aktub-ii	2f
÷-aktub	1

The prefix /t-/ is characteristic of both third person feminine and second person. It is difficult to see how these form a natural class - it certainly can't be referred to by Noyer's feature system. His response is to treat these prefixes as instances of homophony.

6.1.1.

6.1.2. ***4.1.1 Readjustment rules eliminating homophony***

Noyer's treatment of the Arabic paradigm immediately above and earlier splits the Mº into a prefix and an affix, two positions of exponence that need to be filled. Nothing in his treatment requires that a certain type of feature be discharged in one or the other position - it is theoretically possible that any feature could fill any position.

---

[14]See also later discussion of person features in section 4.4.

Halle (p.c.) has pointed out that the distribution of features in the prefixes and affixes of these paradigms appears to be somewhat less random than Noyer's theory would predict, however. No person features are discharged in the suffix position whatsoever, and the only non-person feature that is discharged in the prefix position is the [f] feature whose affix is the very /t-/ that is homophonous with the person affix /t-/ that discharges the feature [2]. This suggests, he maintains, a readjustment rule of the following form:

45. f → 2 / ___ third person

If this rule applies, then the instances of [t-] are no longer homophonous, but are both merely spell-outs of the discharged feature [2].

Such a rule mechanism is clearly not constrained enough to account for such data in a principled way. A simple rewriting mechanism could account for any sort of homophony. Would English plural marking and genitive marking map to third sg. present verbal marking (/-s/, /-z/, /-\z/)? If each phonological sequence was the result of only one spell-out rule, would lexical items map to other homophonous lexical items? (DEER → DEAR → /di:r/?) Such a powerful mechanism clearly overcompensates wildly.

The entire issue raises another troublesome point, however. The very gaps in paradigms that represent the action of morphological filters are not in fact blank - they are merely homophonous with another form. How can one distinguish in a principled way between accidental homophony and homophony that is the result of Impoverishment? And if such rules eliminating "accidental" homophony are adopted, like that in (45), how could one distinguish between such rules and the effect of filters? Such rules would have to be powerful enough to duplicate the effect of filters, since they are mapping features onto other features, (even if they were not as powerful as a simple rewriting rule like that in (45), and would introduce huge redundancies into the system. It therefore seems more reasonable to allow limited homophony in cases where the groupings of features are unattested elsewhere in natural language, as in the Arabic case - [2] and [3 f] do not form a natural class in any other paradigm I know of.

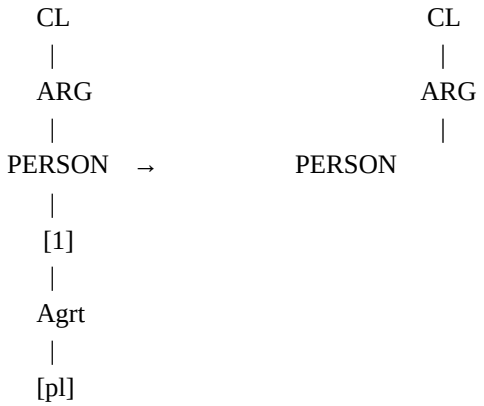
The question then arises - are the cases of Mam and Cree, which Noyer treats as instances where negative values are visible, in fact instances of accidental homophony? Below, I answer this in the negative.

1.1.3.

1.2. 4.2 Mam, Cree and Bonet-style Impoverishment

Bonet proposes impoverishment for clitics of a slightly different type than that described by Noyer. Clitics whose trees are subject to Impoverishment do not lose merely the offending extra feature, but a large portion of the geometry, rendering them completely unmarked and subject to the Elsewhere rule, as in (46) below:

46.



The effect of this type of filter is to render highly complex forms homophonous not with slightly less complex forms, as in the Arabic cases listed by Noyer above, but with the unmarked, or elsewhere class. A cross-linguistic characterization of the effect of this type of filter can be seen in (47) below:

47. Reflexive affixes or clitics in several languages:

	Russian	Papago	Walpiri	Catalan	Piedmt.	?	Valenc.
1st sg.	i		spec.	spec	spec	spec	spec
pl.	i	spec	i	spec	i	spec	i
		spec					
2nd sg.	i		i	spec	spec	spec	spec
pl.	i	i	i	spec	spec	i	i
		i					
3rd sg.	i		i	i	i	i	i
pl.	i	i	i	i	i	i	i
		i					

In the above table, "i" stands for "invariant" or the result of the application of the Elsewhere rule, and "spec" for "special", indicating the existence of a special rule spelling out a different form for these feature combinations.

Evidently, these filters must have some sort of classificatory marking on them, as they apply only to clitics, and not to, say, agreement features - Russian has pronouns and agreement marking that make distinctions for all of the above feature combinations.

Given that filters can have such marking, Noyer's argument for α-notation from the Mam data above loses much of its force. The elements whose distribution must be captured using α-notation are clitics, as can be seen in the paradigms and rules (33) and (39). If, instead, we assume a Bonet-type Impoverishment for these elements, the clitics with the most complex feature geometry will fall together with the less complex clitics in the Elsewhere class.

Thus, a rule like (48) will apply to the clitic cluster that includes a first person inclusive node in Mam, Impoverishing it and causing it to fall together with the Elsewhere class:

48. PERSON → PERSON
|
prt
|
spkr
|
inc

Then the following spell-out rules can apply:

49. Mam: Enclitic distribution
a. [+prt] →/-a/
b. Elsewhere → /-Ø/

Because the inclusive forms have undergone radical Impoverishment, like the plural forms in the table in (47) above, they are distinct from the other first person forms.[15]

1.3.

1.4. 4.3 The gap in Bonet's table

Bonet lists the patterns of clitic distribution from many languages in the table in (47), and maintains that the pattern represented by the column headed "?" is a logical possibility, given her feature geometry, but that she hasn't come across it. Given the feature geometry posited in (19), this gap is predicted. A filter that Impoverishes second person plural forms but not first person plural forms is impossible, given that such a rule would have to refer to representations containing a [prt] node, and representations of both first pl and second pl contain such a node. Any pattern in which second pl is Impoverished will of necessity have first pl Impoverished as well, like the Valencian pattern.[16]

---

[15]The treatment of Potawatami clitics in Halle and Marantz (1993) also raises problems for this particular feature geometry. The distribution of the clitics is characterized by the rules in i. below:

i. Potawatami Clitics
a. [+2] → /k-/
b. [+1] → /n-/
c. [-obv] → /w-/

The problem arises in that the ordering of the rules realizing [2] and [1] is precisely the opposite of that predicted by the feature geometry (see the discussion in 3.3.2 above). (A similar problem arises for Cree, as can be seen from (41) above - the spell-out of the 3rd person clitic will be the elsewhere class, not that of the 2nd person clitic.) The solution advocated for Mam clitics above doesn't seem immediately possible here, because the affix that occurs with [1] arguments is not the Elsewhere affix - radical impoverishment would result in ordering c. before b. above. Closer study is required to discover if such an ordering would be problematic, and to create a tree that can accomodate the natural class captured by the feature [-obv] in c. above.

[16]Harris (p.c) points out that Latin-American Spanish exhibits just the pattern of clitic distribution that the feature geometry predicts to be impossible, above - that is, the second person pl forms are Impoverished while 2sg and first pl forms are not. However, this pattern holds not only for the clitics, but for the rest of the morphemes of the language as well - the Agreement features, and the pronominal system. This Impoverishment is therefore not of the clitic type, but of a more general type. Nonetheless, it is still predicted to be impossible for this feature system.

7. 5 Conclusion

I have shown that the effects of Noyer's feature hierarchy can be derived from a feature-geometric representation of morphological features, by translating relative prominence in the hierarchy into a notion of geometrical markedness, sketching an alternative treatment of the Arabic verbal paradigm. I argue that a feature-geometrical approach has the advantage of capturing cross-linguistic groupings of features and deriving the necessary rule-ordering and filter-constraining effects from the representation rather than by employing an additional mechanism. Further, by classifying clitic-type realizations of morphosyntactic features separately from inflectional-affix-type realizations and adopting Bonet-style Radical Impoverishment I eliminate the use of the over-powerful α-notation device, which Noyer uses to account for Mam enclitic distribution.

References:

Anderson, S. (1992). *A-morphous Morphology*. Cambridge: Cambridge University Press.

Anderson, S. (1984) `On Representations in Morphology: Case, Agreement and Inversion in Georgian', *Natural Language and Linguistic Theory* 2, 157-218.

Arlotto, A. (1972), *Introduction to Historical Linguistics,* University Press of America: Lanham

Bonet, E. *Morphology After Syntax*. MIT PhD dissertation. Cambridge: MITWPL.

Calabrese, A. (1988) *Towards A Theory of Phonologcial Alphabets*. MIT PhD dissertation. Cambridge: MITWPL.

Carnie, A. (1994) `Whence sonority? Evidence from epenthesis in Modern Irish,' *Papers in Phonology and Morphology*, MITWPL 21, Cambridge: MITWPL.

Clements, G.N. (1985) `The Geometry of Phonological Features'. *Phonology Yearbook* 2, 225-252.

Halle, M. and S. Bromberger. (1989) "Conceptual Issues in Morphology". ms., MIT.

Halle, M. (1992) "The Russian Declension: An Illustration of Distributed Morphology," to appear in *The Organization of Phonology: Features and Domains*, CSLI, Stanford University.

Halle, M. and A. Marantz. (1993) "Distributed Morphology and the Pieces of Inflection," in K. Hale and S.J. Keyser, eds., *The View From Building 20: Linguistics Essays in Honor of Sylvain Bromberger*, Cambridge, MA: MIT Press, 111-176.

Hockett, C. F. (1966) ` What Algonquin is really like.' *International Journal of American Linguistics.* 32:1 pp 76-83.

Lumsden, J. (1987) *Syntactic Features: Parametric Variation in the History of English.* MIT PhD dissertation. Cambridge: MITWPL

Lumsden, J. (1992) `Underspecification in grammatical and natural gender', *Linguistic Inquiry* 23:3 pp 469-486.

Marantz, A. (1992) `What kind of pieces are inflectional morphemes?' BLS February 1992.

Noyer, R. (1992) *Features, Positions and Affixes in Autonomous Morphological Structure*, MIT PhD dissertation. Cambridge: MITWPL

Rice, K. & P. Avery. (1991) *Segmental Complexity and the Structure of Inventories*. Paper presented at the GLOW workshop on the acquisition of phonology.

Rice, K. (1992), `On deriving sonority: a structural account of sonority relationships', *Phonology* 9: 61-99.

Sagey, E. (1986) *The Representation of Features and Relations in Non-Linear Phonology*. PhD Dissertation, MIT. Distributed by MITWPL

Silverstein, M. `Hierarchy of features and ergativity.' In *Grammatical Categories in Australian Languages,* ed. R.M.W. Dixon, Australian Institute of Aboriginal Studies, Canberra, pp 112-171.

Stump, G. T. (1989) `A note on Breton pluralization and the Elsewhere Condition', *Natural Language and Linguistic Theory* : 261-273

Dept of Linguistics and Philosophy
20D-219, MIT
Cambridge MA 02139

hharley@mit.edu