Phonologization and the typology of feature behavior

Keywords: features, typology, phonologization, assimilation, dissimilation, coarticulation

4781 words

Jeff Mielke
Department of Linguistics
University of Ottawa
Arts 401
70 Laurier Ave East
Ottawa, Ontario K1N 6N5
Canada
jmielke@uOttawa.ca
(613) 562-5800 ext. 2701

1 Introduction

One of the successes of distinctive feature theory has been the identification of a number of phonetically defined features which are able to describe groups of sound that are phonologically active in many unrelated languages. This study measures the crosslinguistic frequency of occurrence of classes defined by particular features which have been proposed and examines the phonological behavior of these classes. The characteristic behavior profiles of particular features are explored in terms of two approaches to feature effects, one of which draws on representations for explanation, and the other of which draws upon phonologizable phonetic effects for explanation.

Innatist approaches to distinctive features have accounted for the crosslinguistic recurrence of particular types of sound patterns by building crosslinguistic generalizations into the representations used for phonological patterns (Chomsky and Halle, 1968; Clements, 1985; Sagey, 1986[1990]). In this view, representations are explanatory. Recurrent classes are definable using innate features, and the behavior of particular classes is attributed to the organization of the mental representation of phonology. The phonological activity of only some logically possible classes of sounds in sound patterns is accounted for by positing that only the features which define the active classes exist in a universal feature set. More specific observations are accounted for by positing that certain feature values do or do not exist, and that features are organized in a hierarchy which restricts the ways in which they can interact. This approach may be summarized with the slogan *"Things happen because of features."*

Another view is emergent features (Mielke, 2008), in which feature effects are accounted for in terms of the historical development of sound patterns, as in *Evolutionary Phonology* (Blevins, 2004) accounts of markedness generalizations and other patterns. Recurrent phonologically active classes are defined by features whose phonetic correlates are involved in commonly-phonologized phonetic effects, and the behavior of particular classes is attributed to the nature of the phonetic effects from which they developed. This approach can be summarized with the slogan *"Features happen because of things."*

These two approaches to feature effects are schematized in Figure 1. In innate feature theories, recurrent sound patterns are built out of distinctive features from the universal feature set, which are in turn grounded in phonetics, so that features serve as a link between phonetics and sound patterns. In the case of sound patterns which are not easily captured using a particular feature set (e.g. sound patterns involving unnatural classes of sounds), recourse can be made to phonetic effects or historical accidents (the dotted line connecting "sound pattern" and "phonetics"). In emergent feature theory, this is the only connection between sound patterns and phonetics, i.e. all sound patterns are historical accidents, but some of these accidents, such as the phonetically natural ones which form the primary data for innate feature theories, are more frequent than others. Features, in emergent feature theory, are posited by learners in response to observed sound patterns.

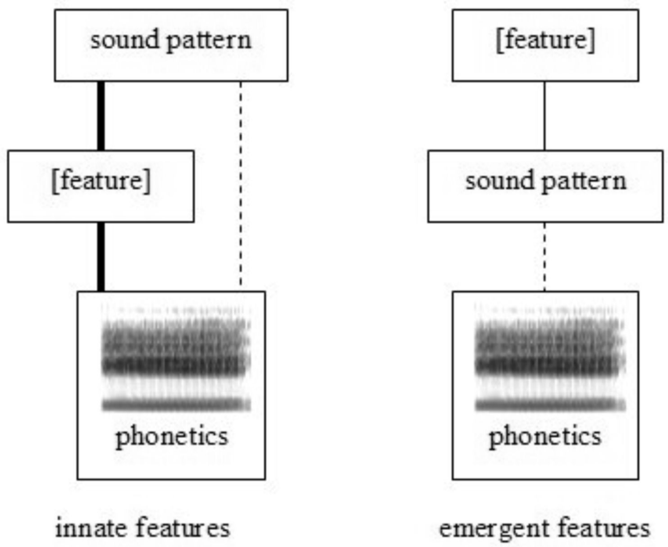

Figure 1: Relationships between phonetics, features, and phonological patterns (Mielke, 2008, 8)

The purpose of the present study is to investigate whether the features which are frequently required to describe sound patterns appear to be attributable to frequently phonologized phonetic effects, i.e. to the dotted lines in Figure 1. Among the goals is to tease apart features from their phonetic correlates as sources of explanation. Emergent feature theory predicts that features are needed in rules only insofar as they are related to the origin of the sound pattern they are involved in. This is a prediction of *any* approach to phonology in which sound patterns are attribute to phonologized phonetic effects rather than to synchronic representations.

To explore the connection between phonetic effects and features involved in sound patterns, the features needed to describe a sample of phonologically active classes were counted and categorized according to their behavior. The next sections describe how this counting and classification were performed and what the results show for the questions such as how often particular feature values are active in phonological patterns, and what the classes defined by these features typically do, and what aspects of these patterns can be accounted for in terms of the relationship between the features' phonetic correlates and phonologization precursurs.

2 Methods

Counting and categorization of classes defined by particular features were conducted on the sound patterns included in P-base[1] (Mielke, 2008), a database of sound patterns found in language grammars available on library shelves. It includes 628 language varieties, which are grouped into 549 languages. Dialects were considered to be one language if they shared an entry in *Ethnologue* (Grimes et al., 2000). All phonologically active classes involving more than one but fewer than all of the segments in an inventory reported in these grammars were recorded. The definition of a phonologically active class given in (1) is based entirely on phonological patterning, as opposed to the traditional definition of natural class which also involves phonetic or featural naturalness.

(1) Phonologically active class (Mielke, 2008, 48-49): any group of sounds which, to the exclusion of all other sounds in a language's inventory, do at least one of the following:
   a. undergo a phonological process;
   b. trigger a phonological process; or
   c. exemplify a static distributional restriction.

See Mielke (2008, chapter 3) for a more detailed description of the survey methods.

The 6077 phonologically active classes matching (1-a,b) (excluding the static distributional restrictions) were classified according to features based on those proposed in *The Sound Pattern of English* (Chomsky and Halle, 1968), listed in (2). The features [syllabic], [long], and [extra (long/short)] were used to capture prosodic distinctions which are not considered to be the responsibility of the segmental feature system, and the classes defined by these features are not discussed here.

[1]P-base is freely available from *http://www.oup.com/uk/companion/mielke*

(2) Features used for categorization of classes and changes:

[consonantal]	[anterior]	[delayed primary release]
[vocalic]	[distributed]	[delayed release of secondary closure]
[sonorant]	[strident]	[glottal (tertiary) closure]
[continuant]	[lateral]	[heightened subglottal pressure]
[voice]	[back]	[movement of glottal closure]
[nasal]	[low]	
[tense]	[high]	([syllabic])
[coronal]	[round]	([long])
[covered]		([extra (long/short)])

Features are being used as a familiar descriptive labeling convention in order to group together different phonologically active classes of sounds from different languages, for the purposes of counting them. The use of features for classificatory purpose is orthogonal to the question of whether features are primitives in phonological patterns. The particular feature set in (2) was chosen because it was able to represent the greatest number of phonogically active classes using conjunctions of distinctive feature values (Mielke, 2008, ch. 7), compared to features from *Preliminaries to Speech Analysis* (Jakobson et al., 1952) and Unified Feature Theory (Clements and Hume, 1995). Some of these features are no longer widely used, but this is of little consequence for the present study, because the primary concern of this study is the behavior of very frequent classes, the features which are easily handled, often in the same way, by many different feature systems.

Featural descriptions for phonologically active classes were generated by an algorithm which constructs a feature matrix for the inventory and produces the minimal set of feature values which can define the class, if any such class exists. 4313 (71.0%) of the 6077 classes could be represented by a conjunction of the features in (2) (Mielke, 2008, 147), and these classes are considered further. In addition to defining the classes automatically, the changes involved in the sound patterns were defined featurally by hand.

Occurrences of features were categorized according to the types of behavior in (3). These examples illustrate the four types of feature behavior using [+voice] as an example. In (3-a), [+voice] is involved in the change and also present in the environment triggering the change. This is classified as *spread*. In (3-b), [+voice] is involved in the change, and its opposite value ([−voice]) defines the environment triggering the change, so this would be classified as *dissimilation*. In (3-c-(i)), [+voice] defines a class of sounds undergoing a change but is not involved in the change itself, making this an example of a feature being used to *partition* an inventory into undergoers and nonundergoers of a sound pattern. Partitioning an inventory into triggers and nontriggers of a sound pattern without being involved in the change is also classified as partitioning, as in (3-c-(ii)), where [+voice] sounds trigger a change which is unrelated to voicing. Any use of a feature that does not fit into one of these three categories, such as being involved in a change that is neither dissimilatory nor assimilatory, as in (3-d), is classified as *other*.

(3) Categorizing feature occurrences by phonological behavior

a. Spread: [−son] → [+**voice**] / ___ [+**voice**]
b. Dissimilate: [−son] → [+**voice**] / ___ [−**voice**]
c. Partition:
   (i) [+**voice**] → [+cont] / [+syl] ___ [+syl]
   (ii) [+cont] → [+strident] / ___ [+**voice**]
d. Other: [−son] → [+**voice**] / ___ [+cont] etc.

3 Results

The results are presented here in terms of the features used to define classes and changes, beginning with a look at the frequency of the most frequently-used features, and proceeding to their behavior.

Figure 2 shows the 18 features that are used in the descriptions of the most sound patterns, and the activity of their + and − values. The dark bars represent cases where a single feature value defines a class or change, and the light bars represent cases where the feature value is used in conjunction with other feature values to define a class or change. [voice] and [high] are used the most, followed by [continuant], [back], [nasal], and [sonorant]. Occurrences of [voice] are divided roughly equally into [+voice] alone, [−voice] alone, [+voice] as part of a larget feature bundle, and [−voice] as part of a larger feature bundle. Occurrences of high, however, are dominated by cases where it is used as part of a larger bundle (such as [+high, +vocalic]).

The features are sorted according to the total number of occurrences of either feature value, although it is apparent that some features are more symmetrical than others in the occurrence of their + and − values. [voice] is fairly symmetrical, but other features are used more for one value than for the other. For instance, [+nasal] is more than twice as frequent as [−voice], and [−sonorant] and [+distributed] are much more frequent than their

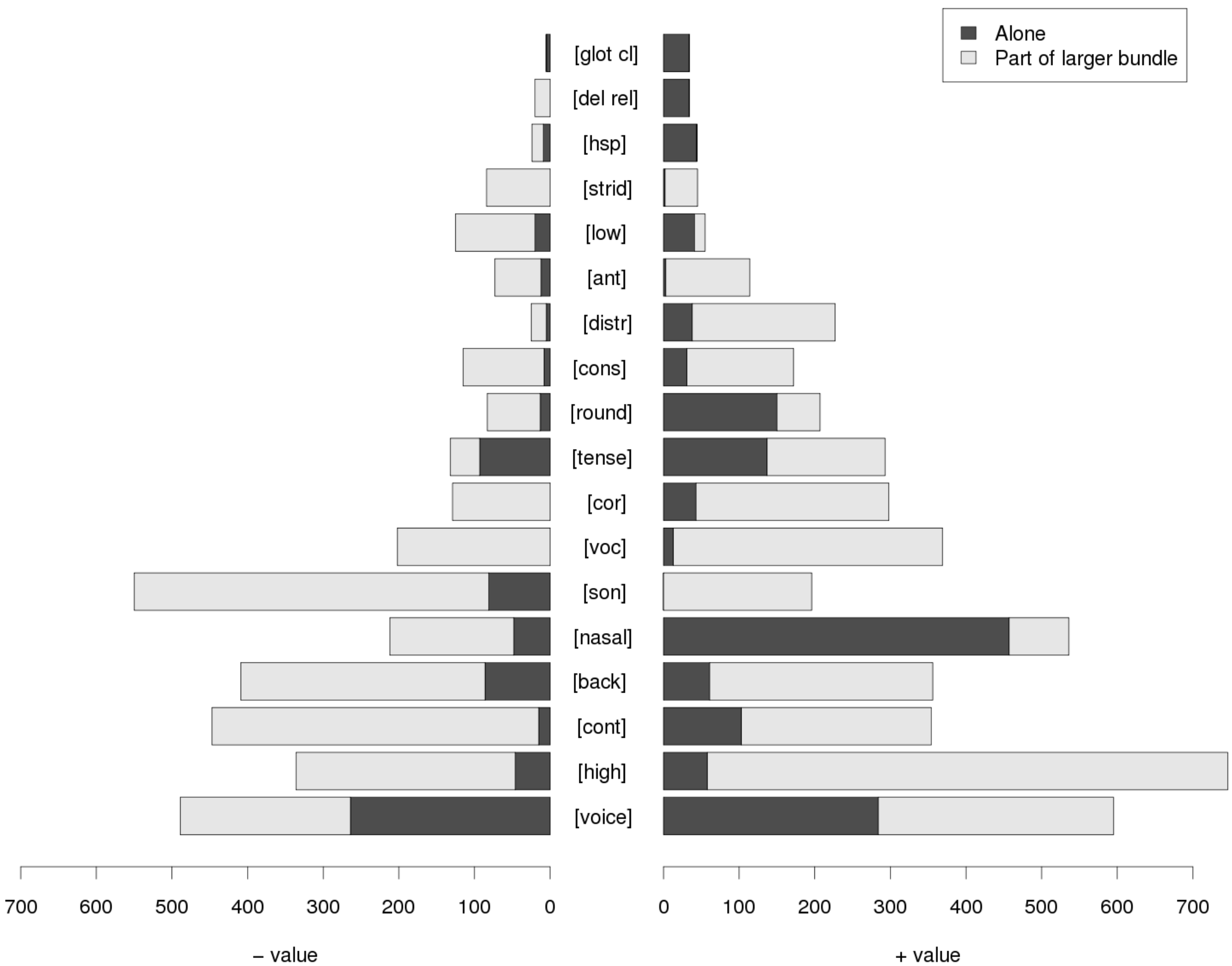

Figure 2: The most commonly-used features

counterparts.

Figure 3 shows the number of occurrences of spreading, dissimilating, partitioning, and other behavior for classes defined using the most frequently-used feature values. The chart is based on *all* of the occurrences of each value, not just the ones where the feature is used by itself. Both are interesting things to count. Counting only single-feature bundles (instances where a feature is used by itself) yields more differences between features, but counting all occurrences of each feature (as in the figure) provides results that are more applicable to whether phonetic effects can account for the need for a particular feature, because the figure shows *all* of the instances where the feature is needed to describe a phonological pattern.

As seen in Figure 3, much of the spreading is concentrated among a small number of feature values, particularly [±voice], [+nasal], [+high], and [±back]. Other feature values, such as [±sonorant] and [±continuant] seldom if ever spread, but are required in other capacities (partioning, other). Dissimilation is much less frequent than assimilation, and is concentrated primarily among [−nasal] and [−voice]. All of the features are used extensively in partitioning, and there are features, many of them major class features, which do almost nothing else.

Figure 4 shows the rates of spreading, dissimilating, and partitioning for the same features, as percentages of the occurrences of each feature. The overall average rates of spreading, dissimilating, and partitioning, and other behavior are indicated. Some feature values, such as [+round] and [+distributed] have high rates of spreading, although they do not account for a very large proportion of all the instances of spreading, because the feature values are used less overall. Figure 5, in the appendix, shows the behavior of single-feature bundles.

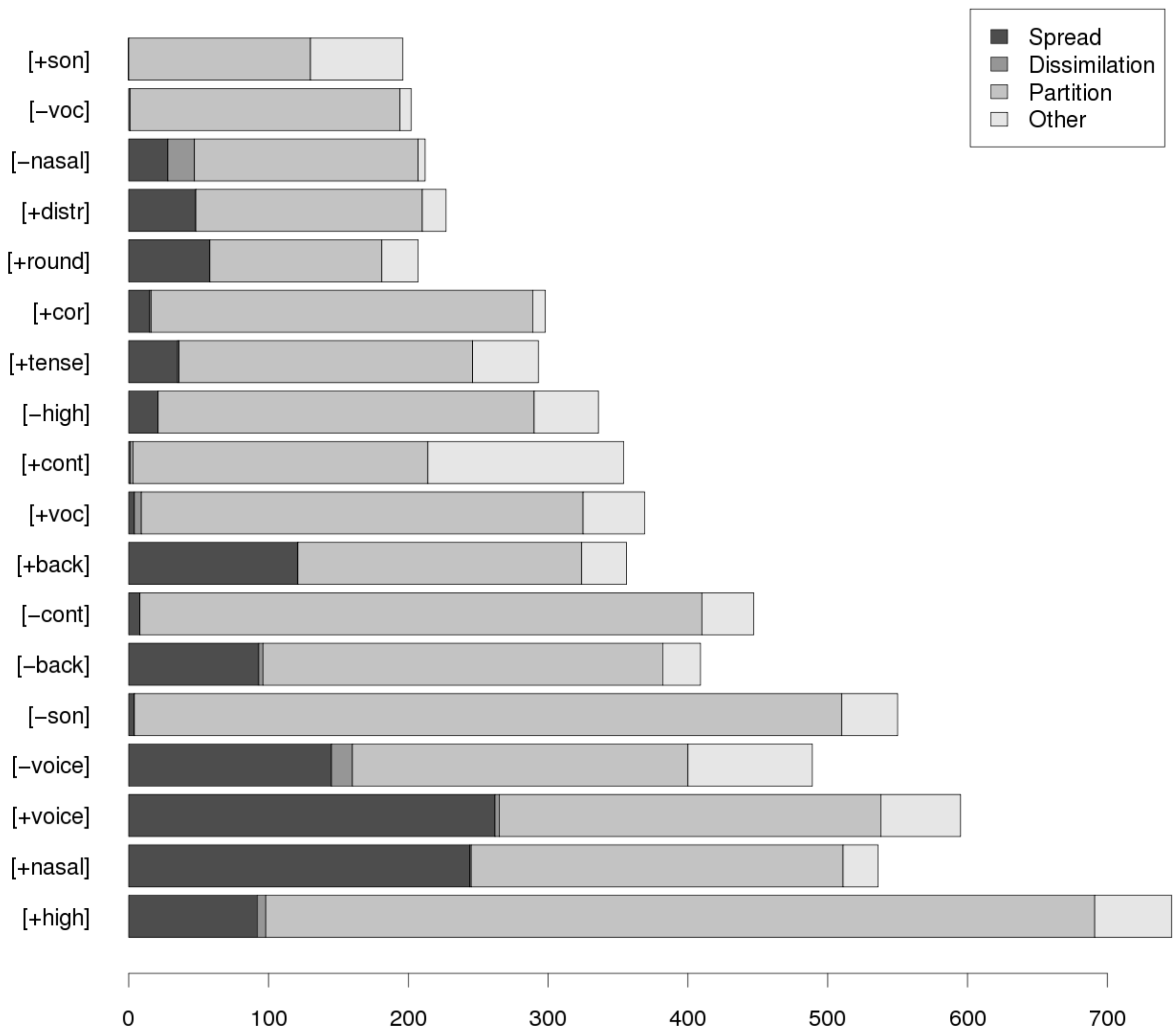

Figure 3: Total number of occurrences of each type of feature behavior for the most frequent feature values

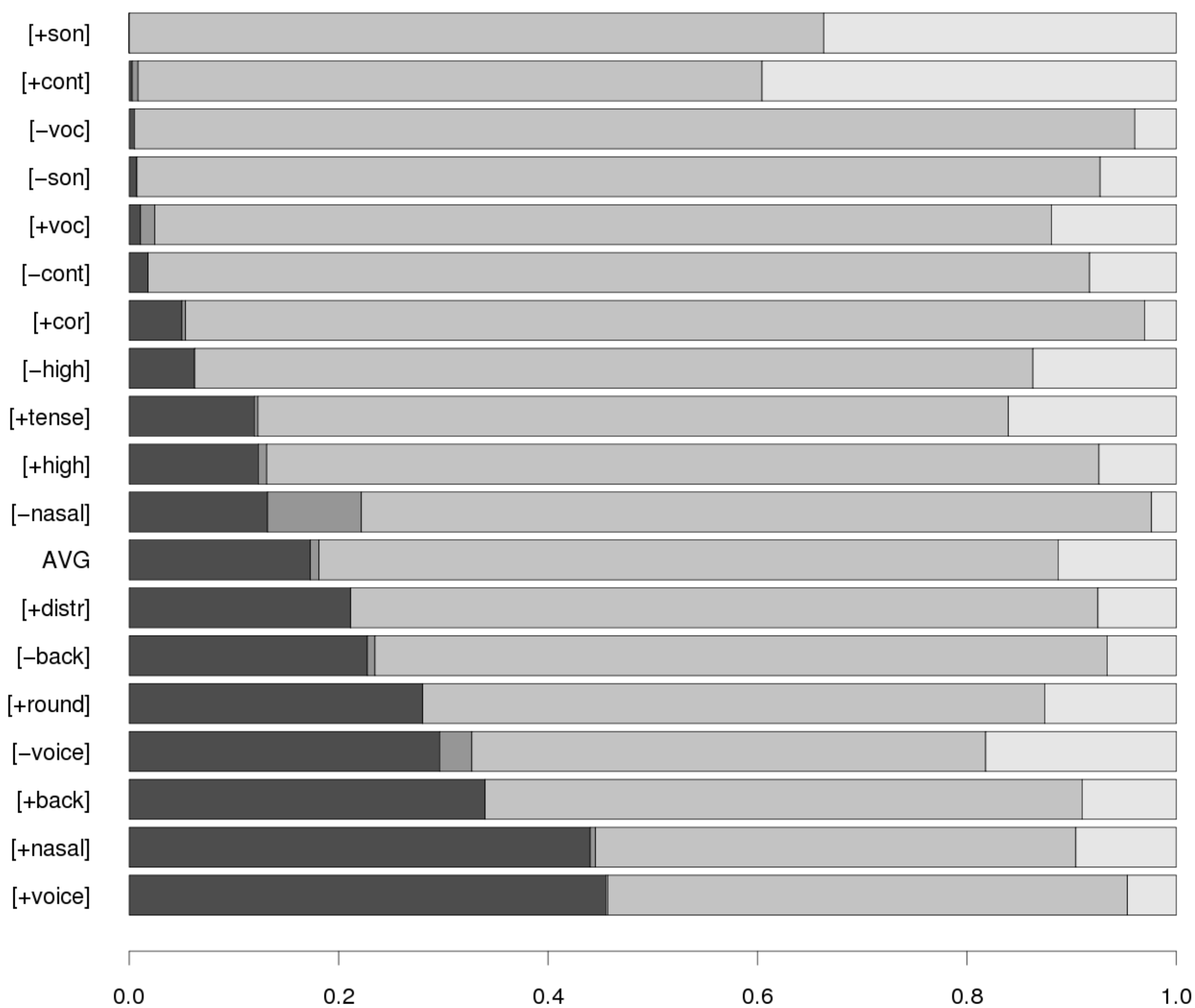

Figure 4: Common features sorted by spread rate (showing spreading, dissimilation, partitioning, and other behavior)

4 Discussion

There are several ways in which the results show different behavior for different features. These are differences in spreading, dissimilation, partitioning, the way particular values of the same feature behave.

4.1 Spreading values vs. non-spreading values

In representational approaches, spreading has been taken as particularly strong evidence that a feature exists, and features that do not seem to spread have had their existence called into question. For example, Hume and Odden (1996) argue that [consonantal] lacks clear cases of being involved in a change, and that instances where the feature has been used to define segmental contrasts or define classes involved in patterns (partitioning) can be reanalyzed in terms of other features. One practical reason for the favored status of spreading (or otherwise defining a change) in establishing the need for a feature is that changes, particularly changes appearing to involve only one feature, are difficult to reanalize, while there are often many alternatives for defining a phonologically active class or a segmental contrast.

One thing that is clear in the results above is that most of the work done by features involves partitioning rather than spreading, and many feature values which are not involved in spreading are needed for their role in defining classes. The classes involved in phonological patterns most likely cannot be reduced to just the spreading feature values in Figure 3. Therefore the distinction between robustly-spreading features and non-spreading features cannot be handled by positing a universal feature set which only includes features with evidence of spreading. This distinction can, however, be made by appealing to the origins of sound patterns involving spreading, if spreading feature values have phonetic correlates which are involved in coarticulatory effects which are precursors for phonologization. In other words, if assimilation is treated as phonologized coarticulation, then it is expected that some feature values (those whose phonetic correlates are prone to noticeable coarticulation, perhaps those with long, simple gestures that can overlap a lot of segments and be prone to misparsing) will be spread more often in phonological patterns. Other feature values may be necessary for other purposes but simply not be involved in sound patterns in this way.

The feature values which are resonsible for most of spreading depicted in Figure 3 ([±voice], [+nasal], [+high], and [±back]) are known to be involved in coarticulatory effects, unlike many of the relative non-spreaders such as [±sonorant] and [±continuant]. By attributing the difference between spreading and non-spreading values to the phonetic effects which tend to get phonologized, it is possible to account for this distinction while leaving a role in the theory for other features.

The features which do the most partitioning as opposed to spreading are [±sonorant], [±continuant], and [±vocalic]. These features are frequently used to define phonologically active classes, partioning inventories into undergoers and non-undergoers or triggers and non-triggers. These features have a history of being difficult to find precise phonetic definitions for. For example, Kenstowicz and Kisseberth (1979, 21) observe that some features are hard to define phoneticaly but are still necessary to describe sound systems:

> There are no truly satisfactory articulatory or acoustic definitions for the bases of these two different partitions [consonant and sonorant]. Nevertheless, they are crucial for the description of the phonological structure of practically every language.

Chomsky and Halle (1968, 318) observe that it is not obvious how laterals should be defined with respect to the feature [continuant], and Mielke (2005) gives evidence that they pattern both ways, and that what has been treated as a single feature [continuant] is perhaps better treated as a bundle of related phonetic parameters opposing stops to fricatives and/or vowels but treating phonetically ambiguous sounds (e.g. laterals and nasals) differently. This is consistent with Kenstowicz and Kisseberth (1979)'s comments about [consonantal] and [sonorant]. Since there is often more than one way to define a class, crucial evidence about the definitions of partitioning features is hard to find. Consequently, it is more difficult to make a case for the universality of partitioning features, and Mielke (2005) argues that [continuant] is only as predictable crosslinguistically as the phonetic properties of the sounds involved, and that the feature effects which are the most consistent across languages seem to be the ones with the most direct phonetic basis.

The connection between being hard to define phonetically and not spreading can be treated as an issue of analysis, i.e. that it is easier to identify phonetic correlates of features that spread, because the unassimilated and assimilated segments can be directly compared. When features define a partition rather than spread, it is often in conjunction with other features, and often there are multiple alternative feature bundles which can define the same class.

The murkier phonetic dimensions are also less straighforwardly involved in coarticulation or less straighforwardly reinterpreted as phonological patterns when they are involved in coarticulation. The spreading-partitioning distinction coincides with Halle and Stevens

(1991)'s distinction between articulator-bound and articulator-free features. It is potentially harder to have coarticulation which corresponds to a feature that is not tied to a particular articulatory implementation. McCarthy (1988) accounts for the lack of spreading by [consonantal] and [sonorant] by placing them inside the root node in feature geometry, so that they do not project their own tiers.

So far the class of partitioning features have been identified negatively, in their failure to be spread in assimilatory patterns (due perhaps to the failure of their phonetic correlates to be involved in coarticulation or result in phonologization) and the failure of the classes they define to submit easily to unambiguous feature analysis. There are also positive reasons for being involved in partitioning, such as when a feature's phonetic correlates cause sounds to be more susceptible to certain changes. [−sonorant] has such a relationship with voicing: certain supralaryngeal configurations require more effort to produce vocal fold vibration, amplifying the tendency of voicing to spread, and the cases of [−sonorant] partitioning include a lot of cases involving [voice].

[±sonorant] is used 665 times as part of a larger feature bundle to define classes which almost exclusively define partitions rather than spreading. This is consistent with other major class/manner features with murky phonetic correlates. [+sonorant] is never used by itself to define any class or change (i.e. the class of sonorants is never active). [−sonorant] is used by itself 81 times to define the active class of obstruents. Close to half of these involve voicing or devoicing. In 32 cases the class is voiced or devoiced, and in six cases the class triggers devoicing. In addition to having reasons not to spread, [sonorant] has particular reasons *to* partition. This also accounts for part of the asymmetry in the behavior of the + and − values: [−sonorant] is associated with phonetic properties that have a particular affect on voicing, and [+sonorant] is not.

In proposing the feature [sonorant], Chomsky and Halle (1968, 300-302) draw a distinction between spontaneous voicing (for sonorants) and non-spontaneous voicing (for obstruents). A cluster of a voiced sonorant and a voiceless obstruent is a cluster containing two segments with voicing specifications that are most compatible with their supralaryngeal configuration. A (non-spontaneously) voiced obstruent, which requires more effort to voice, is more likely to affect or be affected in a cluster with a voiceless obstruent. See also Halle and Stevens (1967); Westbury and Keating (1986); Jansen (2004).

4.2 Dissimilating feature values

Another finding is that the most frequent dissimilating feature values are the opposites of the most frequent assimilating feature values, which provides support for Ohala (1981)'s account of dissimilation as the phonologization of mistakenly undone assimilation. As shown in Figure 3, [−nasal] and [−voice] account for 34% and 27% of the cases of dissimilation, respectively. These are cases where a class of segments become denasalized next to a nasal segment or devoiced next to a voiced segment. Ohala (1981) argues that dissimilation arises as the result of a listener misapplying a correction for coarticulation or assimilation that did not actually occur. This means, for example, that dissimilation to [−nasal] would arise as a result of a listener hearing a sequence such as [ãn] and understanding it as /an/ with coarticulatory or assimilatory nasalization, and subsequently pronouncing it as [an] in careful speech, thus "undoing" nasalization which did not occur in the first place. In order for a listener to attempt to correct an apparently coarticulated sequence, there must be reason to suspect the coarticulation in the first place. If this is right, then dissimilation is dependent on assimilation or coarticulation, because it results from the misapplication of a correction for assimilation or coarticulation. The fact that the two leading dissimilating feature values ([−nasal] and [−voice]) are the opposites of the two leading *assimilating* feature values is consistent with Ohala's account of dissimilation. Since the most assimilation happens with [+nasal] and [+voice], then it is not surprising for the misapplication of the correction for assimilation to happen, as it does, with [−nasal] and [−voice].

4.3 Different behavior of different values

The asymmetrical behavior of + and − feature values may also have accounts in terms of phonetic effects that can become phonologized. For example, while [+voice] and [−voice] are involved in similar numbers and types of sound patterns, [+sonorant] and [−sonorant] are asymmetrical, in that the class [−sonorant] is active in many languages, but the class [+sonorant] is not. Both feature values are used as part of larger bundles, but the − value is use more often. The observation that [+nasal] is more active than [−nasal] and [−sonorant] is more active than [+sonorant] could potentially inform feature models in which one pole of an opposition is treated as marked and the other as unmarked. Asymmetrical behavior has also been accounted for by positing that certain features have only one value, although this leaves the question of why some features have different numbers of values than others in the first place (see Flemming (2002, 131-32) for discussion), and does not address relative activity, since the absence of a value predicts *no* activity. In the case of [sonorant], an explanation is that the phonetic correlates of [−sonorant] are associated with patterns such as voicing assimilation, while the phonetic correlates of [+sonorant] do not predispose features to such patterns. However, while [+sonorant] is less active than its opposite value and is involved in no sound patterns on its own, it is still involved (in conjunction with other features) in defining phonologically active classes. Appealing to the relationship between phonetic effects and feature behavior allows for a more nuanced account of feature behavior than is available from positing the presence or absence of particular feature values.

4.4 The appearance of universality

The similar patterning of segments in unrelated languages has given the impression of a universal feature set. Drawing upon phonetic effects for explanation, the answer to the question of why there seem to be universal features varies from feature to feature. [+voice] is needed in language after language in large part due to the tendency for the vocal fold vibration of one segment to overlap another segment. This is seen in the large number of cases of assimilation involving this feature value. [−sonorant] is active in many languages because certain supralaryngeal configurations require more effort to produce vocal fold vibration, amplifying the tendency of voicing to spread, and making [−sonorant] a frequent partitioning feature. [+high] is common for different reasons, including the tendency for tongue height gestures to overlap, and the tendency for non-sonorous vowels to be reinterpreted as glides. Features are good for referring to groups of sounds that have the same behavior within a particular language. What the sounds do depends on what sound patterns they are likely to be involved in in the first place. Tracing the appearance of universality to particular phonetic effects is consistent with the observation that in synchronic patterns the segments with the most consistent behavior with respect to a particular phonetic dimension are the segments which are most unambiguous with respect to that dimension (Mielke, 2005, 2008).

4.5 Parallels in sound pattern typology

Accounting for the behavior of different features in terms of the phonetic effects which lead to particular sound patterns has parallels in the explanation of different types of sound patterns in terms of their development or phonetic basis. For example, assimilation often involves recurrent phonetically-related sets of features, which motivated feature hierarchies in which articulatorily-independent features are segregated from one another (Clements, 1985). The feature hierarchies can be interpreted as models of the articulators involved in the coarticulation which could give rise to assimilatory patterns. Dissimilation is usually structure-preserving, because it is driven by listeners who are positing plausible lexical representations as they mistakenly undo assimilation (Ohala, 1981), and metathesis is also usually structure-preserving, for the same reason (Blevins and Garrett, 1998; Hume, 2004). Epenthesis is usually traceable to a few natural sound changes (Blevins, to appear).

5 Conclusion

In summary, the survey has provided evidence that different features have different behavior, which in many cases can be attributed to their phonetic correlates' involvement in phonologization precursurs. This is expected if features are abstractions from sound patterns (Mielke, 2008): different features have different reasons for existing, but it is surprising if features are treated as explanatory primitives. Understanding how phonologization gives rise to certain sound patterns is key to understanding the sound patterns themselves. Representational approaches have been developed with the purpose of accounting for feature behavior, but universal representation is often too blunt an object to account for the behavior of features, often placing too much emphasis on whether a particular feature or value does or does not exist. Rather, the life of features seems to be richer than can be compressed into a model based on presence or absence of universal features.

Appendix: Single-feature bundles

Figure 5 shows the behavior of feature values acting alone. This is the same as Figure 4 but with all the occurrences of the feature value in conjunction with other features removed. This results in the appearance of more dramatically different behavior between features, but excludes instances of spreading, dissimilating, partitioning and other behavior which these features are involved in in conjunction with other features.

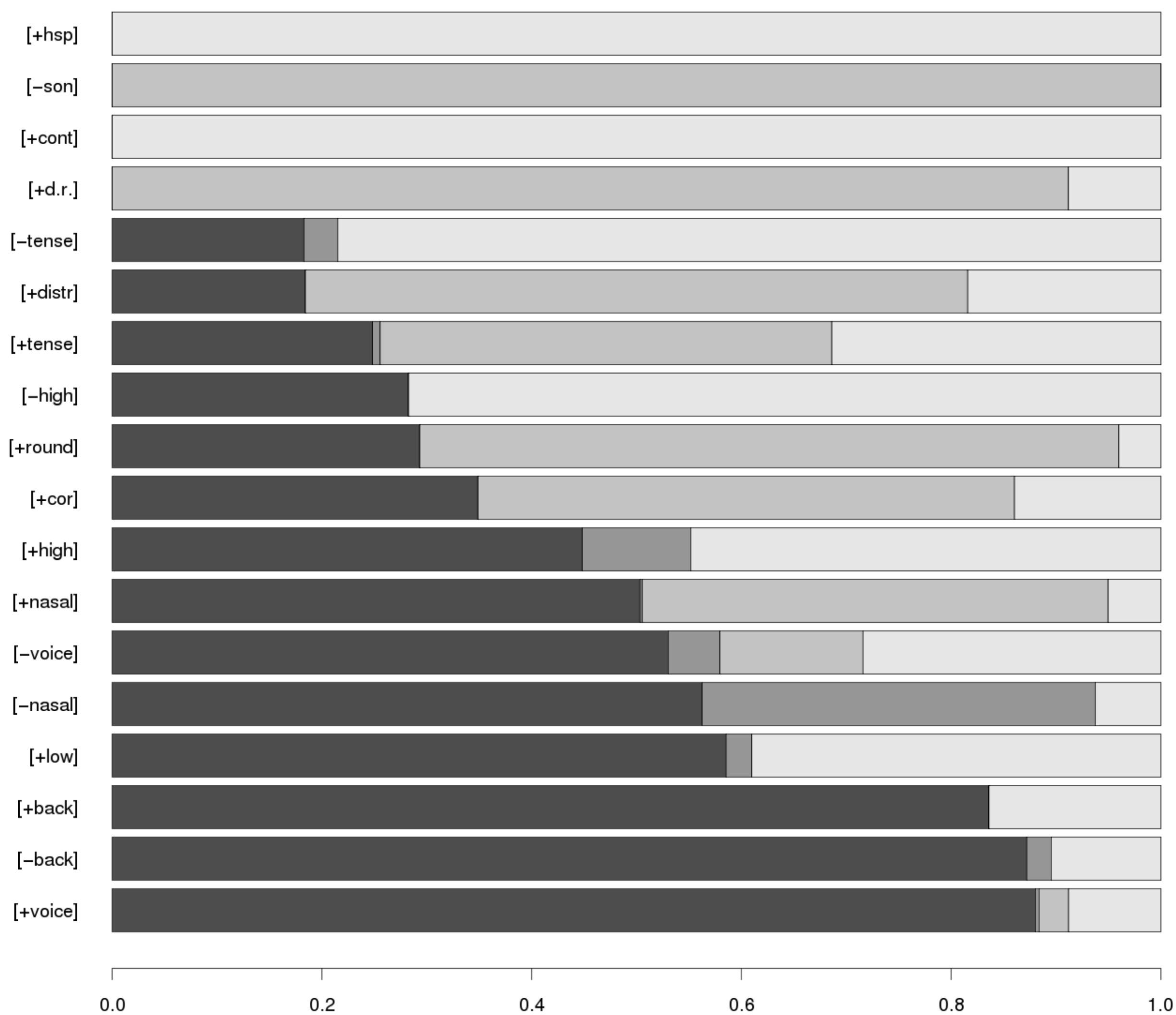

Figure 5: Behavior of single-feature bundles

References

Blevins, Juliette. 2004. *Evolutionary phonology*. Cambridge: Cambridge University Press.

Blevins, Juliette. to appear. Consonant epenthesis: natural and unnatural histories. In *Proceedings of the workshop on explaining linguistic universals*, ed. Jeff Good. Oxford: Oxford University Press.

Blevins, Juliette, and Andrew Garrett. 1998. The origins of consonant-vowel metathesis. *Language* 74:508–556.

Chomsky, Noam, and Morris Halle. 1968. *The Sound Pattern of English*. Cambridge, Mass.: MIT Press.

Clements, G. N. 1985. The geometry of phonological features. *Phonology Yearbook* 2:225–252.

Clements, G.N., and Elizabeth V. Hume. 1995. The internal organization of speech sounds. In *The handbook of phonological theory*, ed. John Goldsmith, 245–306. Cambridge Mass.: Blackwell.

Flemming, Edward S. 2002. *Auditory representations in phonology*. New York: Routledge.

Grimes, Barbara F., Joseph E. Grimes, and Richard S. Pittman, ed. 2000. *Ethnologue: Languages of the world, 14th edition*. Dallas, Tex.: Summer Institute of Linguistics. URL `http://www.ethnologue.com/`.

Halle, Morris, and Kenneth N. Stevens. 1967. On the mechanism of glottal vibration for vowels and consonants. *M.I.T. Q.P.R.* 85:267–70.

Halle, Morris, and Kenneth N. Stevens. 1991. Knowledge of language and the sounds of speech. In *Music, language, speech and brain*, ed. J. Sundberg, L. Nord, and R. Carlson, 1–19. London: MacMillan Press.

Hume, Elizabeth, and David Odden. 1996. Reconsidering [consonantal]. *Phonology* 13:345–376.

Hume, Elizabeth V. 2004. The indeterminacy/attestation model of metathesis. *Language* 80:203–237.

Jakobson, Roman, C. Gunnar M. Fant, and Morris Halle. 1952. *Preliminaries to speech analysis: the distinctive features and their correlates*. Cambridge, Mass.: MIT Press.

Jansen, Wouter. 2004. Laryngeal contrast and phonetic voicing: A laboratory phonology approach to english, hungarian, and dutch. Doctoral Dissertation, University of Groningen.

Kenstowicz, Michael, and Charles Kisseberth. 1979. *Generative phonology*. San Diego: San Diego.

McCarthy, John J. 1988. Feature geometry and dependency: a review. *Phonetica* 43:84–108.

Mielke, Jeff. 2005. Ambivalence and ambiguity in laterals and nasals. *Phonology* 22.2:169–203.

Mielke, Jeff. 2008. *The Emergence of Distinctive Features*. Oxford: Oxford University Press.

Ohala, John J. 1981. The listener as a source of sound change. In *CLS 17: papers from the parasession on language and behavior*, ed. C.S. Masek, R.A. Hendrik, and M.F. Miller, 178–203. Chicago: CLS.

Sagey, Elizabeth C. 1986[1990]. *The representation of features and relations in non-linear phonology*. New York: Garland Publishing.

Westbury, John, and Patricia Keating. 1986. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22:145–166.