How the phoneme inventory gets its shape-cognitive grammar's view of phonological systems

Geoffrey S. Nathan

Cognitive Grammar has proposed that linguistic categories are organized along prototype lines. Earlier work has shown that the internal structure of phonemes instantiates this organization. If we assume that the basic allophone is determined by principles of phonological prototypicality, we can explain why certain sounds are more common in the languages than others. As further support for this view it can be seen that numerous phonetic and phonological theories (ranging from Jakobson to Underspecification Theory) have principles that can be reinterpreted as universal prototypicality statements about kinds of sounds.

1. Introduction.

1.1. Since Cognitive Grammar came into existence it has claimed that a major feature of language structure is the assignment of linguistic forms to categories. This has been true from the very first explorations of meanings of English prepositions such as OVER (Brugman 1981) to recent research on the history of grammatical morphemes (Winters 1992). In an early work within this framework I suggested (Nathan 1986) that phonemes are categories too, and that they follow the same categorization principles as do grammatical morphemes. In a more recent study I enlarged on this theme, exploring how prototypicality principles are at work throughout the phonological system of a language (Nathan in press). In this paper I would like to explore the ways in which prototypicality principles determine the overall shape of phoneme categories, and how they determine the shape of phoneme inventories at the same time. The prototypicality principles I am discussing are not new – they have their origins in in a number of different, sometimes competing, phonological theories that we have seen

* This paper has benefited from the feedback and assistance of a number of people. Margaret Winters read and commented on numerous versions, and Linguistics gatherings at the Universities of Edinburgh, Munich and Vienna all contributed to its development. Many thanks to all, who are, of course, not responsible for any imperfections that remain.


(ricavato nel novembre 1994)
over the past seventy years. All of these theories, I believe, can be understood as exploring the idea that some sounds are better than others for the purposes to which language puts them.

1.2. The earliest comprehensive work exploring the shape of phonological systems was that of the Prague school phonologists, as seen in, for example, Trubetzkoy (1939), and as further elaborated in the classic presentation in Jakobson (1968). Their work dealt with the overall shape of phoneme inventories.

Later research within the framework of Natural Phonology (see Stampe 1979, Donegan & Stampe 1979 for discussion) proposed that the language capacity of human beings included a set of innate processes called fortition that defined the set of phonemes within a language, by systematically eliminating more marginal sounds in favor of less marginal ones.

Within the phonetically-based framework of the Quantal Theory, Stevens and his colleagues have begun exploring how certain features may acoustically 'assist' others, and thus reinforce their properties (Stevens 1989, Stevens & Keyser 1989).

Underspecification theory (as presented, for example, in Archangeli 1988), can also be considered a theory of possible phoneme inventory shape, especially if, as is suggested in Kenstowicz (1993), default rules are part of universal grammar, and thus statements about the overall nature of language, rather than being language-specific.

2. The Phoneme as a Prototype Category.

2.1. In Nathan (1986, in press) I suggested that phonemes could be thought of as prototype categories similar to categories of other types in human cognition. Thus, an aspirated [ʰ] and a voiced flap [D] are, in English, both sorts of /l/---that is, they are both instantiations of the category /l/, in much the same way as sparrows and turkeys are instances of the category bird. To put it another way, allophones are like bird species---sparrows, robins, turkeys, seagulls... We recognize them as 'kinds of birds'.

However, as we know from extensive research on categorization of all kinds, we do not perceive sparrows and turkeys as equally good representatives of 'bird-dom'. A sparrow is a better example of a bird than a turkey is, and this perception of relative 'quality' can be demonstrated in numerous ways, ranging from introspection (if you ask yourself, or anyone else, to imagine a bird, the image will look much more like a sparrow than like a turkey) to reaction time experiments (subjects recognize a picture of a sparrow as being a bird much faster than they recognize a turkey as being one). These 'prototype effects' have been seen in virtually all instances of categorization, and, as has been suggested in Lakoff (1987) and elsewhere, tell us that the pure 'Venn diagram' view of categorization is inadequate for an understanding of how human beings view the world. Specifically, categories are not mere containers into which we throw instances of entities, but rather are like family systems, with central, nuclear members and other, less essential members (such as cousins n-times removed). They have considerable internal structure within the category itself, and the membership itself is graded. Thus fathers and daughters are prototypical relatives, but great aunts are much less prototypical members of our families.

It is much more difficult to decide why particular members of a category are more or less prototypical than other members. In the case of metaphorical extensions of meanings it is relatively simple. For example, given the meanings of the word over as 'finished' and as 'located vertically' Brugman (as quoted in Lakoff 1987) shows not only that the latter is the center of the prototype category, but also why it is. But why, for example, is a sparrow a better example of a bird than a turkey? One could easily construct a set of characteristics of what the 'typical' bird looks like and does. A typical bird has wings, and can use them to fly. A typical bird is smaller than a cat and larger than a mouse. Where these judgements come from is a very complex question, but the answer seems to be that they are in part, but only in part, culturally determined. It is more likely the case that evolutionary forces cause the diversity of bird species to be clustered more closely around certain overall characteristics, and hence that the prototype effects themselves have a natural origin in the nature of animal species. Thus there is a family resemblance among birds because they evolved from the same ancestor. Since evolution follows laws of statistical probability, we can expect that species of birds will be distributed along a bell-shaped curve, with the vast majority having general similarities, but with there being exceeding oddities at either end (such as hummingbirds, or ostriches). Similar arguments can be made about virtually all living species (trees, dogs, ants...). On the other hand, evolution can also produce convergent development, such that expert knowledge is required to differentiate members of different phyla (such as sharks and dolphins, or birds and bats). In general, however, the fact that at the level of the basic category (see Rosch 1983, 1978) most cultures agree about animal and plant species indicates that categorization is not solely arbitrary, but not, on the other hand, completely determined by external forces. Instead, it seems to be an interaction between the objects in the world and human perception of them.

1 It is important to point out that I am not talking about whether membership itself is graded. Whether categories are fuzzy sets is a different question from whether they have internal structure. As I will argue, phonemic categorization is an all or none affair.
2.2. Forces dictating inventory shape.

2.2.1. If we assume that allophones are the members of the category represented by the phoneme, there are two interacting issues involved in talking about how phoneme inventories are structured. On the one hand, we have the question of which of the allophones represents the phoneme per se. This is a question of the internal structure of the phoneme. But, I am arguing, it is the selection of this allophone (or the minimal set of distinctive features which define it) that establishes the place of the phoneme itself among the other phonemes in the system. So, for example, if we decide that a language has a set of stops which are allophonically either voiced or voiceless (such as in Mohawk), then we can either decide that this language has voiceless stop phonemes, or, following various underspecification theories, that it has simply a set of stops. In either case, this definition will place the phonemes themselves within the phonemic system. So in a sense, it is the distribution of the allophones within each phoneme that determines the overall structure of the phoneme system itself, because it is the decision of which features are distinctive that shapes the overall system. This is because it is the distinctive features that set the shape of the phoneme system by establishing the other sounds against which each sound is contrasted. And it is a fact about the internal distribution of the allophones that tells us which features are distinctive and which are supplied by rules of one sort or other (or altered by rule, depending on one's theory).

Consequently, we need to explore what forces determine the choices about the internal structure among allophones, in order to understand what forces determine the external relationships among the phonemes. We can make an analogy here to political relationships among nations. While a nation's political system is primarily determined by forces internal to the country, it is those decisions that determine the external relationships that the country experiences—the alliances that it chooses to join and so on. Consequently, the study of natural phonological processes (postlexical phonological rules and default rules in most current Generative terms) will enlighten us as to the reasons for the content and shape of the inventories that we find. So the forces shaping the internal structure of categories determines the interrelationships among those categories.

2.2.2. I would like to suggest that the internal forces acting on the phoneme itself can be seen as determining the placement of the phoneme among others within the system. Consider the following image: All the possible sounds available in human language constitute a field of evenly spaced points, like the pile on a carpet. However, not all points have equal probability, because for varying reasons, to be discussed at greater length below, some of them are better suited to be sound units in a natural language than others. If we treat each principle favoring some feature value of a sound over some competing feature (say, that oral vowels are superior to nasalized ones) as a principle raising the height of individual sounds above its neighbor, we will develop a carpet with hills rather than one resembling a smooth plain. Each peak will constitute a sound which, through a number of independent driving principles, has emerged as an ideal one. The tips of the peaks themselves will correspond to the typical phoneme inventories that we find, with voiceless stops higher than voiced ones, peripheral vowels higher than centralized ones and so on. We should find that individual peaks tend to be somewhat distanced from each other, but that this is not being caused by an independent set of principles distanced peaks (i.e., 'contrast' is not driving the sounds apart), but rather the fact that sounds inherently have ideal states (defined by what we have been calling prototypicality principles) which will spontaneously organize any random set of possible human sounds into evenly spaced rows and columns with sizable distances between them.

2.2.3. The fact that there will be rows and columns is also derivable directly from the prototypicality statements, since these statements (although, importantly, not the sounds that they apply to) are stated in terms of features only. This is a principle admirably first in Donegan & Stampe (1983) which they refer to as 'asegmental phonology'. The principle, which I subscribe to here, is that the sounds of language are real, individual segments (perhaps organized hierarchically into syllables, feet and so on), but that the principles governing those sounds, modifying them allophonically and/or morphophonemically etc., are constrained to refer only to features— that is, only to classes of sounds by virtue of their characteristics—and that the domains over which the principles apply are not, under normal circumstances, segment sized, but are rather prosodic units of various sizes, ranging from the syllable to the phonological phrase. The result of this binary constraint, namely that the sounds are represented qua sounds, but the rules manipulating them are limited to expression in features, is that the inventory of sounds will be lined up in neat rows and columns if the rules are themselves prototypicality principles.

To see how this could be so, consider a limiting set of prototypicality principles, such that they express absolute constraints. For example, suppose that we have a simplistic point of articulation constraint: all consonants are either labial, coronal or velar. If we apply this to the IPA chart (even including all possible modifications permitted by the associated diacritics), the result will be a consonant chart consisting of exactly three columns. Then suppose that we extend our consonantal
constraints by saying all consonants are either stops or fricatives, and finally, all fricatives are voiceless. This will result, through the application of three general constraints stated in (unofficial) feature terms to the entire set of possible human consonants, in a chart of the following:

```
p  t  k
b  d  g
f  s  x
```

Consequently, a set of principles defining the 'best quality' sounds, defined solely in features, operating on sounds, considered a mass of different kinds, will inevitably lead to a nice, neat feature chart.

Let me give an illustration of a different, but related sort. All languages of the world have coronal consonants (even Hawaiian, which lacks a coronal obstruent, has /h/). But, with very few exceptions, if a language has a coronal consonant or series, it is anterior, and no language has non-anterior coronals while lacking anterior ones. That is, the 'best' coronal consonant is clearly dental or alveolar, while alveopalatalals, retroflexes and, at the absolute extreme, linguo-labials, are dispreferred. Again, then, we can say that the prototypical tongue-tip consonant is articulated either at the teeth or at the alveolar ridge.

3. Sources of prototypicality.

3.1 Jakobsonian laws.

The first extensive presentations of what we could now call 'prototypicality' principles within phonological systems can be found in the works of the Prague school phonologists. A particularly good example is Jakobson's *Kindersprache, Aphasie und allgemeine Lautgesetze*, originally published in 1941, and easily available now as Jakobson (1968). In it he suggests that the order in which sounds are acquired as phonemes by children, irrespective of native language, corresponds to an order of relative frequency of the same sounds among the phoneme inventories of the world. Thus, he claims, children first acquire an opposition between high and low vowels, followed by a front/back opposition. This corresponds to a minimal possible phoneme inventory either of the classic /i,a,u/ type found in many languages, or to a system consisting solely of height contrasts, such as those "found in the West Caucasian languages, in which the choice between different vowels of the same degree of opening [...] depends only on the adjoining consonants" (Jakobson 1968:50).

Jakobson refers to these laws as 'irreversible laws of solidarity', a term he takes from Husserl and which we would now call implicational

hierarchies. He considers them to be universal (more properly 'general'-*allgemein*). Thus '[T]hey retain their validity at every stage, and in the course of every change of all of the languages of the world' (p.59).

3.2. Fortitions.

Natural Phonology has argued that the inventory of phonological rules of language are of two fundamental types: fortitions and lenitions. Lenitions are the kinds of phonological rules most familiar to phonologists of all theoretical persuasions – assimilations and reductions. In general they define the non-central allophones and syllable structures of a language (nasalized vowels in English, derived consonant clusters in French, reduced vowels in Russian). However, there is another class of rules in NP known as fortitions. These processes have no exact counterpart in other theories. Donegan, in a recent work refers to them as rules of 'selection'. They eliminate classes of sounds by ruling out incompatible combinations of features. They often have the opposite effect to lenitions. For instance, while there is a lenition nasalizing vowels adjacent to nasal consonants, there is a fortition which says that vowels should be non-nasal. Similarly, there are fortitions inserting vowels within consonant clusters and processes polarizing vocal qualities – increasing palatality or labiality (hence the opposite of vowel reduction, which normally removes distinctive vowel qualities.) As Donegan (MS:20) says:

fortitions result in the strengthening (maximization or optimization) of some particular phonetic property of an individual segment. These are the constraints that limit our perceptions of what is a possible speech sound. They represent the creation of categories of sounds. Fortitions eliminate distinctions, creating fewer, more-inclusive categories of sounds.

3.2.1. Fortitions thus express, in process terms, the insights originally explored by Jakobson, about the preferential organization of phonological systems. They constitute the integration of markedness relationships into the synchronic grammar by making them part of the living phonology of the language. Within NP it is believed that processes represent real, active constraints on pronounceability. Fortitions function both within the lexicon as 'gatekeepers', eliminating impossible underlying forms, and also function during the perception

footnote: Recall that NP makes a fundamental distinction between 'rules' (which correspond roughly to Lexical rules within Generative grammar) and 'processes', which correspond (again roughly) to Postlexical rules (and possibly some Postcyclic Lexical rules). Processes are those Postlexical rules which appear to apply during speech production, and which apply after speech errors. They constitute a set of constraints on what is 'pronounceable', defining what is 'hard to say'. As a consequence they are generally below the level of consciousness.
of ongoing speech by constraining what we hear, in order to make it fit the phonological system of our native language, whatever the input. The effects of fortitons are not normally visible except when the input (that is, what we hear) violates the bounds of our phonological system, at which time the set of fortitons 'corrects' what we hear. For example, French nasalized vowels are perceived as oral vowels plus a nasal consonant by native speakers of English or Japanese, and the implosive voiced stops found in many dialects of Southern American English are simply classified as plain voiced stops, even by the vast majority of native speakers of American English who lack these particular allophones.4

The view that there is a set of processes that select the set of 'ideal' sounds is not limited to Natural Phonology. In his dissertation, Andrea Calabrese (1988) provides a history within Generative Phonology of processes such as 'morpheme structure rules' and 'default rules' which, to a certain extent attempt to constrain the set of phonemes found in a language. However, the idea that they constitute active processes available and ongoing in the real time production of speech is a unique view of NF, which holds that they constrain the actual perception of speech in real time, the speech of both native and non-native speakers.

3.3. Quantal effects.

3.3.1. In recent years Stevens and his coworkers have developed a theory of phonological features which suggest that there are natural (spectral) acoustic boundaries between kinds of segments that are determined by the structure of the articulatory apparatus. Within limited parameters slight articulatory variation produces no significant acoustic differences, but at other articulatory points a very slight difference produces a very noticeable perceptual difference. For example, Stevens states that:

there are some articulatory states or configurations or gestures that give rise to well-defined patterns of auditory response in a human listener, such that these patterns are not strongly sensitive to small perturbations or inaccuracies in the articulation. These patterns are distinctive in the sense that if some articulatory parameter crosses over a threshold region there will be a significant change or a qualitative shift in the auditory response. The multidimensional space that devides acoustic-articulatory or auditory-acoustic relations, rather than showing continuous and monotonic variation, exhibits quantal attributes characterized by rapid changes in state over some regions and less abrupt variations or greater stability over other regions. (Stevens 1989:5)

As a result of this sensitivity to boundary zones, features are constrained to combine in certain ways such as to augment the distinctiveness of particular sounds:

The strength with which each of the primary features is represented in the sound is influenced by the combination of secondary features that co-occur with the primary features [...] An outcome of this exercise is a specification of a number of feature combinations or segments that are in some sense optimal, since they provide the strongest representation of the contrasts defined by each of the three primary features 6. (Stevens & Keyser 1989:89)

These specific combinations of features that produce the most salient sets of contrasts explain why languages choose particular sets of combinations of features over others. This of course, as we have been arguing, leads to particular phoneme inventories being preferred over others: "We suggest that this tendency toward particular feature groupings characterizes the languages of the world because of the properties of saliency and enhancement which these groupings exhibit." (Stevens & Keyser 1989:104)

3.4. Default rules and underspecification theory.

3.4.1. Although it is not normally envisioned in this way, Underspecification theory could also be considered a version of a theory of prototypicality. Radical underspecification theory (Archangeli 1988) holds that phonemes are characterized only by their marked values, while contrastive underspecification theory holds that they are characterized by their distinctive features. In either case, those features not added by garden variety phonological rules (through assimilations, reductions and the occasional dissimilation) are filled in at the 'bottom' of the derivation by Default Rules. The difference between Radical and Conservative Underspecification theory is that in the Radical version the least marked segments are not specified at all, so that for example there may be a stop (often /l/) that is marked only as [+consonantal]. Since it is the least marked stop, all other features are predictable. Within the less radical theory, if there are contrasts, the contrasting features must be listed.

It should be pointed out here, incidentally, that Stevens and his coworkers would not endorse the segment-based approach that I am advocating here. The fact that sounds are processed and categorized by the rules of the language using features does not rule out the existence of segments as storage units, and in particular phonemes as potential objects of perception. As mentioned above, storage (and hence categorization) is in terms of units the size of segments, although the processing carried out on them is determined by feature-sized considerations.
3.4.2. One of the functions of this theory is to allow spreading (assimilation) rules to operate correctly, missing the redundant features intervening between triggering sounds and the targets of the assimilations. Classic examples include retroflexion in Sanskrit, which spreads past labials but not past coronals, and Latin l → r dissimilation, which spreads over all segments except l or r. (For extensive discussion on this issue see Kenstowicz 1993:506-513)

3.4.3. The result of the interaction of the default rules and the requirement that only contrastive (or fewer) features be present in underlying representations is to define a set of shapes for phoneme inventories of exactly the kind we are discussing. Radical underspecification theory is a little harder to squeeze into this mold because there is some question of where in a phoneme inventory we could place the most minimally specified sounds. Without any place features, for example, where in a chart would we place /i/? Archangeli’s analysis of Yawelma argues that /i/ is the minimally specified vowel, which would make its sole defining characteristic [-consonantal], and in terms of issues such as minimal distance, inventory shape and so on, it is unclear how this could be interpreted. Contrastive underspecification theory, on the other hand, reduces phonemes to exactly the right set of features to push the system into the familiar shapes originally proposed by Trubetzkoy.

3.5. Calabrese’s Theory of Phonological Alphabets.

In a recent dissertation, Andrea Calabrese (1988) proposed a similar theory to the one we have been exploring, although not at all within the framework I am proposing. Calabrese proposes that UG includes a set of feature filters that specify incompatible feature values. In exactly the way we have been discussing here, however, these filters are not absolute constraints, but rather relative sets of dispreferences for particular combinations of features. Thus a particularly bad combination of features for a given vowel would be [+low, +high], while a mildly dispreferred one would be [-low, -high], which merely reflects the fact that /i,u,a/ is the most common three-vowel system. The ranking of these rules, ranging from ones like the former, which reflect what seems to be completely impossible, to the latter, which reflect what seems merely to be slightly dispreferred, seems to be exactly a set of prototypicality principles for possible sounds of a human language.

4. Conclusion

4.1. There is clearly a convergence of opinion on the forces determining the most and least likely combinations of features within a set of phonemes. From the viewpoint of a number of different frameworks we can see general agreement that some sounds are ‘better suited to be phonemes of a language than others, and that this can be explained by tendencies to augment certain features or combinations of features at the expense of others. This in turn leads languages to select sets of features that lead to an inventory shape of a certain type. The overall shape emerges out of the preference for the combinations of particular features without the independent operation of ‘system-shaping’ forces such as a drive for symmetry, or even a minimal distance principle. The emergence of order out of the individual ‘preferences’ of the atomic units involved is, of course, a common notion in studies of spontaneous order, and has been extensively discussed by Keller (1985). We can thus find explanations for such structurally motivated sound changes as drag chains and push chains, not in terms of a drive towards a better-looking system, but in terms of the fact that each sound will tend to optimize its own characteristics.6 The ultimate result is that a system lacking /a/ or any other low vowels, will gain one, not because the other phonemes need Lebensraum, but because the existing sound with the greatest remaining sonority will prefer to increase sonority, hence lowering to fill the void caused by the loss of the low vowel.

4.2. There is some question, given this way of looking at the internal structuring of sound systems, whether we can reconstruct any viable way of talking about ‘push chains’, but of course it has often been questioned whether any such changes in fact occur. Certainly it is far more often the case that when a sound approaches a neighbor the result is a merger rather than a concomitant fleeing of the threatened sound. Chain shifts are much rarer than neutralizations, at least in historical change. Neutralizations in rapid speech processes are very common, but push chains in rapid speech seem to be virtually unknown.

4.3. In sum, it seems that each sound within a language is striving to become the best it can be (that is, there is a set of prototypicality principles for each feature of any sound), and each sound is consequently moving towards an ideal state. This fact leads to phoneme inventories structuring themselves to take more or less symmetrical shapes that have preferentially three major points of articulation for

6 We do not, of course, really mean that the sounds themselves prefer particular combinations, but that we, as speakers and hearers, find certain combinations better for our purposes, and as a consequence use them. Noting the possible anthropomorphizing of the terminology I will continue to use it.
obstructs, three heights for vowels (with the front ones unrounded and the back ones rounded) and so on in the ways we have known about since (at least) Trubetzkoy. What I am suggesting in this paper is that these prototypicality principles, which have been investigated within other frameworks by many phonologists and phoneticians over the past sixty years, are responsible for the fact that inventories have shapes, and that they have the shapes they do.

Address of the Author:

Geoffrey S. Nathan
Southern Illinois University at Carbondale
Carbondale, IL, 62901-4305
USA
E-mail: geoffn@siu.edu

References


Dinnisen, D., ed. (1979), Current Approaches to Phonological Theory, Bloomington, Indiana University Press.

Donegan, P. J. (MS), "The acquisition of phonetic perception".


