Testing Feature Economy

G. N. Clements
Laboratoire de Phonétique et Phonologie (UMR 7018)
CNRS / Sorbonne-Nouvelle, Paris, France
E-mail: clements@idf.ext.jussieu.fr

ABSTRACT

Why do speech sounds tend to be organized along just a few phonetic dimensions in each language, such as “nasals”, “voiced fricatives”, etc.? The classic answer to this question was proposed by Martinet [8]: “Une nouvelle économie est réalisée en faisant résulter [les phonèmes] de combinaisons de traits phoniques non-successifs, ce qui réduit encore le nombre des éléments de base” (p. 95). In other words, features introduce a further level of economy by allowing the set of phonemes P to be defined in terms of a smaller set of features, F. Economy exerts pressure on systems to maximize the combinatory possibilities of features by increasing the ratio P/F. This paper outlines a method for testing feature economy, and presents illustrative results involving the patterning of labial consonants.

1 INTRODUCTION

According to the principle of feature economy, proposed by Martinet [8] after an earlier suggestion by de Groot [5], phoneme systems tend to use a small number of features (F) to define a larger number of phonemes (P). The pressure to increase the ratio P/F accounts, in this view, for historical effects such as the shift of isolated phonemes into correlative series. As also noted by Martinet, feature economy is subject to limits imposed by functional constraints, such as the avoidance of articulatory complexity and the preference for perceptually salient contrasts.

Feature economy can be tested at the synchronic level. This paper first identifies two predictions of feature economy, Mutual Attraction and Cumulativeness. Following this it outlines a quantitative method for testing these predictions, and applies it to selected comparisons involving labial consonants.

2 PREDICTIONS OF FEATURE ECONOMY

The basic insight underlying feature economy is that if a feature is used once in a system, it will tend to be used again. Several predictions follow from this view. One is that sounds will tend to attract other sounds bearing the same features. This prediction can be formulated as follows:

> **Prediction 1** (Mutual Attraction): A given sound will have a higher than expected frequency of occurrence in languages having other sounds bearing one or more of its features.

Another prediction is that if a feature is present in a system at all, it will tend to be employed not just once but several times, maximizing its effectiveness.

> **Prediction 2** (Cumulativeness): The frequency of occurrence of a given sound will increase in proportion to the number of other sounds in the system bearing the same features.

3 DATA BASE

The data base used for this study is UPSID-92, a sample of 451 phoneme systems from the world's languages [7]. This data base was selected for several reasons. First, it is the largest data base of this type currently available. Its 451 languages represent over 7% of the estimated 6,000 languages of the world. Second, UPSID-92 was constructed by selecting just one language from each moderately distant genetic grouping in order to insure some measure of genetic and typological balance. Third, the use of a computerized data base facilitates rapid searches using query languages such as the one packaged with UPSID itself.

Nevertheless, no data base is ideal. One problem is that true genetic balance is an unattainable goal, given the inevitable skewing toward units containing large numbers of languages. For example, Niger-Congo is represented in UPSID-92 by 55 languages, Sino-Tibetan by 21, and isolates such as Basque by only 1; insofar as the daughter languages of Niger-Congo preserve features of the parent language, UPSID counts the same language 55 times. A second problem is that the inventories included in UPSID have been compiled by many researchers using phonetic symbols and labels differently, and assuming different theoretical models and criteria for phonemic analysis as well as for the choice of allophone selected to represent each phoneme; no rigorous standardization of these source materials is possible. A third problem is that UPSID is riddled with clerical errors involving miscopying, miscoding, etc. [1]. Such errors may not be significant in the case of broadly-based trends, but can become important in the case of generalizations based on just a few languages

ISBN 1-876346-48-5 © 2003 UAB

(e.g. “some languages have minimal contrasts between velar stops and fricatives”). In such cases, the data should be checked against the original sources and interpreted in the light of each author's usage and theoretical approach.

4 FEATURE CODING

To test Predictions 1 and 2, the phoneme systems of UPSID-92 were coded in terms of a current model of distinctive features [4]. For consonants, these include the following:

one-valued: labial, coronal, dorsal, radical, spread glottis, constricted glottis

two-valued: sonorant, consonantal, distributed, anterior, strident, lateral, voice, nasal, continuant

It might be objected that the use of a contemporary feature set does not provide a fair test of a hypothesis that was first conceived in terms of another set of features. In fact, as far as consonants are concerned the feature system shown above does not differ radically from that of Martinet [8]. The two systems share the following characteristics:

- most features are defined in articulatory terms
- most manner features are two-valued
- most place features are one-valued

Natural classes defined by the two systems are also very similar. The newer system differs from Martinet's mainly in its inclusion of the articulator-based features [coronal] and [dorsal] instead of more specific place-of-articulation features such as apico-dental and velar [9]. Even with this difference, the systems are easily comparable.

5 STATISTICAL ANALYSIS

Frequencies of selected sets of sounds in the UPSID inventories were compared in order to determine the number of languages having 1) each member of the set, 2) all members of the set, and 3) no members of the set. Numbers were arrayed in contingency tables and the resulting distributions were tested for significance by the chi square (χ^2) test. This test is commonly used to test whether two characteristics are associated in such a way that high frequencies of one tend to be coupled with high frequencies of the other. It is therefore appropriate here. However, several precautions must be taken in interpreting χ^2 test results.

First, tables containing cells whose *expected* values are 5 or less must be avoided as far as possible, since in such comparisons one of the assumptions underlying the chi square test, that the values correspond to a continuous rather than a discrete frequency distribution, is not valid. In cases where such tables must be used, χ^2 values are calculated using Yates's continuity correction which yields smaller values and makes a more conservative projection of significance levels.

A further precaution is that a positive association between A and B does not necessarily demonstrate a direct relation between A and B. The association may instead be due to the covert influence of a third term, C, which is positively associated with both A and B. Such “covert attractors” can usually be detected by performing a new comparison within the subset of the data base from which languages having the suspected attractor have been removed. A related problem is that a positive association between A and B may be due to a subset of A rather than to A as a whole; such “subset effects” must also be searched for to avoid over-generalization.

Third, if a statistically significant association is found to hold primarily within the languages of one genetic unit, the association is better explained by common genetic heritage or language contact. To confirm that a given result is not an effect of common retention or areal diffusion, the results should be tested in all major genetic groupings for which reliable data are available.

6 TESTING PREDICTION 1

Prediction 1 (Mutual Attraction) will be illustrated by comparing the frequencies of voiced labial fricatives V and voiced coronal fricatives Z. Here and below, upper-case letters are used to designate general sound types (labials, etc.) as opposed to specific phonetic categories (bilabials, labiodentals, etc.). Our expectation under feature economy is that voiced labial fricatives V should be positively associated in frequency with voiced coronal fricatives Z by virtue of their shared features [+continuant] and [+voiced]. In other words, there should be more languages having both V and Z than would be expected if their frequencies were independent of each other. To test this hypothesis, a 2x2 contingency table was constructed as shown in Table 1.

		Z present	Z absent	T =
V	present	110 (57)	37 (90)	147
	absent	65 (118)	239 (186)	304
	T =	175	276	451

Table 1: Contingency table showing the observed frequencies (F_O) of voiced labial fricatives (V) and voiced coronal fricatives (Z) in UPSID phoneme systems. Expected frequencies (F_E) are shown in parentheses.

This table shows that of the 451 languages in the sample, 110 have both V and Z, 37 have V but not Z, 65 have Z but not V, and 239 have neither. These are the observed frequencies (F_O). The *expected* frequencies (F_E) of V and Z under the null hypothesis that they are

ISBN 1-876346-48-5

mutually independent are shown in parentheses. Since the observed frequency of languages having both V and Z (110) is higher than the expected frequency (57), the association between V and Z is positive. In other words, languages having one of these sounds tend to have the other.

We next ask whether the association is statistically significant. The probability that a given association is due to chance is estimated by comparing the differences between observed and expected frequencies. In this case, the discrepancies are very large and prove to be highly significant ($\chi^2 = 119.203$, $p < .0001$).

This result confirms Prediction 1 for this pair of sounds. We must ask, however, whether it could be due to genetic skewing. An examination of the V/Z distribution in each of the eleven major genetic groups represented in UPSID-92 reveals a positive association between V and Z in nine of them. The remaining groups (Australian, Nilo-Saharan) show neither a positive nor a negative association. We conclude that the result is not due to the undue influence of any one group in the sample, and that Prediction 1 is confirmed at the cross-linguistic level for this pair.

7 TESTING PREDICTION 2: CUMULATIVENESS

To test prediction 2 (Cumulativeness), a comparison was made of the frequency of voiced labial fricatives V in systems with 0, 1, 2, and 3 other members of the set of labial obstruents, P B F. The following frequencies, expressed as percentages, were calculated for V:

- F0: the frequency of V in systems having *no* other labial obstruent
- F1: the frequency of V in systems having *just one* other labial obstruent
- F2: the frequency of V in systems having *just two* other labial obstruents
- F3: the frequency of V in systems having *all* other labial obstruents

Numbers were arrayed in a 2x4 contingency table and the resulting distribution was tested for significance. The prediction is that the frequencies F0, F1, F2, F3 will constitute a rising scale. The result is shown in Table 2.

The first row of this table shows the number of languages having V together with 0, 1, 2 or 3 of the other members of the set of labial obstruents, P B F. It shows that no languages have V alone, 17 have V and one other member of the set P B F, and so forth. The second row shows the total number of languages with 0, 1, 2 or 3 members of this set (including those with and without V); thus 5 languages have no labial obstruents at all, 102 have P, B, or F (with or without V), and so forth. The third row shows values of F0, F1, F2, and F3 as defined above. As this row shows, Prediction 2 is strongly confirmed: values of F0, F1, F2, and F3 form a rising scale. This distribution is highly significant (conflating columns 1 and 2 to achieve sufficient numbers in each cell, $\chi^2 = 47.697$, $p < .0001$). In other words, the frequency of V across phoneme systems increases in proportion to the number of other members of the set of labial obstruents (P B F) that are present.

	number of members of the set of labial obstruents P B F:				
	0	1	2	3	T=
languages with V	0	17	47	83	147
total	5	102	187	157	451
frequency of V(%)	0	16.7	25.1	52.9	32.6

Table 2: Frequency of voiced labial fricatives V in systems containing 0, 1, 2, and 3 other members of the set of labial obstruents (P B F).

To test for the generality of this effect, the other labial obstruents P, B, F were tested in the same way. The frequencies F0, F1, F2, and F3 for each of these sounds and their total frequencies in UPSID are shown in Table 3. (Results for V are included in the first row for comparison.)

	number of other members of the set P B F V:				
	0	1	2	3	overall
V	0	16.7	25.1	52.9	32.6
F	0	32.0	44.8	76.1	48.3
B	*66.7	62.7	76.9	82.2	71.4
P	*93.8	*92.9	81.9	96.5	90.2

Table 3: Frequency of occurrence (in %) of the labials V, F, B, P in systems containing 0, 1, 2 and 3 other members of the set, and in UPSID overall. Asterisks mark cells in which Prediction 2 fails..

As this table shows, Prediction 2 is also confirmed for F (row 2): frequencies increase steadily as we read from left to right. However, as shown by the asterisks, Prediction 2 fails in one cell for B and in two for P. Thus Prediction 2 holds only for the marked labial sounds: the fricatives and (except for one cell) B.

Post hoc analysis shows that the anomalous cell values of B and P are due to interacting markedness and feature economy effects. First, the unmarked labial obstruent P occurs as the sole labial with a much greater than expected frequency (41 systems expected, 75 occur); this skewing lowers the frequency value of the second cell of B and raises that of the first two cells of P. Second, the fricatives F and V do not occur as the sole

 ISBN 1-876346-48-5 © 2003 UAB

labial obstruent, another markedness effect, which further raises the frequency value of the second cell of P. Third, by an interacting feature economy effect, the systems /P B F/ and /P B V/, characterized by a missing fricative, are less frequent than would be expected on the basis of chance, lowering the frequency value of the third cell of P. The frequencies in Table 3 thus result from several interacting factors, explaining why the cumulativeness principle (Prediction 2) holds only for the marked labial sounds, especially the fricatives.

8 LIMITS ON FEATURE ECONOMY

If there were no limits on cumulativeness, sound systems would be infinitely large. As Martinet observed [8], feature economy is constrained by factors imposed by the nature of speech communication: a sound will tend to be avoided if its cost in terms of articulatory complexity or low contrast with other sounds is too high.

Let us take an example. Given the cumulativeness principle just discussed, we might expect labial systems to acquire further contrasts *within* each of the classes P, B, F, and V. However, this expectation is not generally confirmed. Labial obstruent systems tend to be "saturated" when they acquire four members. For example, as Table 4 shows, bilabial and labiodental fricatives show a strong negative association in our sample: 36 languages would be expected to have both if their frequencies were independent of each other, but only 20 languages actually do.

	Exp	Obs	χ^2	p <
BIL vs. LBD FRICATIVES	36	20	15.83	.0001

Table 4: Expected and observed frequencies of bilabial (BIL) and labiodental (LBD) fricatives across languages. These two sound types are negatively associated.

Languages having one of these fricative types tend not to have the other. This may be because these sounds, though easy to produce, represent "fairly subtle distinctions that are not easy to hear" (Ladefoged [6]: 141).

9 CONCLUSIONS

This paper has outlined a method for studying feature economy effects and has illustrated its application to the study of mutual attraction and cumulativeness in certain consonant systems. It has found strong evidence for feature economy, but has also confirmed the existence of functionally-motivated limitations on its full generality. Other comparisons are currently being tested in order to determine the full extent of feature economy effects in the organization of sound systems ([2], [3]). Questions being studied include: Which features show economy effects? Do both values of binary features exhibit economy effects? What about redundant features, and unmarked features? What are the consequences of feature economy for areas such as speech perception and first and second language acquisition? Answers to these questions will do much to advance our understanding of phonological universals.

ACKNOWLEDGEMENTS

I thank Jean-Yves Dommergues and Hannah Rohde for invaluable help in developing the statistical method and computational tools used here, and the Féderation de Typologie et Universaux (CNRS) for its support.

REFERENCES

[1] H. Basbøll, review of I. Maddieson, *Patterns of Sound*, *Phonology Yearbook*,vol.2, pp.343-53, 1985

[2] G. N. Clements, "Feature economy as a phonological universal," this volume, 2003.

[3] G. N. Clements, "Feature economy," Ms., CNRS / Sorbonne-Nouvelle, Paris, 2003.

[4] G. N. Clements and E. Hume, "The Internal Structure of Speech Sounds," in *Handbook of Phonological Theory*, J. Goldsmith, Ed., pp. 245-306. Oxford: Basil Blackwell, 1995.

[5] A. W. de Groot, "Phonologie und Phonetik als Funktionswissenschaften," *Travaux du Cercle Linguistique de Prague*, vol. 4, pp. 116-147, 1931.

[6] P. Ladefoged, *Vowels and Consonants: an Introduction to the Sounds of Language,* Oxford: Basil Blackwell, 2001.

[7] I. Maddieson and K. Precoda, "Updating UPSID," *UCLA WPP* vol. 74, pp. 104-111, 1989.

[8] A. Martinet, *Économie des changements phonétiques*, Berne: A. Francke, 1955.

[9] E. Sagey, *The Representation of Features in Nonlinear Phonology: the Articulator Node Hierarchy*, New York, NY: Garland, 1990.