A BEHAVIOURAL CORRELATE OF PHONOLOGICAL STRUCTURE*

DANIEL NETTLE

University College London

Data from a cross-cultural study of conversational behaviour were compared with the phonological inventories of the languages being spoken. There was a highly significant inverse relationship between the number of contrastive vowels and the average volume of native speakers' voices. It seems that different languages are not equally salient perceptually.

Key words: phonological structure, speech volume

INTRODUCTION

There is a long tradition of attempting to explain patterns of linguistic sounds using the assumption that languages make the most efficient use possible of human articulatory and perceptual mechanisms. For example, Lindblom (1986) has argued that the vowels of a language tend to be distributed in perceptual-acoustic space such as to produce maximal perceptual contrast, whilst minimising the number of articulatory dimensions involved. A similar account has been attempted for consonants (Lindblom and Maddieson, 1988). It is clear, however, that even if human languages choose their consonants and vowels according to similar optimising principles, the resulting systems vary widely in terms of total inventory size, vowel/consonant ratio, and syllabic structure (Maddieson, 1984). The question thus arises whether there are differences in perceptual-acoustic efficiency between languages that affect the behaviour of speakers.

A relevant data set in this context is that gathered by Watson (1970). Watson observed conversations between pairs of subjects of 31 nationalities and found marked cultural differences on five measures of proxemic behaviour. These were the orientation of the two people to each other (Axis), the distance between them (Distance), the amount of face-to-face gaze (Gaze), the amount of touching (Touch), and the volume of their speech (Voice). His analysis led him to group cultures as basically "contact" (close, loud voice, extensive visual contact) and "non-contact" (distant, quiet, less visual contact). The contact group consisted of all Arab, Latin American, and southern European nationalities. The non-contact group consisted of all subjects from northern Europe, the Indian subcontinent, and southeast and east Asia. The groups differed significantly from each other on all five of the measures.

---

* This research was funded by the UK Medical Research Council. I am grateful to Robin Dunbar, Michael Studdert-Kennedy, and Bruno Repp for helpful comments on an earlier version of this paper.

Address correspondence to Daniel Nettle, Department of Anthropology, University College London, London WC1E 6BT, UK.

Watson interpreted the differences as aspects of a learned cultural matrix. It seems likely, however, that speech volume is constrained by the salience of the language being spoken. In this study, the Voice scores from Watson's data are compared with aspects of phonological structure. It is hypothesised that the greater the functional load borne in the language by vowels, which are highly sonorous, the lower the volume required for conversation will be.

METHOD

Watson made his observations from behind a one-way mirror during conversations staged for experimental purposes in same-nationality dyads. The subjects were 122 students of the University of Colorado. The subjects conversed in their native tongues. Although the information is not given, it has been assumed that these were the standard varieties of the national or majority languages of their countries. The variables were all measured on scales that increased with relative distance, so that a low score meant closer contact. Voice volume was sampled every 10 seconds using a microphone and the decimeter (*sic* – presumably the sound level meter) of a tape recorder. As the absolute accuracy of the meter was unknown, the scale was divided into six ranges from Very Loud (1) to Very Soft (6). Watson's data, which are grouped by nationality, are regrouped here by language.

Phonological inventories were obtained for all the languages involved from standard sources in the literature, principally Campbell (1991). These were used to calculate the following three indices:

(i) V: The total number of contrasting syllabic nuclei, derived from the number of monophthongal vowels and diphthongs multiplied by the number of contrastive vowel lengths, tones, or stresses where applicable.

(ii) C: The number of contrastive consonants in the inventory, counting pairs of segments distinguished by length as distinct. This was intended to capture the maximum number of contrasts possible at a non-nuclear slot.

(iii) V/C: The ratio of V to C.

Turkish presented a difficulty. Although it has a total of eight contrastive vowels, vowel harmony restricts the choice to two in non word-initial syllables. The average value of V will therefore be somewhere between 2 and 8, depending on the length of the phonological word. To obtain an estimate of a typical value for this length, three articles from the Turkish newspapers *Hurriyet* and *Türkiye* were analysed. One was the text of an informal interview, and the other two were columns on sport and current affairs. The averages for the number of syllabic nuclei in the phonological word (which does not always coincide with the orthographic word in Turkish) were 2.54, 2.6 and 3.1 for the three articles. The overall mean was 2.75. Whilst this value is likely to differ from that occurring in spontaneous conversation, it is a reasonable basis for estimation. In every word, there will be one nucleus with eight possibilities, and an average of 1.75 nuclei with two possibilities. The average V is thus $(1*8 + 1.75*2)/2.75 = 4.3$.

TABLE 1

The average behavioural measures and the phonological indices for the fifteen languages

Language	Voice	Gaze	Distance	Axis	Touch	V	C	V/C
Arabic	3.96	1.26	3.53	2.57	6.59	6	28	0.21
Hindi/Urdu	4.39	2.05	3.94	3.60	6.99	20	32	0.63
Spanish	4.15	1.41	4.96	2.47	6.74	10	23	0.43
French	5.05	1.19	4	2.08	6.89	38	19	2
Italian	4.56	1.83	5.64	1.79	6.84	21	36	0.58
Turkish	3.6	1.4	2.83	3.2	6.97	4.3	20	0.22
Mandarin	4.62	2.55	3	4.3	6.82	32	21	1.52
Indonesian	4.57	1.83	5.15	3.15	7	8	22	0.36
Japanese	4.6	2.04	5	3.02	7	10	15	0.67
Tagalog	5.47	2	8	3	7	55	15	3.67
Thai	4.92	1.89	5.03	3.03	7	54	22	2.45
English	4.41	2.5	7.48	3.08	7	20	25	0.8
German	4.63	1.97	5.13	2.88	7	20	21	0.95
Dutch	4.8	1.55	3	4.43	7	21	19	1.11
Norwegian	4.79	1.9	4.55	4.31	7	23	18	1.28

RESULTS

The phonological indices used are shown in Table 1, along with the average values for Voice and Watson's other behavioural measures for the different language groups.

Voice correlates highly significantly with V ($r = 0.77$, $p < 0.001$, *d.f.* = 13, two-tailed). As the number of contrastive vowels increases, the average level of speech volume decreases. This relationship is shown in Figure 1.

The correlation of Voice and V/C is slightly higher ($r = 0.83$, $p < 0.001$, *d.f.* = 13, two-tailed). However, in this sample of languages there is much more variation in vowel systems than consonantal systems. V/C in fact correlates almost perfectly with V ($r = 0.95$). It is therefore clear that it is the number of vowel contrasts which is driving the correlation.

The correlation between Voice and C is not significant ($r = -0.39$). The correlations between V and the other behavioural measures are all positive, though none is significant.

DISCUSSION

Speech volume is clearly related to phonological structure. There is, however, considerable variation between the different nationalities that make up the English, Spanish, Arabic and Hindi/Urdu groups. Part of this may be due to the fact that the sample size for each nationality was very small (2 – 16). Confounding variables such as degree of intimacy of the subjects may thus have a significant effect on the national average.

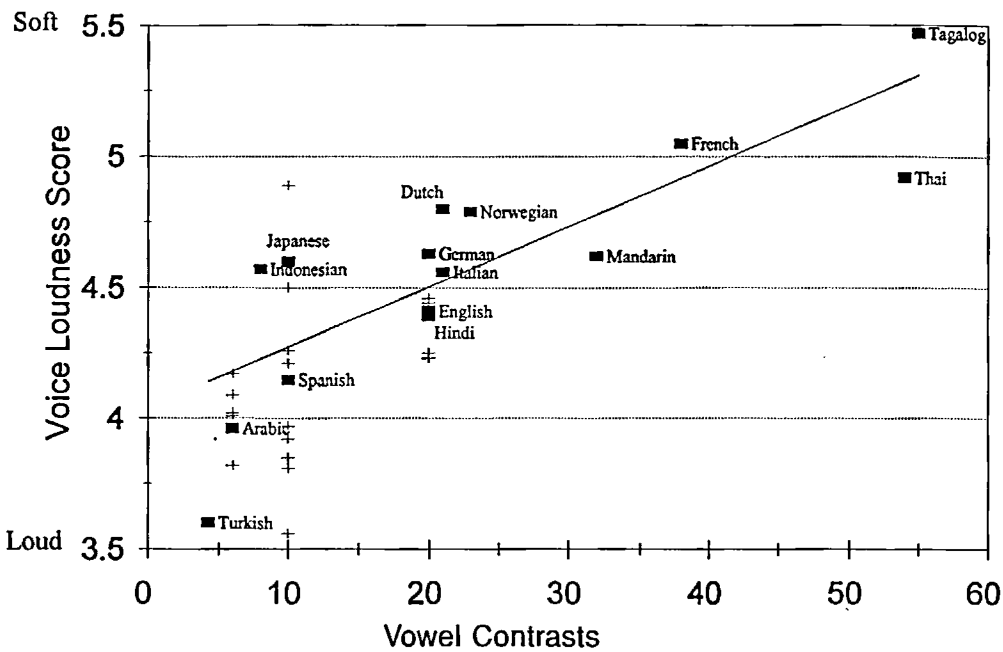

Fig. 1. Average voice loudness score for the 15 language groups plotted against V, the total number of vowel contrasts in the language. A low score indicates a loud voice, and a high score a soft one. Where a language is represented by several nationalities (Arabic, Spanish, Hindi/Urdu, English) the scores for the separate national groups are shown as a small cross to indicate the spread of points.

Error may also have been introduced from the phonological side, as the compilers of the inventories used may differ in their approach to phonetic diphthongs and consonantal clusters. The indices used are probably too crude to pick up all the relevant aspects of phonological structure.

Across the fifteen languages, the consonantal inventories are much less variable in size than the vowel inventories, differing at most by a factor of two. It follows that an increase in the vowel inventory will usually decrease the functional load borne by the consonantal contrasts and increase that borne by the more sonorous vowels. This seems likely to be the explanation for the decrease in volume as V increases.

We may ask what significance these findings have in terms of the constraints on the phonological structure of languages. Having a larger vowel inventory clearly permits softer speech and thus saves energy. However, there are plenty of languages with very small vowel systems. Must we then conclude that some languages are simply more efficient than others, and abandon the functionalist tenet (cf. Lindblom, 1984) that every aspect of linguistic structure has a communicative advantage?

The conclusion might be avoided if there were some benefit to having a small vowel system which compensates for the loss in perceptual salience. One possibility is that there is an increase in articulatory ease. Manuel (1990) has found that speakers of languages with small vowel systems exhibit more coarticulation than speakers of languages with large ones. Small systems therefore have the advantage that fluency in production is less inhibited by the need to keep a large number of segments distinct.

Another possible advantage is robustness. Andersen (1988) has argued that the maintenance of elaborate systems requires high-fidelity transmission, which is only possible where the speech community is relatively closed. Dialects used as vehicles of communication over large areas or between different groups of people will tend to develop simple systems, because of the need for reliable decoding despite accent differences. Such an effect is likely to apply to vowel systems more strongly than consonantal systems, as the articulatory vowel space is continuous, whereas the consonantal space is to some extent quantal (Stevens, 1989). This means that slight differences in pronunciation are less likely to make consonants lose distinctiveness than vowels. However, although such a factor may have been historically important, it is unclear if it could explain the current differences among the fifteen languages in this study, all of which are used over large areas and by millions of people.

Perhaps most importantly, it should be remembered that languages have a social-indexical function. The adoption and maintenance of norms may often be more socially then linguistically motivated. After all, if the tendency to perceptual-acoustic and articulatory efficiency were entirely unconstrained, the existence of so many different phonological systems in the world would be inexplicable.

(Received January 27, 1994; accepted August 16, 1994)

REFERENCES

ANDERSEN, H. (1988). Center and periphery: Adoption, diffusion and spread. In J. Fisiak (Ed.), *Historical Dialectology: Regional and Social* (pp. 39 – 84). Berlin: Mouton de Gruyter.

CAMPBELL, G.L. (1991). *Compendium of the World's Languages*. 2 vols. London: Routledge.

LINDBLOM, B. (1984). Can the models of evolutionary biology be applied to phonetic problems? In M.P.R. Van den Broecke and A. Cohen (Eds.), *Proceedings of the Tenth International Congress of Phonetic Sciences* (pp. 67 – 81). Dordrecht: Foris.

LINDBLOM, B. (1986). Phonetic universals in vowel systems. In J.J. Ohala and J.J. Jaeger (Eds.), *Experimental Phonology* (pp. 13 – 44). Orlando, FL: Academic Press.

LINDBLOM, B., and MADDIESON, I. (1988). Phonetic universals in consonant systems. In L.M. Hyman and C.N. Li (Eds.), *Language, Speech and Mind* (pp. 62 – 80). London: Routledge.

MADDIESON, I. (1984). *Patterns of Sounds*. Cambridge, UK: Cambridge University Press.

MANUEL, S. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *Journal of the Acoustical Society of America*, **88**, 1286 – 1298.

STEVENS, K.N. (1989). On the quantal nature of speech. *Journal of Phonetics*, **17**, 3 – 45.

WATSON, O.M. (1970). *Proxemic Behaviour: A Cross-Cultural Study*. The Hague: Mouton.