ON THE PHONETICS OF SMALL VOWEL SYSTEMS: EVIDENCE FROM AUSTRALIAN LANGUAGES

Andrew Butcher
Department of Speech Pathology
Flinders University, South Australia

The majority of indigenous Australian languages have a 'triangular' system of vowel quality contrasts, said to consist of /i/~/a/~/u/ only. Formant frequency data, both previously published and recently gathered, suggest that the basic phonetic space for Australian vowels is smaller than is generally taken for granted in the literature on the phonetics of vowels. Two possible explanations for this are examined. (1) Australian aboriginal speakers may have vocal tracts which are of different proportions from those of European speakers, on whom conventional notions of phonetic vowel space are based (2) Australian languages may not conform to the principle of 'maximum dispersion', which is widely assumed to be a universal one, whereby the vowels of a language are said to be dispersed maximally and evenly within the available phonetic space.

INTRODUCTION

One of several unusual characteristics of Australian languages (though perhaps not, at first glance, a particularly exciting one) is that over 50% of them have a 'triangular' system of vowel quality contrasts. In most accounts, these have been simply described as two HIGH vowels, one FRONT UNROUNDED and one BACK ROUNDED, and one LOW CENTRAL vowel (Busby, 1980:96f). Opinions differ as to how common the three-vowel system is amongst the world's languages. According to Crothers (1978:105), this is the third most frequently occurring system, found in 13.3% of the Stanford corpus. According to Maddieson (1984:127), on the other hand, only 5.4% of languages in the UPSID database have just three distinct vowel qualities. The five-vowel /i/~/e/~/a/~/o/~/u/ pattern, on the other hand, (easily the most frequently occurring in the world) is much less common in Australia (although its prevalence has perhaps been underestimated at 9% of languages - cf Busby, 1980:97ff).

THE DATA

A typical set of formant frequency data from an Australian 3-vowel system is shown in Figure 1. These are mean values from a single female speaker of Warlpiri. Each data point represents the average of 84 tokens. Generally there were six tokens for each possible segmental environment, defined in terms of preceding and following consonantal place of articulation. Each point is enclosed by an ellipse representing two standard deviations on either side of the mean. Areas corresponding approximately to the eight primary Cardinal Vowel qualities are indicated within dotted lines (Ladefoged 1967:88f). This pattern, besides being typical of my own data on 9 Australian 3-vowel languages, is also confirmed by previous published data (e.g. Busby 1979:186ff, Treffry 1984). All of these languages appear to have very compact and rather symmetrical vowel spaces, 'anchored' by the LOW central vowel /a/, which, in terms of mean values, has a rather consistent realization in an area of the vowel space centred on an F_1 at around 800 Hz and F_2 at around 1500 Hz. These high-first-formant sounds have a quality which could be symbolized phonetically as [ɐ] and which is very similar to the /ʌ/ of many accents of English. Three-vowel systems contrast this with two sounds with a lower first formant, centred in the area of 450 to 500 Hz. These two sounds in turn contrast with one another along the second formant dimension: one has a high F_2 (generally between 1900 and 2400 Hz) and the other a lower F_2 (in the area 800 to 1100 Hz). This means that these 'HIGH' vowels have a phonetic quality of [ȩ] and [ɔ̜] respectively - similar to but mostly lower than the /ɪ/ and /ʊ/ of many English accents.

The point to note about these vowel realisations is that they don't seem to be nearly as widely spaced as much of the literature would predict (e.g. Liljenkrants & Lindblom 1972, Crothers 1978:125ff, Disner 1984). According to these authors the vowels of a language are said to be dispersed maximally and evenly within the available phonetic space, i.e. realised by maximally different configurations of the vocal tract. The kind of space generally assumed to be universal would be one which extended approximately

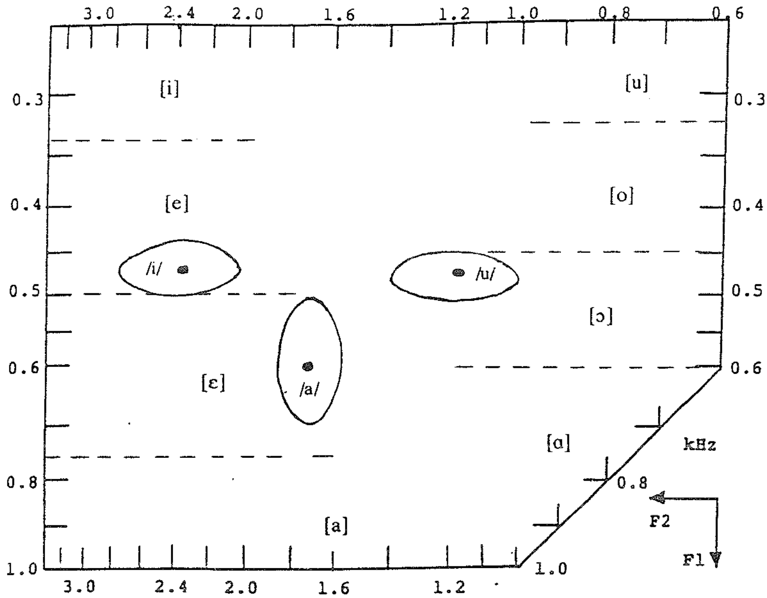

Figure 1. Formant plot of stressed short vowels in a typical 3-vowel Australian language (Warlpiri)

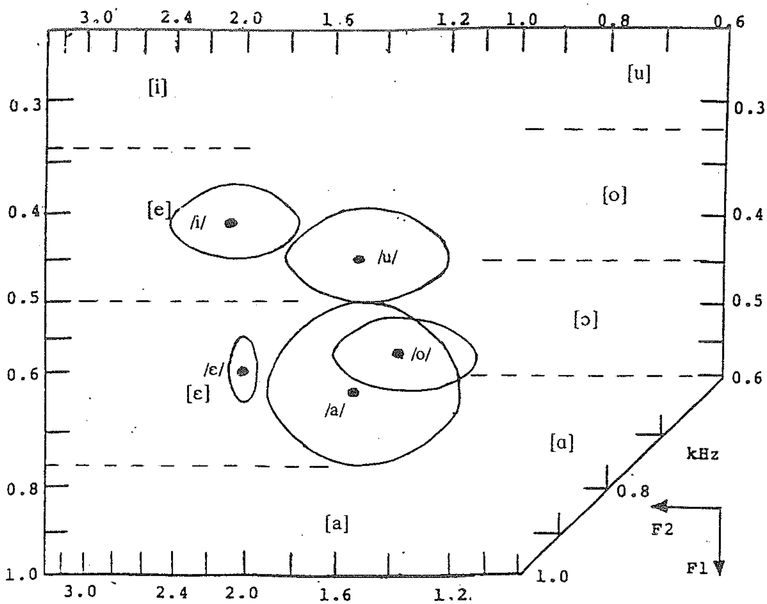

Figure 2. Formant plot of stressed short vowels in a typical 5-vowel Australian language (Burarra)

from 200 to 1000 Hz on the first formant axis and from 700 to 2600 Hz on the second formant axis. Vowel spaces for the 3-vowel Australian languages seem to use a space running roughly from 450 to 800 Hz on the F_1 dimension and from 800 to 2400 Hz on the F_2 dimension. Thus while speakers appear to be using more or less the full range of second formant values, the space is significantly 'flattened' in the first formant dimension.

VOCAL TRACT CONSTRAINTS?

One possible explanation for this apparent difference in the extent of the phonetic vowel space between Australian languages and most others described in the literature is that Australian aboriginal speakers may have differently proportioned vocal tracts from those of speakers of other languages - in particular European languages - on which most such descriptions are based. Proffit & McGlone (1975:217), for example, present measurements based on x-ray photographs of vocal tracts which suggest that male Warlpiri speakers (n=10) have an oral cavity which is somewhat longer than that of their (presumably white) American counterparts (n=7), and a pharyngeal cavity which is considerably shorter (apparently by some 30%). It is possible that such a difference in relative pharyngeal cavity length could give rise to a difference in vowel space, in particular with regard to the degree of variation in the first formant frequency. This is not a directly testable hypothesis, but it has a corollary which is: If aboriginal speakers' vowel space is constrained by the proportions of their vocal tracts when speaking a language with only three vowels, then speakers of aboriginal languages with larger vowel systems and aboriginal people speaking English should have a similarly small phonetic vowel space.

Mean values and standard deviations from speaker of a typical Australian five-vowel language are reproduced in Figure 2. These values are are based on a similar number of measurements to those taken for Warlpiri, and are once again consistent with those of the other two five-vowel languages studied and also with previously published data (e.g. Busby's 1979:190f, Trefry 1983:23) These configurations too are quite compact and there is often considerable overlap, as between the /o/ and /a/ realisations of this speaker. However, whereas /a/ has very similar formant values to the /a/ of the smaller systems, both /i/ and /u/ are closer than the HIGH vowels of those systems (the F_2 values of the mid vowels /e/ and /o/ falling roughly half way between these two levels). This results in a pattern of realizations with the approximate qualities [e], [ɛ], [ɐ], [ɔ], [o]. In other words the basic 5-vowel space is an 'upwards' extended version of the 3-vowel space, with the same low central anchor vowel, but with higher HIGH vowels. In this space, the /i/ vowel has a mean F_1 centred around 370 to 420 Hz, with the /u/ often slightly more open, in the area of 390 to 450 Hz. These sounds have [e] and [ọ̈] qualities which are the same as /ɪ/ and /ʊ/ in many English accents. The 'mid' vowels /e/ and /o/ are usually equidistant between the realizations of the HIGH and the LOW phonemes along the F_1 dimension, with mean values at around 550 to 600 Hz. They thus have phonetic qualities in the [ɛ] and [ɔ] areas, the former quite close to Australian English /æ/ (and RP /ɛ/) and the latter near to the /ɒ? of many English accents. Four-vowel systems, of which there are a few (11% according to Busby (1980:97,99), appear to use a mixture of these two spaces, extended at the front to accommodate the mid vowel, but compact at the back, with the same general position for ?u/ as in the three vowel systems. Systems with more than five vowels are somewhat rare. Thus far I have only looked at the six-vowel system of Mpakwithi. This language seems to use the basic 5-vowel space, with the extra LOW FRONT /æ/ vowel squeezed in by extending the space very slightly 'leftwards' along the F_2 dimension.

I have as yet very limited data on the pronunciation of English vowels by indigenous Australian speakers, but Figure 3 compares mean values from 5 'corner' vowels of English, as pronounced by a single female Eastern Arrernte speaker, with Bernard & Manell's (1986) data for the same five vowels as spoken by 63 male native speakers of Australian English. In both cases citation form monosyllabic words were used ending in /-d/. Eastern Arrernte has three vowel quality contrasts (arguably reducible to two). Australian English has at least ten contrasting monophthongs. Clearly, however, the indigenous speaker is employing a formant frequency space of very similar size (and shape) to that found for the native Australian English speakers (despite the gender difference). This space extends approximately from 300 Hz to just below 800 Hz on the F_1 dimension, and from 900 Hz to 2400 on the F_2 dimension. Thus there appear to be no differences between the vowel spaces of indigenous and non-indigenous speakers, as would be predicted by a hypothesis based on assumed differences in pharyngeal cavity length.

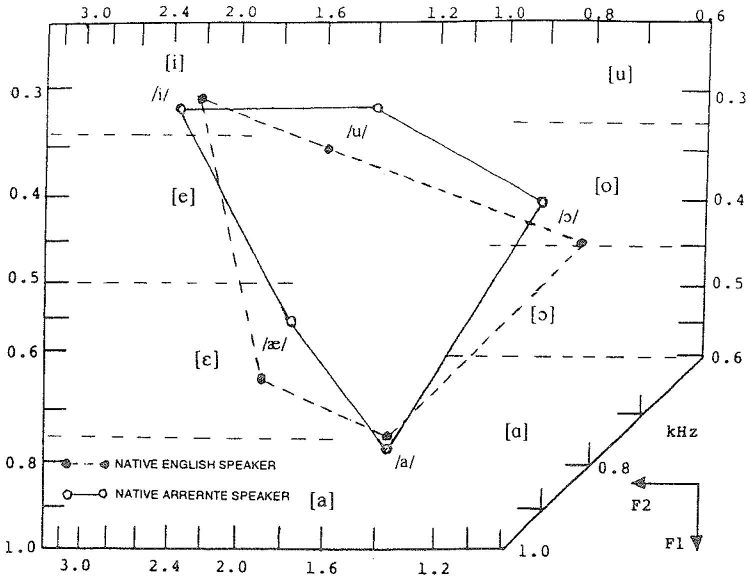

Figure 3. Formant plots of 5 stressed vowels of English as spoken by Aboriginal and non-Aboriginal speakers

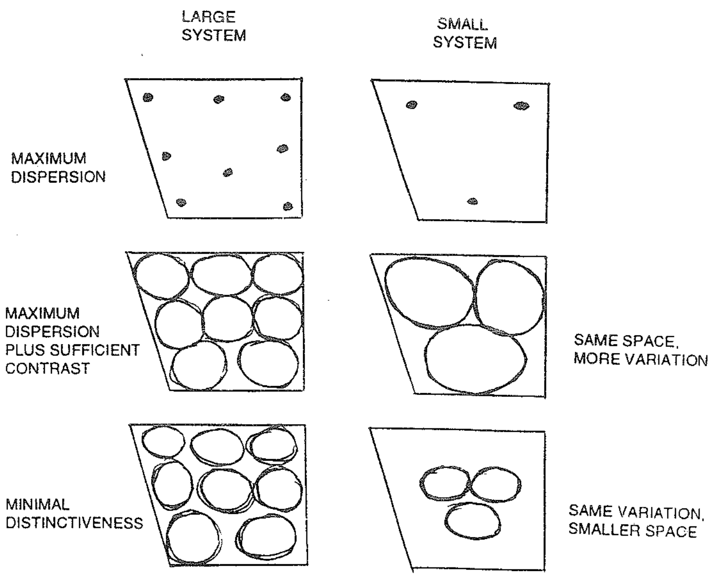

Figure 4. Opposing principles of phonetic patterning in natural vowel systems

MAXIMUM DISPERSION?

The principle of 'maximum dispersion' is fairly widely accepted (e.g. Liljenkrants & Lindblom 1972, Crothers 1978:125ff, Disner 1984). According to this principle the vowels of a language will be dispersed maximally and evenly within the available phonetic space, i.e. realised by maximally different configurations of the vocal tract. Thus, even speakers of languages with small vowel systems would be expected to spread out their vowels over an acoustic space of the size usually assumed in the literature.

The general impression from the data on Australian three-vowel systems, however, is that the phonetic realisations, as defined by the mean first and second formant frequencies, are not widely dispersed within the available acoustic phonetic space. In seeking evidence from other three-vowel systems outside Australia, it becomes clear that data on the phonetics of small vowel systems is somewhat scanty. But, what there is suggests that there is a good deal of variation in the phonetic spaces of these systems and that the pattern revealed by the Australian data may not, after all, be so unusual. Ahmad & Butcher (1987:161), for example, found that the short-vowel system of Iraqi Arabic had very similar realizations in the speech of their three (male) subjects, with /i/ at around F_1 500 Hz\F_2 1900 Hz, /a/ slightly higher than in the Australian systems at 650 Hz/1700 Hz and /u/ rather more rounded at about 500 Hz/750 Hz. Data on the 4-quality system of short vowels in Navajo, recorded by McDonough, Ladefoged & George (1993), show a somewhat larger space, however, with (in women speakers) /i/ at F_1 400 Hz/F_2 2000 Hz, /a/ at 800 Hz/1300 Hz, and /ɔ/ (the closest back vowel) at approximately 550 Hz/1200 Hz.

MINIMAL DISTINCTIVENESS?

There seems to be some evidence, then, that these vowel contrasts do not conform to the principle of maximum dispersion. In fact a plausible explanation of the data reviewed in this paper would involve invoking a diametrically opposite principle, namely one of *minimal distinctiveness.* According to this principle a language would realise its basic vowel quality distinctions in terms of the minimum excursions from the neutral vocal tract configuration consistent with maintaining the necessary perceptual contrasts. (Note that this is not the same as Lindblom's (1986:33) criterion of 'sufficient contrast', which says that variation in vowel articulations will be limited mainly by the necessity to keep phonemes distinct. This criterion is seen as compatible with the maximum dispersion principle and predicts greater phonetic variation within smaller systems). Maximum dispersion predicts that within a universal vowel space, the mean realisations of vowels in small systems will be acoustically further apart from one another than those in larger systems. Minimal distinctiveness predicts that smaller systems will use smaller spaces than larger ones, and that the acoustic distances between adjacent vowels will remain about the same across systems of varying sizes. The differences between these predictions are shown diagrammatically in Figure 4. As the diagram makes clear, the magnitude of the difference between the outcomes predicted by the two principles increases in inverse proportion to the size of the system - i.e. large systems take up a lot of space, whichever principle is in operation. The fact that the possibility of a minimal distinctiveness principle has been overlooked is thus perhaps a reflection of a lack of interest in the phonetics of small vowel systems. As can be seen from the two lower right-hand diagrams in Figure 4, support for such a principle depends crucially on the amount of variation to be observed in the realisation of contrasts in small systems.

Much of the argument in favour of maximum dispersion appears to be based on the assumption that even in the smaller (5- and 3-vowel) systems, the closest vowels are phonetically [i] and [u]. Maddieson (1984:125) says that 91.5% of UPSID languages have /i/ and 83.9% have /u/, whereby these symbols are said (p.204) to denote "high" (as opposed to "lowered high" or "higher mid") vowels. Crothers (1978) makes much the same assumption and, when discussing the phonetic value of these symbols in three-vowel systems, states that "[t]hese oppositions are not simply relative; they involve specific areas of the vowel space; it is just that in a three vowel system the areas are larger and more vague than in a more complex system". And yet, even in terms of the coarse impressionistic transcription that forms the input to these two databases, it is clearly not the case that the majority of three-vowel systems have /i/and /u/. Of the 28 three-vowel languages, 25% are Australian, 5 of which have been shown not to have such close vowels in this study and one (Nungubuyu) in which the close vowels are actually transcribed as /ɪ/

and /ɯ/. Of the non-Australian languages, another 25% of the total have /ɪ/ and /ʊ/, and 18% have /e/ and /o/. A further 14% have /i/ but /ʊ/, /ɯ/ or /o/, leaving just 18% (four languages) with /i/ and /u/ as their closest vowels. Even on the evidence of the transcriptions, it seems open to question whether the principle of maximum dispersion is operating in the majority of the world's three-vowel systems.

REFERENCES

Ahmad, K. & Butcher, A.R. (1987), Some acoustic and aerodynamic characteristics of pharyngeal consonants in Iraqi Arabic. *Phonetica* 44, 156-172.

Bernard, J.R. & Mannell, R.H. (1986), A study of /h-d/ words in Australian English. *Working Papers, Speech Hearing & Language Research Centre.*

Busby, P.A (1979) A classificatory study of phonemic systems in Australian Aboriginal languages. MA thesis, Australian National University, Canberra.

Busby, P.A. (1980) The distribution of phonemes in Australian Aboriginal Languages. *Papers in Australian Linguistics* 4, (Pacific Linguistics: Canberra) A-60, 73-139.

Butcher, A.R.(1982) Cardinal vowels and other problems. In Crystal, D. (ed): *Linguistic Controversies,* (Edward Arnold: London), 50-72.

Crothers, J. (1978) Typology and universals of vowel systems. In Greenberg, J.H., Ferguson, C.A. & Moravcsik, E.A. (eds), *Universals of Human Language, Vol 2: Phonology*, (Stanford University Press: Stanford, CA) 93-152.

Disner, S.F. (1984) Insights on vowel spacing. In Maddieson (1984), 136 -155.

Dixon, R.M.W. (1980) *The Languages of Australia*, (Cambridge University Press: Cambridge, UK).

Fant, G. (1973) *Speech Sounds and Features*, (The MIT Press: Cambridge, MA).

Ladefoged, P. (1967) *Three Areas of Experimental Phonetics*. Oxford: Oxford University Press.

Liljenkrants, J. & Lindblom, B. (1972) Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* 48, 839-862.

Lindblom, B. (1986), Phonetic universals in vowel systems. In Ohala, J.J. and Jaeger, J.J. (eds) *Experimental Phonology*. Orlando, FL: Academic Press, 13-44.

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

McDonough, J., Ladefoged, P. & George, H. (1993) Navajo vowels and phonetic universal tendencies. *UCLA Working Papers in Phonetics* 84, 143-150.

Proffit, W.R. and McGlone, R.E. (1975) Tongue-lip pressures during speech of Australian Aborigines. *Phonetica* 32, 200- 220.

Trefry, D. (1983) Discerning the back vowels /u/ and /o/ in Burarra, a language of the Australian Northern Territory. *Working Papers of the Speech and Language Research Centre, Macquarie University* 3/6, 19-51.

Trefry, D. (1984), Diari segmental phonology. *Papers in Australian Linguistics* 16, Canberra: Pacific Linguistics A-68, 171- 327.