ACOUSTICAL DATA BASE AS A TOOL FOR THE RESEARCH OF VOWEL SYSTEMS

Antti Iivonen

Department of Phonetics, Helsinki, Finland

1. ABSTRACT

A formant data base representing ca. 20 languages has been collected. The main purpose is to use this data base for a *comparison of language specific vowel qualities and vowel systems*, but it also can be used as a research tool to *avoid sources of errors due to research methods and materials.*

For comparison of vowels, an *F1/F2-plot on a Bark-scale* has been utilized. This representation can be considered to be an approximation for psycho-acoustical vowel space. The vowels are presented as *1 Bark-sized circles* in order to show the auditive distances between them (cf. [3]).

2. APPROXIMATION OF THE PSYCHO-ACOUSTICAL VOWEL SPACE

Several studies have indicated that the simulation of vowel space using the first two formants F1 and F2 on a Bark scale is a strong approximation of vowel perception. Fig. 1 shows a gliding vowel series [i-e-ɛ-æ-a-ɔ-o-u], produced as a continuous utterance by the author according to the Finnish articulation base. The glide was analyzed in 20 ms steps, and it forms a trajectory on the F1/F2-plot which correponds well to the traditional location of vowel qualities on a vowel quadrilateral. FFT spectra (with a 30 ms time window) were used. The first approximation of the glide is presented in Fig. 1.

The power of the F1/F2 representation might be explanable on the basis of motor perception theory: motor facts correspond to the perceptual ones in the sense that the listener ´hears´ an F1/F2-pattern as a corresponding tongue/lip gesture. A vowel quality is presented as a freely mobile, 1-Bark-sized circle on the F1/F2-plot. Its position is calculated from its measured Hz-values. A Bark circle can be understood as a point that scans its surrounding space to check if there is per-

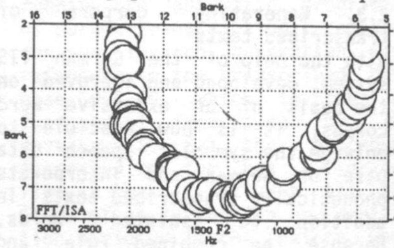

FIG. 1. Continuous vowel series [i-e-ɛ-æ-a-ɔ-o-u] produced by the author. The circles represent the F1/F2 points in 20 ms time intervals.

ceptually distinct psycho-acoustic distance from other vowels. In most cases a circle covers an area smaller than the distribution of the single occurrences of the same phoneme. It can be assumed that if two circles overlap, the listener may have difficulty distinguishing the vowels considered. Fig. 2 shows the East Central Bavarian vowel means measured by Traunmüller [11]. The rounded, front vowels represent earlier lateral sounds in the dialect. According to Traunmüller's auditive judgement, it is difficult to distinguish the vowels of the pairs [e, ɛ], [ø, œ], and [o, ɔ]. This effect has the correspondence in Fig. 2: The circles of these vowels overlap.

Lindblom [6] calculated the first four formant values of 19 "quasi-cardinal vowels" representing psycho-acoustically equal quantization steps. Fig. 3 presents these vowels on an F1/F2-plot. It can be seen that the vowels are mainly equidistant on the plot concerning each formant separately, but the distances are generally greater for F2 than for the F1.

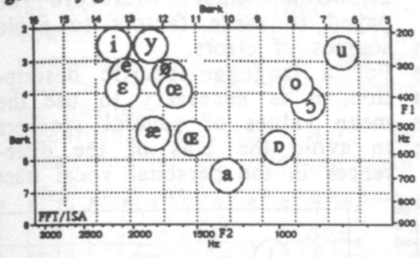

FIG. 2. East Central Bavarian vowels. Data from [11]. Means of three speakers. Note that some mid-vowels overlap which corresponds to auditory confusion.

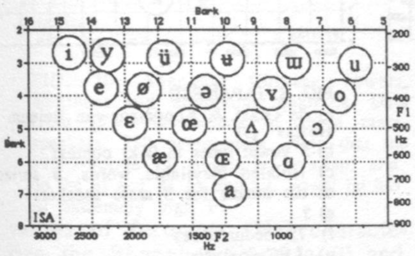

FIG. 3. Theoretical quantizing of a vowel space, formant data from Lindblom [6]. The positioning of 19 quasi-cardinal vowels.

Fig. 3 also illustrates that an empty space remains between the vowel circles. This is understandable if, for example, we consider the total number of the possible

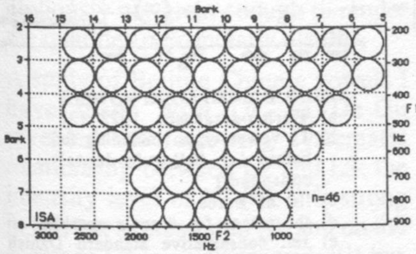

FIG. 4. If the articulatorily possible vowel space is filled with 1 Bark-sized circles, the result is 46 circles. This number seems to be near the amount of universally distinguishable vowels.

phonetic vowel symbols. The newest IPA chart (1989/1990) contains 25 vowel signs, the Stanford Phonological archive has even more: 37. Articulatorily, the number of possible vowels is unlimited. Hence, the explanation for the maximal number of vowel qualities must lie in the human auditive capacity. When the articulatorily possible F1/F2 vowel space is filled with 1 Bark-sized circles, the result is 46 circles (Fig. 4). This number corresponds well to the number of the possible vowel qualities (quoted above) (if height, frontness and rounding are considered; cf. [5]).

A diphthong can be depicted as a F1/F2-glide from beginning to end (excluding the transitions). The glide can be measured in 10 ms intervals (Fig. 5).

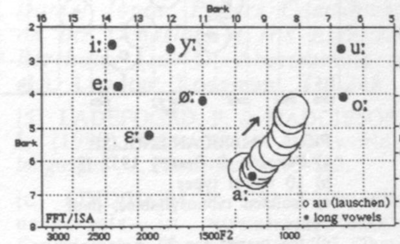

FIG. 5. One possible way to present a diphthong. The glide of a single occurrence has been displayed in 10 ms steps with the speaker's long vowels (=means) as the background. A North German male speaker, [aʊ] in *lauschen*.

3. LANGUAGE DATA

Most of the formant data included in this study have been collected from the literature. The following are the main features that were included in the data base: *a*) author(s), *b*) vowel phonemes (allophones, types) considered, *c*) utterance type used (isolated words, list of words, the carrier sentence, etc.) plus the consonant context of vowels, *d*) number and sex of the informants, *e*) language (dialect, regiolect, sociolect), *f*) number of occurences, *g*) equipment utilized for analysis, *h*) formant measurement principles, and *i*) formant values. Figures 6-9 illustrate four language examples. Considering the research features, it can be argued that very few language comparisons can be made without a bias that is a result of the differences in research methods.

4. CRITICAL REMARKS

In some cases, the representation based on F1/F2 proves to be problematic. The areas of concern are (1) the non-phonemic factors influencing the vowel positioning on the F1/F2-plot and (2) the role of formants in vowel quality characterization.

3.1. Non-phonemic factors

F1/F2 positioning is influenced by vowel reduction (due to the stress degree), vocal tract length, larynx height, allophonic variation, several voice quality types, pure chance, and formant measurement principles. Special attention must therefore be payed to these factors to avoid sources of errors.

For a language-specific description, it is necessary to use the mean values of several speakers to avoid the bias of the differences in the personal vocal tract length. The means are also necessary, because the single occurrences show considerable variation.

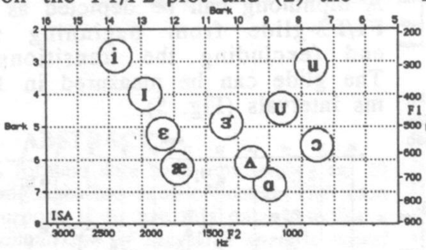

FIG. 6. AMERICAN ENGLISH
a) Peterson & Barney 1952 [8]
b) 10 vowel types
c) isolated monosyllables, [h-d] context
d) the means for 33 male speakers
e) majority General American
f) 666 occurrences
g) spectrograms and sections.

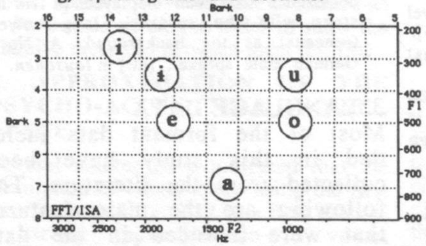

FIG. 7. POLISH
a) Jassem 1964 [4]
b) 6 vowel types
c) 44 "items" in a word list
d) the means for 3 male speakers
e) "educated Polish"
f) 132 (?) occurrences
g) Kay El. Sonagraph 661, broad band sonagrams, broad and narrow band sections

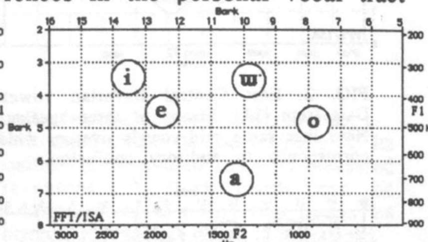

FIG. 8. JAPANESE
a) de Graaf & Koopmans-van Beinum 1982/83 [7]
b) 5 vowel types in [k-k] context
c) isolated bisyllabic words, 5 series
d) the means for 3 male speakers
e) ?
f) 75 occurrences
g) LPC-analysis

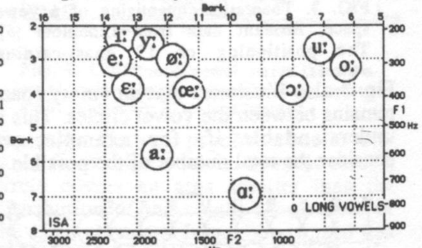

FIG. 9. DANISH LONG VOWELS
a) Fischer-Jørgensen 1972 [2]
b) 11 vowel types, including [ɑː] before /-r/, [h-l] or [h-dental consonant]
c) list of words
d) the means for 8 male speakers
e) rel. conservative standard Danish
f) 88 occurrences
g) narrow and wide band spectrograms

3.2. Difficulties in formant approach

In languages like Swedish, Danish (cf. Fig. 9), and Chinese (Fig. 10), the corner of the front, close vowels is crowded, so that parameters other than F1 and F2 may be needed for distinguishing the qualities.

According to the data in Svantesson [9], the following mean

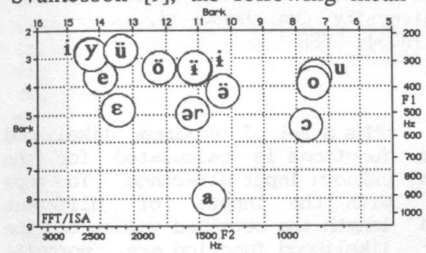

FIG. 10. Shanghai Chinese vowels, dominant and allophonic qualities. Data from [9]. The means of three male speakers, 230 occurrences. Note that the F1/F2-points of [i] and [y] overlap, and do the [i] and [ɨ], [o] and [u]. Vowel [a] has an extremely high F1.

values of formants (Hz) characterize the Shanghai Chinese [i] and [y] by three male speakers:

	F1	F2	F3
i	274	2442	3387
y	270	2455	3465
diff.	4	13	78

The differences of the formants are obviously not great enough to render the qualities perceptually distinct.

A study of Beijing Chinese vowels I have made with Dr. Li De-Gu revealed the following systematic relationship between L2 and L3: the intensity level of F2 (=L2) is strong in [y], but supressed in [i], and on the contrary L3 is very strong in [i] but weak in [y] (cf. also the discussion in [1] and [10].

REFERENCES

[1] AALTONEN, O. (1985) "The effect of relative amplitude levels of F2 and F3 on the categorization of synthetic vowels", *J. of Phonetics* 13, 1-9.

[2] FISCHER-JØRGENSEN, E. (1972) "Formant frequencies of long and short Danish vowels", *Studies for Einar Haugen presented by friends and colleagues* (ed. E. S. Rirchow et al.), The Hague/Paris: Mouton, 189-213.

[3] IIVONEN, A. (1987) "Regional differences in the realization of Standard German vowels", *Proceedings of the XIth International Congress of Phonetic Sciences, Tallinn.* Vol. 4, 161-164.

[4] JASSEM, W. (1964) "A spectro-graphic study of Polish speech sounds", *In Honour of Daniel Jones: Papers Contributed on the Occasion of His Eightieth Birthday* (ed. D. Abercrombie & al.), London: Longman, 334-348.

[5] LADEFOGED, P. & MADDIESON, I. (1990) "Vowels of the world's languages", *J. Phonetics 18*, 93-122.

[6] LINDBLOM, B. (1986) "Phonetic universals in vowel systems", *Experimental Phonology* (ed. J. J. Ohala & J. J. Jaeger), Orlando etc.: Academic Press, 13-44.

[7] de GRAAF, T. & KOOPMANS-van BEINUM, F. J. (1982/83) "Vowel contrast reduction in Japanese compared to Dutch", *Proc. of the Institute of Phonetic Sciences*, Amsterdam, 27-38.

[8] PETERSON & BARNEY (1952), "Control methods used in a study of the vowels", *J. Amer. Acoust. Soc.* 24, 175-184.

[9] SVANTESSON, J.-O. (1989) "Shanghai vowels", *Working Papers 35* Lund University, Dept. of Linguistics, 191-202.

[10] SCHWARTZ, J. L. & ESCUDIER, P. (1987) "Does the human auditory system include large scale spectral integration?", *The Psychophysics of Speech Perception* (ed. M. E. Schouten), Dordrecht etc.: Nijhoff, 284-292.

[11] TRAUNMÜLLER, H. (1982) "Der Vokalismus im Ostmittelbairischen", *Zeitschrift für Dialektologie und Linguistik*, IL. Jahrgang, Heft 3, 289-333.