A Cross-Dialect Study of Vowel Perception in Standard Indonesian

E. Van Zanten and V.J. Van Heuven
Leyden, the Netherlands

1. Introduction

Shortly after World War II Indonesia gained its independence. One of the governmental acts was to impose one variety of Malay as the official standard language, Bahasa Indonesia, throughout the archipelago. As a result the Indonesians now speak the standard language (almost as a second language) as well as a local vernacular, which though obviously related to the standard language may differ from it in many respects. As a case in point consider the vowel systems of the three vernaculars that are dealt with in the present paper:

Toba Batak (5)

i u
e o
a

Javanese (6)

i u
e ə o
a

Sundanese (7)

i y u
e ə o
a

The 6-vowel system of Standard Indonesian (disregarding diphthongs) is equal to that of Javanese. Toba Batak lacks the central mid vowel, whereas Sundanese has two central vowels, viz. one mid and one high(er).

In our study of the Indonesian vowel system we are interested in the acoustic *and* perceptual properties of the monophthongs, and the possible influence of the regional substrate on the subjects' performance. The present paper is confined to a perceptual experiment only.

Our method is based on early work by Cohen, Slis and 't Hart (1963) on the perceptual tolerances of Dutch vowels, which was later successfully extended to contrastive vowel studies by e.g. Schouten (1975) and Hombert (1979). In these latter studies subjects were presented a large number of synthesized isolated vowel sounds, regularly sampled from a vowel space essentially defined by F_1 and F_2. Their task was to label each vowel sound in terms of one of the vowels of their language. This method proved sensitive enough to reveal differences between the internal representation of British-English vowels of native speakers and that of advanced Dutch learners of English (Schouten, 1975). Hombert (1979) successfully applied the method to the description of the vowel systems of a number of - strongly related - African Bantu languages. The issue at stake in our present study is whether the

labelling method will also reliably reflect differences between the vowel systems of speakers of a standard language with different dialect backgrounds.

2. Method

On the basis of an acoustic pilot study (Van Zanten and Van Heuven, 1982), realistic formant ranges were defined for Indonesian vowels spoken in isolation. A set of 188 monophthongs were then produced with a Fonema OVE IIIb speech synthesizer whose parameter values were controlled by a DEC PDP11/03 computer. All vowels were given a 350 ms duration including linear onset and offset portions (in dB) of 50 and 100 ms, respectively. During the steady state portion voicing was set at maximum intensity. F_1 and F_2 were systematically varied in steps of 9%, i.e. 3 times the Just Noticeable Difference commonly reported for F_1 and F_2 centre frequency changes (Flanagan, 1955; Mermelstein, 1978; Nord and Sventelius, 1979), sampling the acoustic vowel space in the way indicated in Fig. 1. F_4 and F_5 centre frequencies were set at 3500 and 4000 Hz, respectively, for all vowels; bandwidths B1-B3 were set at mid-range values. The centre frequency of F_3 equalled that of $F_2 + 600$ H_2, with a minimum of 2460 Hz. Two tapes were prepared containing the set of 188 stimuli, preceded by a series of practice items, in counterbalanced random orders.

Three groups of Indonesians with different regional backgrounds participated in the experiment: 4 Toba Batak, 5 Javanese, and 4 Sundanese listeners. All the subjects had completed a university education in their own country, and had only recently arrived in the Netherlands to enroll in a postgraduate program at the University of Leyden. They participated on a voluntary basis, and were paid for their services.

Subjects were instructed to label each vowel stimulus as one of the six monophthongs of Standard Indonesian (forced choice), and to rate each token along a scale of acceptability: 1 (good), 2 (poor, but easily identifiable), 3 (unacceptable and hardly identifiable).

3. Results

After having tried out various weighting procedures, the responses were finally analysed such that 'good' tokens counted twice, 'poor' tokens once, and 'unacceptable' tokens were eliminated altogether. Figure 1 plots the data for the three listeners groups (panel A: Toba Batak; panel B: Javanese; panel C: Sundanese). Areas of preference were defined containing only those stimulus points that were identified as one particular vowel in at least 50% of the responses (small letters), and in at least 75% (large letters). To facilitate the exposition, summary statistics are given in Table I, specifying the number of stimulus points (absolute and relative) contained by each eara of preference.

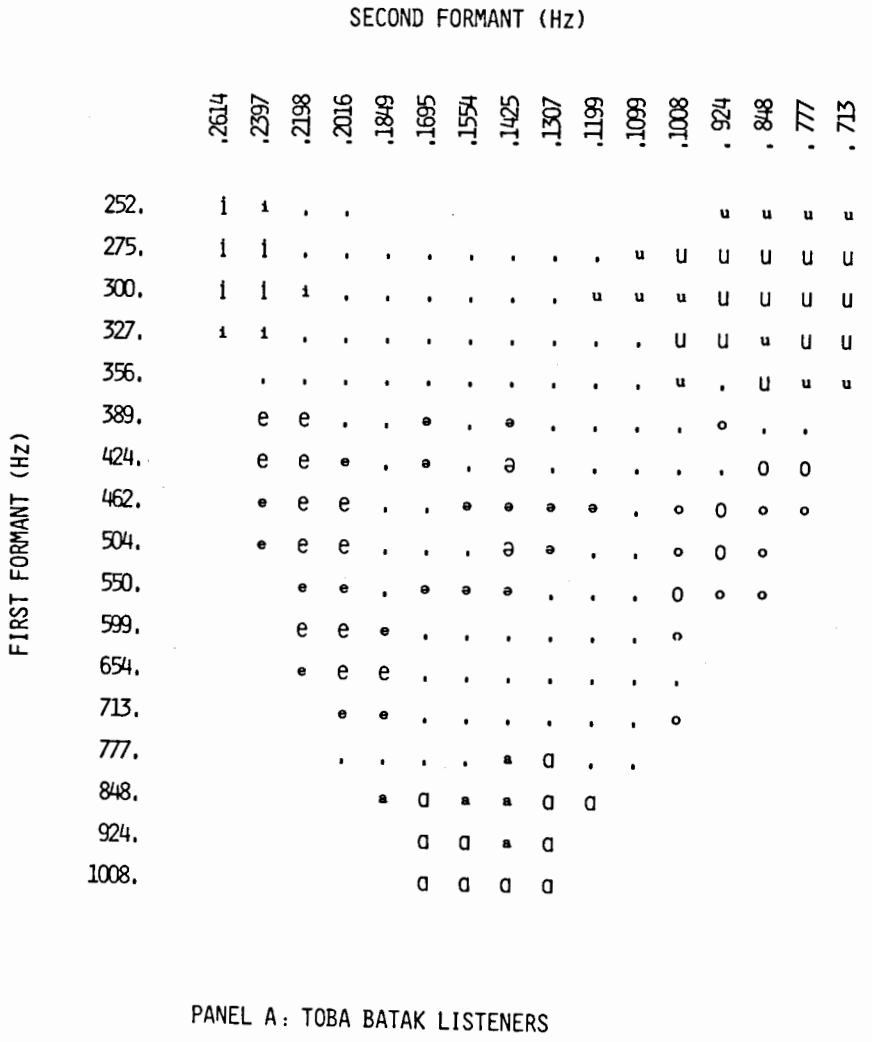

Figure 1a. Distribution of responses to labelling test for three groups of listeners. Panel A: Toba Batak, N = 4, as a function of F_1 and F_2.

4. Conclusions and Discussion

There are several conspicuous differences across the three listener groups in the locations and sizes of the preferred vowel areas, specifically in the way the central region of the vowel space is divided over the competing vowel phonemes. Typically, the /ə/ area is small for the Toba Batak group, intermediate for the Javanese, and largest for the Sundanese. Conversely, the area associated with /u/ is large for the Toba Bataks, intermediate for the Javanese, and smallest for the Sundanese.

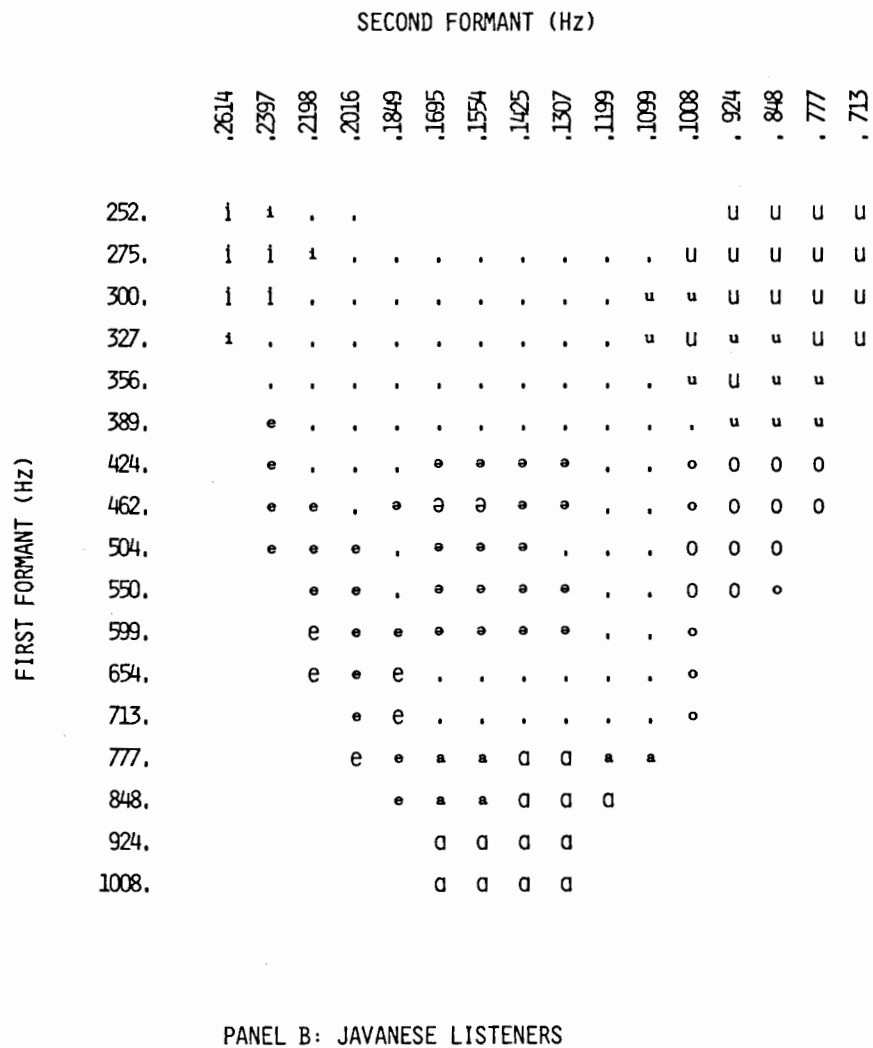

Figure 1b. Distribution of responses to labelling test for three groups of listeners. Panel B: Javanese, N = 5, as a function of F_1 and F_2.

These differences in the distribution of the responses obviously reflect properties of the subjects' regional substrate languages. Remember that the Toba Batak dialect has no central vowel, which explains why /ə/ is the least favoured response category for the Bataks. Also, its area of dispersion is highly irregular, and only 2 out of 188 stimulus points are identified as /ə/ in more than 75% of the responses.

For listeners with a Javanese background, an /ə/ dialect, the preferred area for /ə/ is appreciably larger, and the responses are more regularly distributed. For the Sundanese group, having a background dialect with two central vowels, the preferred /ə/ area is larger still, and, perhaps more

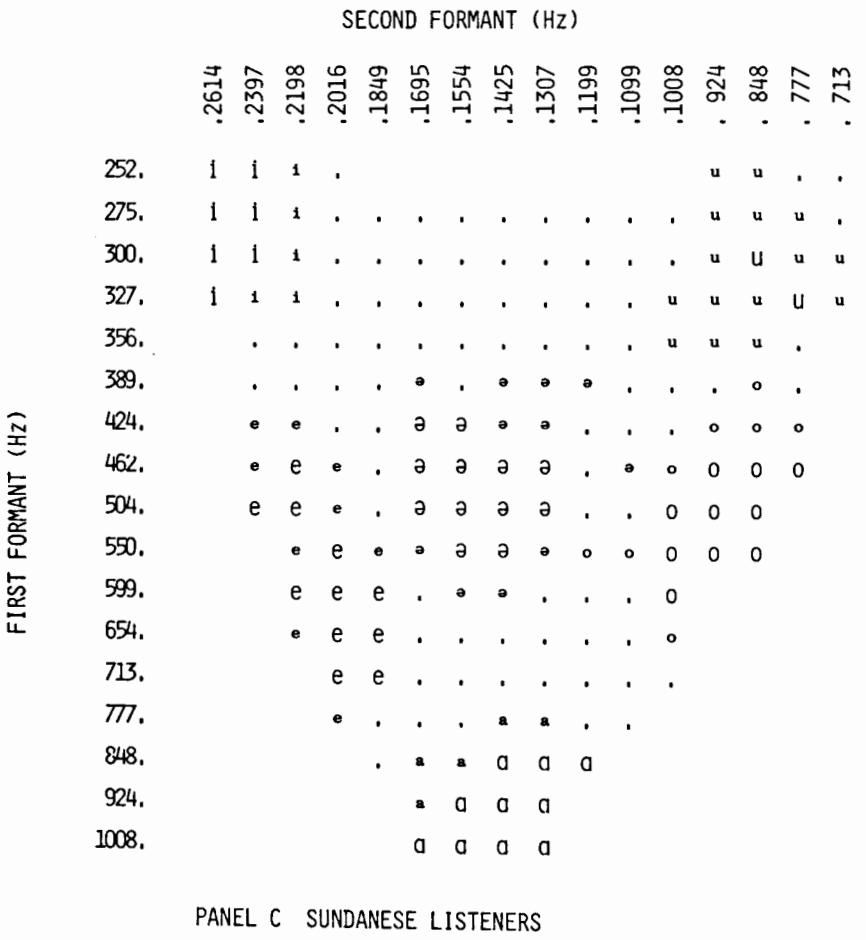

Figure Ic. Distribution of responses to labelling test for three groups of listeners. Panel C: Sundanese, N = 4), as a function of F_1 and F_2. Large letters represent stimuli identified as indicated by the phonetic symbol in at least 75% of the responses (after weighting, see text); small letters represent vowels identified with at least 50% agreement.

Table I. Number of stimulus points identified as one particular vowel in at least 50% (75%) of the responses, expressed absolutely and relatively, per vowel per group of subjects (4 Toba Bataks, 5 Javanese, 4 Sundanese)

Vowel	50% agreement			75% agreement		
	Bataks	Javanese	Sundanese	Bataks	Javanese	Sundanese
/i/	9 (5%)	8 (4%)	12 (6%)	5 (3%)	5 (3%)	7 (4%)
/e/	22 (12%)	20 (11%)	20 (11%)	12 (6%)	5 (3%)	11 (6%)
/a/	15 (8%)	19 (10%)	15 (8%)	11 (6%)	13 (7%)	11 (6%)
/o/	15 (8%)	20 (11%)	18 (10%)	5 (3%)	11 (6%)	10 (5%)
/u/	25 (13%)	25 (13%)	17 (9%)	14 (7%)	17 (9%)	2 (1%)
/ə/	13 (7%)	20 (11%)	23 (12%)	2 (1%)	2 (1%)	12 (6%)
unlabelled	89 (47%)	76 (40%)	83 (44%)	139 (74%)	135 (72%)	135 (72%)

importantly, here the distribution of especially /u/ is much more restricted: only 17 stimulus points are identified as /u/ with more than 50% agreement (against 25 for the other dialect groups), and only 2 with more than 75% (against 14 and 17 for the Batak and Javanese listeners respectively). Presumably, the high(er) central vowel (which was not a response option open to the subjects) 'pushes back the /u/ boundary', i.e. precludes /u/ responses to stimuli with F_2 values larger than 1000 Hz, whereas the preferred /u/ area extends to 1100 Hz for listeners with a Javanese background (1 neutral vowel), and even 1200 Hz for Toba Batak listeners (no central vowel).

As a final observation we would like to point out a difference in overall performance on the part of the Javanese, as opposed to the other groups: it appears that the properties of the stimulus points that could not be adequately identified (i.e. with at least 50% agreement) tends to be smaller for the Javanese (40%) than for either the Bataks (47%) or the Sundanese (44%).

Summing up then, we have shown that the perceptual method of charting a vowel system proved sensitive enough to reflect influences of the regional substrates of listeners when asked to identify vowels in terms of the categories given by their common national language.

It also demonstrates that speakers of a vernacular that is most similar to the standard language (in terms of the inventory of monophthongs) are in a better position to reach high agreement (or: consistency) in the identification task than speakers whose background dialect has a (marginally) richer or poorer inventory.

Finally, we advocate a wider use of the perceptual method outlined here to the study of vowel systems under conditions where sophisticated laboratory equipment is not available for spectral analysis. The test tape we have prepared can be administered (to a large number of subjects in parallel, if necessary) in half an hour, and provides a wealth of easily interpretable and surprisingly stable data.

Acknowledgements

Work supported in part by a grant from the Foundation for Linguistic Research, funded by the Netherlands Organisation for the Advancement of Pure Research (ZWO) under project nr 17-21-20; we thank Jos. F.M. Vermeulen for technical assistance.

We acknowledge the support of the Dept. of Phonetics at the University of Utrecht by putting equipment at our disposal.

References

Cohen, A., Slis, I.H. and Hart, J. 't (1963). Perceptual tolerances of isolated Dutch vowels. *Phonetica*, **9**, 65-78.

Flanagan, J.L. (1955). A difference limen for formant frequency. *Journal of the Acoustical Society of America*, **27**, 613-617.

Hombert, J.-M. (1979). Universals of vowel systems: the case of centralized vowels. In: E. Fischer-Jørgensen, J. Rischel, N. Thorsen (eds.). *Proceedings of the Ninth International Congress of Phonetic Sciences*, Vol. II. Institute of Phonetics, Copenhagen, 27-32.

Mermelstein, P. (1978). Difference limens for formant frequencies of steady state and consonant-bound vowels. *Journal of the Acoustical Society of America,* **63**, 572-580.

Nord, L. and Sventelius, E. (1979). Analysis and prediction of difference limen data for formant frequencies. *Phonetic Experimental Research at the Institute of Linguistics University of Stockholm, Report I: Experiments in speech perception; Phonetics Research Seminar 1978-1979*, 24-37

Schouten, M.E.H. (1975). Native-language interference in the perception of second-language vowels, Doct. Diss. University of Utrecht.

Zanten, E. Van and Heuven, V.J. Van (1983). A phonetic analysis of the Idonesian vowel system: a preliminary acoustic study. *NUSA, Linguistic Studies in Indonesian and other Languages in Indonesia,* **15**, 70-80.