ELLEN VAN ZANTEN AND VINCENT J. VAN HEUVEN

THE INDONESIAN VOWELS AS PRONOUNCED AND PERCEIVED BY TOBA BATAK, SUNDANESE AND JAVANESE SPEAKERS*

Introduction

The primary aim of this study is to give a description of the vowel systems of speakers of Indonesian of different regional backgrounds. Secondly, we will compare our data with two different vowel typologies.

The Indonesian vowel system comprises 6 vowel phonemes (disregarding diphthongs), viz. /i, e, a, o, u/ and a central vowel. The central vowel is taken by Trubetzkoy to be a closed central vowel (cf. Trubetzkoy 1929:49). According to others, it is a mid vowel (cf. Halim 1974:169; Stokhof 1975:269). Grammars of the Indonesian language usually give a rather impressionistic description of the vowel system of that language. Teeuw (1978), for instance, relates the pronunciation of the Indonesian vowels to the Dutch vowels: Indonesian *a* is roughly speaking pronounced somewhere between Dutch *aa* in *maat* and Dutch *a* in *mat*; *i* is

* The research for this paper was supported in part by a grant from the Netherlands Organisation for the Advancement of Pure Research (ZWO), project no 17-21-20 (Stichting Taalwetenschap) We would like to thank Jos F M Vermeulen for his technical assistance and Gerald G Prast for drawing the figures We are grateful to an anonymous reader of the manuscript who made a number of valuable comments

ELLEN VAN ZANTEN, a graduate from the University of Leiden, is currently a research associate in the Department of Linguistics at Leiden Her main academic interest being the phonetics of Indonesian, she is the co-author of 'A Phonetic Analysis of the Indonesian Vowel System', *NUSA* 15, 1983, pp 70-80, and 'A Cross-Dialect Study of Vowel Perception in Standard Indonesian', in *Proceedings of the Tenth International Congress of Phonetic Sciences,* Dordrecht/Cinnaminson 1984

VINCENT VAN HEUVEN obtained his Ph D from the University of Utrecht and is associate professor in the Department of Linguistics at Leiden Interested in experimental linguistics, he is co-author of 'Auditory Discrimination of Rise and Decay Times in Tone and Noise Bursts', *Journal of the Acoustical Society of America* 66, 1979, pp 1308-1315, and 'Effect and Artifact in the Auditory Discrimination of Rise and Decay Time, Speech Versus Non-Speech', *Perception and Psychophysics* 33, 1983, pp 305-313

Both authors may be contacted at the Department of Linguistics/Phonetics Lab , University of Leiden, Cleveringaplaats 1, Leiden

somewhere between Dutch *ie* in *riet* and *i* in *rit*; *u* is somewhere in the region of *oe* in *hoed*, and so on (Teeuw 1978:9). This kind of description may be perfectly adequate for teaching purposes. But a scientific description should be, and indeed can be, much more explicit.

Experimental phonetic methods can be used to give such an explicit and, possibly, objective description of speech sounds. By comparing a fair number of measurements of the physical qualities of those speech sounds, differences between realizations can be revealed which otherwise would not be easily noticeable. These differences, however small, may be meaningful if they are systematically related to linguistic facts, for instance to variations between allophones of the same phoneme or to differences between speakers of different language or social backgrounds. An experimental phonetic description of the Indonesian vowel system may thus be of interest to students of the Indonesian language.

Indonesian is spoken on the basis of a wide variety of substrate languages. Most speakers of Indonesian have an active and passive knowledge of at least one regional language. We will attempt to contribute to the description of the vowel system of Standard Indonesian as well as to determine any possible effects of a particular regional language on the standard language. We obtained the cooperation of informants from among speakers of three different regional languages for our experiments, viz. Javanese, Sundanese and Toba Batak speakers. These three regional languages differ crucially from each other as regards their vowel system: Javanese, like Indonesian, has 6 vowel phonemes (including 1 central vowel), Sundanese 7 (including 2 central vowels) and Toba Batak 5 (no central vowels).

Two mutually complementary methods were used for the collection of our data. First, 13 Indonesians from each of the three regional backgrounds read a set of Indonesian sentences. Each of these sentences included one of the six Indonesian monophthongs pronounced in isolation. Secondly, the same informants were presented with machine-generated vowel-like sounds, which they were asked to label as one of the range of Indonesian vowels. The results of the two experiments are compared with each other below. They are also compared with two typological descriptions of vowel systems.

Vowel typology

Experimental phonetic descriptions of vowel phonemes may contribute to the typology of vowel systems. In recent years a fair amount of attention has been given to the typology and structure of vowel systems. Crothers (1978:95-145) gives a phonological typology of the vowel systems of a large number of languages, together with a list of the phonetic realizations of the vowel phonemes in these languages. According to Crothers, the quality of the vowels in a given system is to a large extent determined by the number of vowels in that system. For any given number of vowel phonemes per system, only one or two arrangements occur with some frequency in the world's languages. For instance, the arrangement in the great majority of languages with a 6-vowel system is /i,ɛ, a, ɔ, u/ and /ɨ/ (closed central) or /ə/ (mid central).

According to Crothers' list, the vowels of Indonesian ("Malay" in Crothers' list, 6) and of Toba Batak (5 vowels), Javanese (8 vowel phonemes in Crothers' inventory, but 6, including the allophonic pairs [e] - [ɛ] and [o] - [ɔ], in the opinion of most other scholars (e.g. Ras 1982:3)) and Sundanese (7 vowels) occupy the following positions:

front	central	back		front	central	back
i		u	high	i		u
			lower-high			
		o	higher-mid	e		o
			mid		ə	
ɛ			lower-mid	ɛ		ɔ
			higher-low			
	a		low		a	

FIG. 1a: Toba Batak FIG. 1b: Javanese

front	central	back		front	central	back
i	ɨ	u	high	i		u
			lower-high			
		o	higher-mid	e		o
	ə		mid		ə	
ɛ			lower-mid			
			higher-low			
	a		low		a	

FIG. 1c: Sundanese FIG. 1d: Indonesian(Malay)

Figure 1.
The positions of the vowels of Indonesian, Toba Batak, Javanese and Sundanese according to Crothers (1978). Contrary to the view of most other scholars, Javanese has 8 vowel phonemes in Crothers' inventory.

Liljencrants and Lindblom (1972) have developed a model for predicting the phonetic structure of individual vowel systems based on the principle of maximal perceptual contrast (cf. Liljencrants and Lindblom 1972:839-862). In this model, vowel space is defined by the perimeters of the vocal tract: position and shape of tongue, jaw, lips and larynx. Consequently, all languages are assumed to have an identical vowel space, regardless of the number of vowels in each language. Within this acoustic space the principle of maximal perceptual contrast is applied. This principle is illustrated by Liljencrants and Lindblom with a simple physical experiment. A set of similarly oriented magnets is attached to cork floats in such a way that they are able to float on a water surface. A non-magnetic boundary is constructed at the water surface. The magnets will move away from each other until an equilibrium is reached where their distances cannot be increased any further. According to Liljencrants and Lindblom, in systems of up to 9 magnets (or vowels) this equilibrium is reached when all the magnets have taken up positions along the boundary. It is only when a 10th magnet is introduced into the floating space that one of the magnets will take a position in the centre area.

For 5-, 6-, and 7-vowel systems the Liljencrants and Lindblom model generates the following positions:

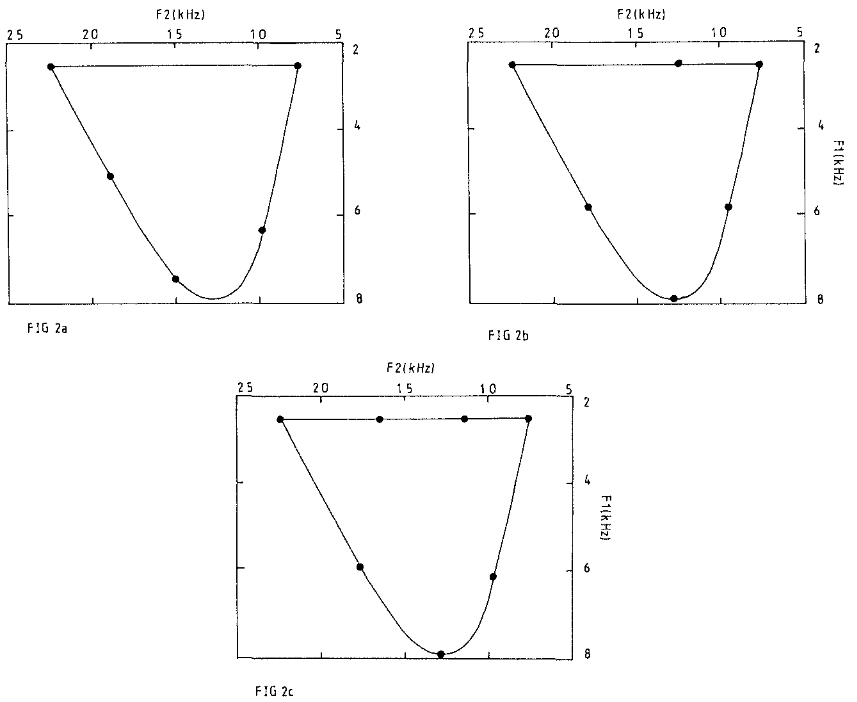

Figure 2.
The positions of the vowels in 5-, 6- and 7-vowel systems after Liljencrants and Lindblom (1972).

Note that the perimeters of the vowel space are identical in all three cases.

The predictions for /i, u, a/ correspond fairly closely to Crothers' data. The correspondence is not so close for /e/ and /o/. The predicted closed central vowel /ɨ/ in the 6-vowel system does not fit in with Crothers' scheme for Indonesian. The 7-vowel system of Figure 2 shows two closed central vowels, whereas Crothers describes the 7-vowel language Sundanese as having one closed and one mid central vowel. With our phonetic description of the vowel system observable in the pronunciation of 13 Indonesians we hope to be able to contribute to a more accurate typology.

Although this is not explicitly stated, it seems that what Liljencrants and Lindblom had in mind were the so-called target positions of vowel phonemes, i.e. the vowels as uttered not in a word but in isolation. In the two experiments which will be described presently we in fact restricted ourselves to the description and analysis of vowels produced in isolation.

The experiments

The research consisted of two experiments: a production experiment, for which the informants were asked to read a set of Indonesian sentences, and a perception experiment, in which the same informants were asked to label vowel-like sounds as one of the 6 Indonesian monophthongs.

There were 13 male informants taking part in the experiments: 4 Toba Batak, 5 Javanese and 4 Sundanese. All but one were university students or postgraduate students in The Netherlands. They all spoke their regional language more or less frequently alongside Indonesian and English and/or Dutch, but spoke no other regional language of Indonesia.

Production experiment

We restricted ourselves here to the six Indonesian monophthongs, viz. /i, e, a, o, u/ and a central vowel, leaving out of consideration the category of diphthongs. With the intention of keeping the prosody as similar as possible, all the stimuli were embedded in a fixed sentence frame, with the isolated vowels occurring in pre-pausal position at the end of the sentence, e.g. *Dalam kata (tutup) terdapat bunyi (u),* 'In the word (*tutup*) is found the sound (u)' (see Appendix). The vowels in isolation in their carrier-sentences were presented 5 times to the subjects in the same random order. The speakers were asked to read all the sentences with the same intonation. They were warned beforehand that the central vowel, which in Indonesian is written 'e', would now appear as 'ə'. All instructions were in Indonesian. The sentences appeared one at a time on a video screen. The experimenter would present the next sentence only after the informant had finished reading the last one without mistakes. The recordings were made in a sound-proof cabin on a Revox 2-track tape-recorder (19 cm/s) with a Sennheiser MKH 416 condenser microphone.

vowel	speakers	N	F1 $\bar{x}$	F1 s	F2 $\bar{x}$	F2 s	F3 $\bar{x}$	F3 s
/i/	Toba Batak	4	291	33	2190	132	3075	137
	Javanese	5	319	52	2174	59	2912	176
	Sundanese	3	274	4	2296	96	3108	150
	Indonesian	12	298	43	2210	110	3015	180
/e/	Toba Batak	4	390	37	2077	135	2727	142
	Javanese	5	506	73	1953	68	2557	58
	Sundanese	4	476	97	2071	126	2647	221
	Indonesian	13	461	88	2028	126	2637	166
/a/	Toba Batak	4	764	50	1365	54	2362	159
	Javanese	5	750	78	1345	75	2480	99
	Sundanese	4	730	138	1383	96	2559	160
	Indonesian	13	748	96	1363	78	2468	160
/o/	Toba Batak	4	496	44	869	111	2569	143
	Javanese	5	544	71	871	54	2533	57
	Sundanese	4	535	118	914	55	2591	110
	Indonesian	13	526	85	884	79	2562	109
/u/	Toba Batak	2	377	30	829	109	2449	111
	Javanese	4	378	26	856	25	2515	91
	Sundanese	4	304	27	744	77	2613	130
	Indonesian	10	348	45	806	87	2541	129
/ə/	Toba Batak	4	389	32	1794	201	2623	182
	Javanese	5	519	63	1367	140	2501	55
	Sundanese	4	425	87	1388	96	2592	276
	Indonesian	13	450	86	1505	245	2567	194

Table I.

Mean formant frequencies ($\bar{x}$) and standard deviations (s) in Hz of the 6 Indonesian monophthongs produced in isolation; 13 speakers, 5 utterances per speaker. The formant values are averaged first for the utterances per speaker and then for the speakers per dialect group and for all speakers. The standard deviation s expresses the scatter of the means over the speakers.
N: number of speakers for whom this particular vowel was analysed. /i/, /e/ . . .: as spoken in the carrier sentence with *titik, tetes* and *bebe*, . . . (see Appendix).

Analysis. The quality of vowels depends mainly on the center frequencies of the formants, i.e. groups af adjacent overtones amplified by the resonance characteristics of the vocal tract. The frequency of the lowest formant, F1, is inversely related to vowel height in a traditional vowel diagram. The second formant, F2, roughly reflects – again inversely – the degree of backness in a traditional vowel diagram. The lowest formants, F1, F2, and to a lesser extent F3, are therefore decisive in distinguishing vowels from each other.

The center frequencies of the first five formants over the steady state portion of the vowels were estimated for each of the vowel tokens by Linear Predictive Coding (LPC-analysis) and averaged per vowel token.[1] In our later analysis only the time averaged steady state formant frequencies of the first three formants (F1, F2 and F3) were taken into consideration, after some data cleaning by hand so as to remove obvious errors in the output of the LPC algorithm. In a few cases, notably /u/, it was not possible to measure the formant structure of the vowel realization.

Results. Table I sets out the means and standard deviations per group of speakers of the measured formant values, averaged first over the number of utterances per speaker and then over the speakers per dialect group. Figures 3a, 3b, 3c and 3d present these values for the 6 vowels in the F1/F2-plane.

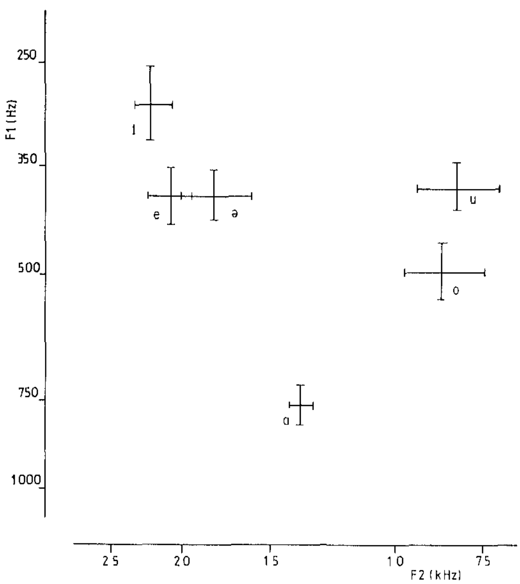

Figure 3a.
Toba Batak (4 speakers).

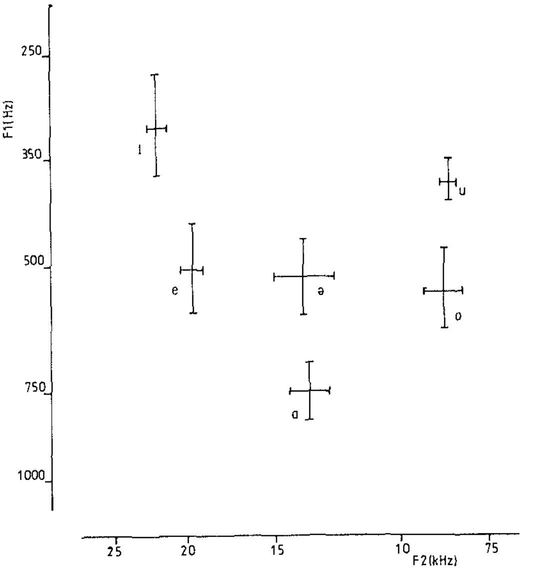

Figure 3b.
Javanese (5 speakers).

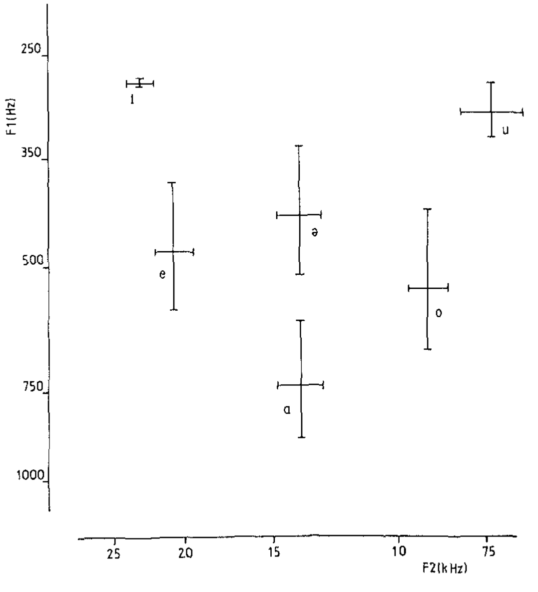

Figure 3c.
Sundanese (4 speakers).

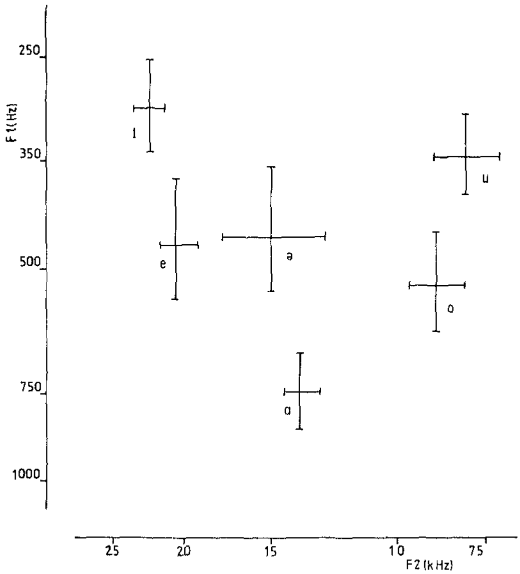

Figure 3d.
All speakers: 13 Indonesian speakers.

Figures 3a-d.
Mean F1 and F2 values (centres of crosses) and standard deviations (bars of crosses) of the formant values of the 6 Indonesian monophthongs produced in isolation; 13 speakers, 5 tokens per speaker. The formant values are averaged first for the tokens per speaker and then for the speakers per dialect group. The standard deviations express the scatter of the means over the speakers. For the number of speakers for whom a particular vowel was analysed see Table I.

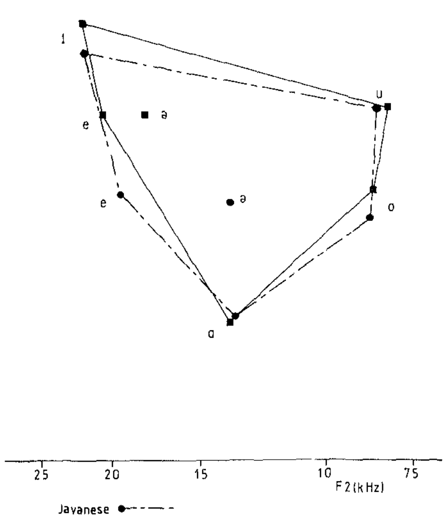

Figure 4a.
4 Toba Batak and 5 Javanese speakers.

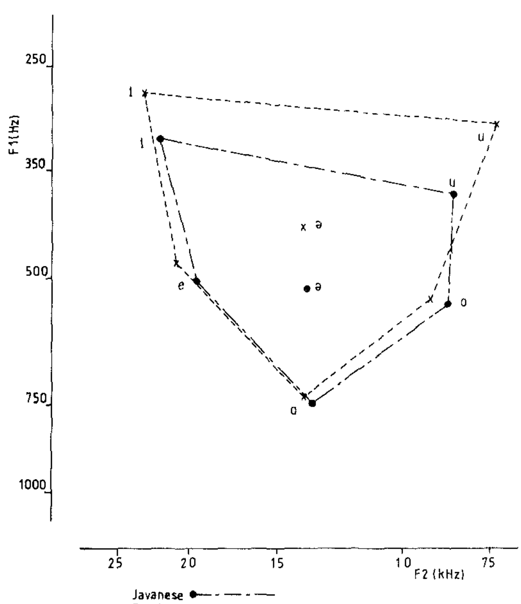

Figure 4b.
5 Javanese and 4 Sundanese speakers.

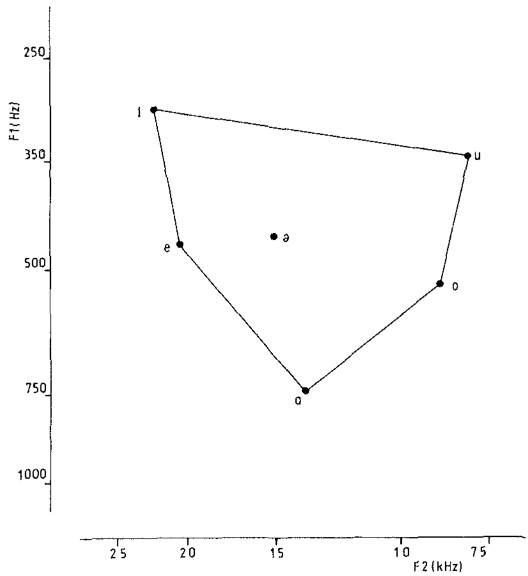

Figure 4c.
All speakers: 13 Indonesian speakers.

Figures 4a-c.
Mean F1 and F2 values of the 6 Indonesian monophthongs pronounced in isolation; 13 speakers, 5 tokens per speaker. The formant values are averaged first for the tokens per speaker and then for the speakers per dialect group.

Discussion. In Figures 4a and 4b the centres of gravity of the 3 groups of speakers are compared. The figures give rise to a number of observations of which the most important seem to be:

- The vowel space taken up by the Sundanese group of speakers is considerably larger than that of the Javanese and Toba Batak speakers; the Sundanese speakers used relatively closed realizations of /i/ and, especially, /u/ (low F1).

- The vowel diagram of the Toba Batak group of speakers is more "slender" than the other two, especially in the mid area.

Differences in vowel positions as between Javanese /i/ and Sundanese /i/ (F1: 15% and F2: 8%), or between Javanese /u/ and Sundanese /u/ (F1: 20% and F2: 13%) (Figure 4b) are clearly perceptible; according to Flanagan (1972:280) differences of 3 to 5% of the formant frequency of synthetic vowel sounds are already noticeable. Both differences can be understood if we take the vowel systems of the regional dialects of the 3 groups of speakers into consideration. Sundanese has 2 central vowels, and it seems likely that the Sundanese informants need a large central area, which causes them to have relatively closed realizations of /i/ and /u/, even when speaking Indonesian. Toba Batak, on the other hand, has no central vowel in its phoneme system. Apparently Toba Batak speakers make do with a small central area, not only in their regional dialect but also when speaking Indonesian.

- The central vowel realizations are in a mid position for the Javanese group of speakers, almost exactly half-way between /e/ and /o/, which, incidentally, showed considerable allophonic variation between realizations in the sentences with *tetes* and *bebe,* and with *totok* and *toto.* The central vowel as pronounced by the Sundanese informants is considerably more closed. We have no exact information on the position of the two central phonemes of the Sundanese language. Consequently, we cannot decide which of the two was chosen, or whether both or some area in between were used for the Indonesian central vowel realizations. The position of the central vowel as pronounced by the Toba Batak informants, indicated in Figures 3a and 4b, is misleading. The Toba Batak informants had difficulties in pronouncing the Indonesian central vowel in isolation. In actual fact two of the four Toba Batak speakers pronounced it as /e/. An analysis of their pronunciation of the central vowel in context is necessary before anything further can be said about Toba Batak pronunciation of the Indonesian central vowel phoneme.

Perception experiment

In the perceptual approach, each informant is presented with the same set of stimuli (cf. Schouten 1975:25). The informants have to indicate which speech sound they recognize in each of the stimuli. It is hoped that this way speaker-dependent variation in vowel quality can be avoided, or at least reduced (cf. Hombert 1979:27-32; Paliwal, Lindsay and Ainsworth 1983:77-83).

Method. On the basis of an acoustic pilot study (van Zanten and van Heuven 1983:70-80), a realistic vowel space was defined for Indonesian vowels spoken in isolation. Within the limits of this space, a set of 188 monophthongs was produced with a Fonema OVE IIIb speech synthesizer.[2] F1 and F2 of these vowel sounds were systematically varied in steps of 9%, i.e. 3 times the just noticeable difference commonly reported for F1 and F2 frequency changes for synthesized vowels (Flanagan 1972:280), sampling the acoustic vowel space in the way indicated in Figure 7. The frequency of F3 equalled that of F2 + 600 Hz, with a minimum of 2460 Hz. F4 and F5 were set at 3500 and 4000 Hz respectively for all vowels. Two tapes were prepared containing the set of 188

vowel	listeners	F_1 $\bar{x}$	F_1 s	F_2 $\bar{x}$	F_2 s
/ɪ/	Toba Batak	300	13	2347	72
	Javanese	295	22	2378	80
	Sundanese	313	27	2408	74
	Indonesian	302	23	2377	49
/e/	Toba Batak	541	29	1979	40
	Javanese	579	29	2005	77
	Sundanese	579	44	2064	107
	Indonesian	567	39	2015	86
/a/	Toba Batak	856	30	1441	16
	Javanese	868	21	1389	67
	Sundanese	848	14	1474	26
	Indonesian	858	24	1431	57
/o/	Toba Batak	499	8	1018	30
	Javanese	495	21	939	13
	Sundanese	534	15	1010	62
	Indonesian	508	24	985	54
/u/	Toba Batak	313	7	999	55
	Javanese	299	5	1005	51
	Sundanese	314	12	960	38
	Indonesian	308	11	989	53
/ə/	Toba Batak	494	13	1569	70
	Javanese	467	33	1464	105
	Sundanese	497	17	1547	49
	Indonesian	484	27	1522	93

Table II

Mean formant frequencies ($\bar{x}$) and standard deviations (s) in Hz of the 6 Indonesian monophthongs as labelled by the informants, 13 listeners (4 Toba Batak, 5 Javanese, 4 Sundanese), 188 synthesized vowel stimuli, 2 identifications per stimulus per listener Responses with the qualification 'identical or very similar to the vowel chosen' were counted twice, 'fairly acceptable' responses were counted once and 'unacceptable' responses were disregarded The formant values are averaged first for the responses per listener and then for the listeners per dialect group and for all listeners The standard deviation s expresses the scatter of the means over the listeners stimuli, preceded by a series of practice items, in counterbalanced random orders.

The same informants who took part in the production experiment participated in the perception experiment. They were instructed to label each vowel stimulus immediately after hearing it as one of the six monophthongs of Indonesian (forced choice) and to rate the stimulus along an acceptability scale, as follows: '2' (identical or very similar to the vowel chosen), '1' (fairly acceptable) or '0' (unacceptable, but there is no other vowel closer to this stimulus).

Results. The responses were weighted according to the following procedure. Each response with the qualification 'identical or very similar to the vowel chosen' was counted twice; 'fairly acceptable' responses were counted once, and '0' 's were disregarded. For each informant weighted means were calculated. Of these weighted means a normal mean and standard deviation over the groups of informants and over all the individual informants were calculated for each vowel. We thus obtained the means and standard deviations as listed in Table II and depicted in Figures 5a, 5b, 5c and 5d. In the figures, the centres of the crosses indicate the vowel means and the bars of the crosses represent the standard deviations.

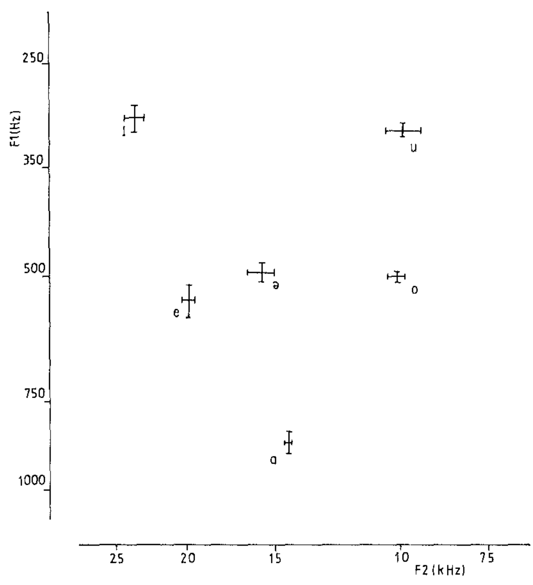

Figure 5a
Toba Batak (4 listeners)

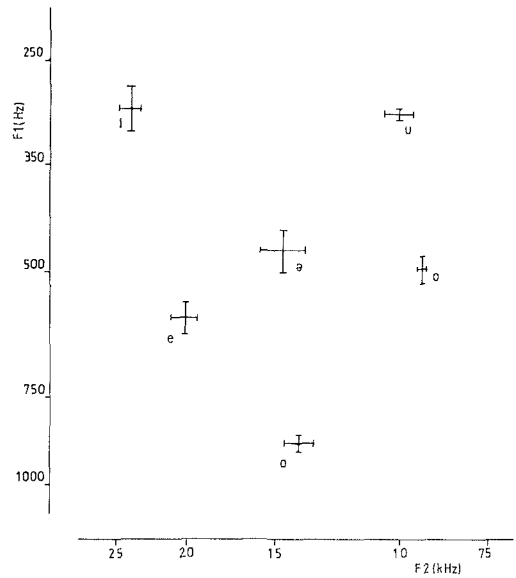

Figure 5b
Javanese (5 listeners)

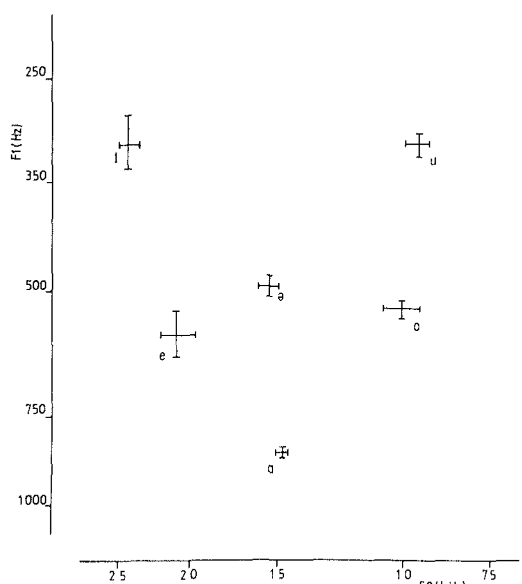

Figure 5c
Sundanese (4 listeners)

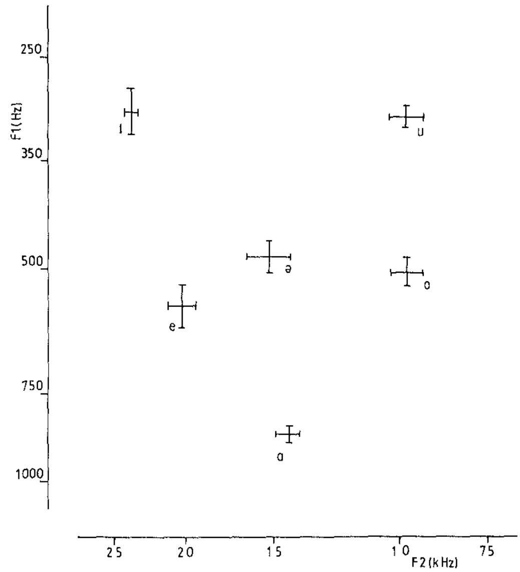

Figure 5d.
All listeners: 13 Indonesian listeners.

Figures 5a-d.
Mean F1 and F2 values (centres of crosses) and standard deviations (bars of crosses) of the 6 Indonesian monophthongs as labelled by Indonesian informants; 13 listeners, 188 synthesized vowel stimuli, 2 identifications per stimulus per listener. Responses with the qualification 'identical or very similar to the vowel chosen' were counted twice, 'fairly acceptable' responses were counted once and 'unacceptable' responses were disregarded. The formant values are averaged first for the responses per listener and then for the listeners per dialect group and for all listeners. The standard deviations express the scatter of the means over the listeners.

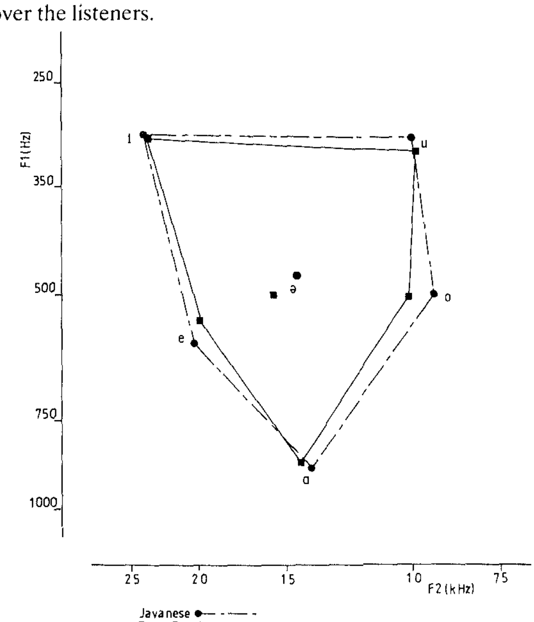

Figure 6a.
4 Toba Batak and 5 Javanese listeners.

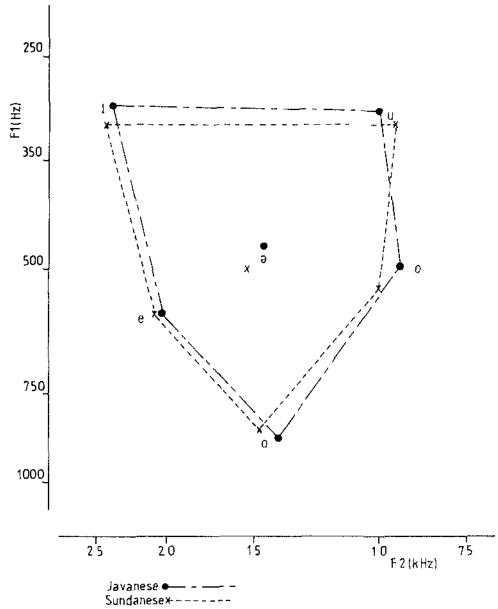

Figure 6b.
5 Javanese and 4 Sundanese listeners.

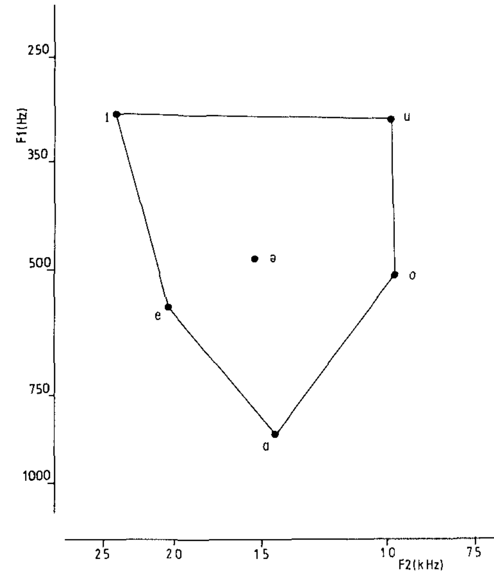

Figure 6c.
All listeners: 13 Indonesian listeners.

Figures 6a-c.
Mean F1 and F2 values of the 6 Indonesian monophthongs as labelled by Indonesian informants; 188 synthesized vowel stimuli, 2 identifications per stimulus per listener. Responses with the qualification 'identical or very similar to the vowel chosen' were counted twice, 'fairly acceptable' responses were counted once and 'unacceptable' responses were disregarded. The formant values are averaged first for the responses per listener and then for the listeners per dialect group and for all listeners.

Discussion. On the basis of the data of Table II we can make the following observations (cf. Figures 6a and 6b):

- The points of gravity of /e/ and /o/ of the Toba Batak listeners are considerably more centralized than the corresponding points of the other two groups. The difference between the F1's of the Toba Batak and Javanese /e/, for instance, is 7%, while the F2's of Toba Batak /o/ and Javanese /o/ differ 9%. As was indicated above, such differences in vowel quality are well audible. The point of gravity of the central vowel appears to be more back (7%) and more closed (6%) for the Javanese listeners than for the other two groups.
  - The vowels of the listeners of a Sundanese background are all, except /u/, more fronted than the corresponding points for the Javanese listeners. /i/ and /u/ are not more closed than in the case of the other two groups, but they are farther separated.

The 'slender' vowel diagram of the Toba Bataks in the perception experiment may, again, be explained by the fact that Toba Batak has no central vowel. The central area is apparently at least partly perceived as /e/ or /o/. The points of gravity of Sundanese /i/ and /u/ are not more closed than those of Toba Batak and Javanese /i/ and /u/. Remarkable, however, is the realization of /u/, which is farther back than Javanese and Toba Batak /u/, whereas all the other Sundanese vowels are more front than the Javanese and (apart from /o/) Toba Batak vowels. The relatively large distance between Sundanese /i/ and /u/ can be explained by the large central area in the Sundanese language, which has to accommodate two central vowels.

We can view the perception data in a different, and possibly more illuminating way by taking into consideration the areas of preference for the response vowels. Figures 7a, 7b, and 7c plot the areas of preference for the three listener groups, i.e. areas containing only those stimuli that were identified as one particular vowel in at least 50% of the responses (small letters) and in at least 75% (large letters). Summary statistics are given in Table III, specifying the number of stimulus points (absolute and relative) contained by each area of preference.

vowel	50 % agreement			75 % agreement		
	Toba Batak	Javanese	Sundanese	Toba Batak	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%)
unlabell.	89 (47%)	76 (40%)	83 (44%)	139 (74%)	135 (72%)	135 (72%)

Table III
Agreement on vowel identification, 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 informants (4 Toba Batak, 5 Javanese, 4 Sundanese)

SECOND FORMANT (Hz)

FIRST FORMANT (Hz)	2614	2397	2198	2016	1849	1695	1554	1425	1307	1199	1099	1008	924	848	777	713
252.	i	i	.	.									u	u	u	u
275.	i	i	.	.	.	.	.	.	.	.	u	U	U	U	U	U
300.	i	i	i	.	.	.	.	.	.	u	u	u	U	U	U	U
327.	i	i	.	.	.	.	.	.	.	.	.	U	U	u	U	U
356.		.	.	.	.	.	.	.	.	.	.	u	.	U	u	
389.		e	e	.	.	e	.	e	.	.	.	.	o	.	.	
424.		e	e	e	.	e	.	ə	.	.	.	.	.	O	O	
462.		e	e	e	.	.	e	e	e	e	.	o	O	o	o	
504.		e	e	e	.	.	.	ə	e	.	.	o	O	o		
550.			e	e	.	e	e	e	.	.	.	O	o	o		
599.			e	e	e	.	.	.	.	.	.	o				
654.			e	e	e	.	.	.	.	.	.	.				
713.				e	e	.	.	.	.	.	.	o				
777.				.	.	.	.	a	ɑ	.	.					
848.					e	ɑ	a	a	ɑ	ɑ						
924.						ɑ	ɑ	a	ɑ							
1008.						ɑ	ɑ	ɑ	ɑ							

Figure 7a.
Toba Batak (4 listeners).

SECOND FORMANT (Hz)

FIRST FORMANT (Hz)

	2614	2397	2198	2016	1849	1695	1554	1425	1307	1199	1099	1008	924	848	777	713
252.	i	i	.	.									u	u	u	u
275.	i	i	i	.	.	.	.	.	.	.	.	u	u	u	u	u
300.	i	i	.	.	.	.	.	.	.	.	u	u	u	u	u	u
327.	i	.	.	.	.	.	.	.	.	.	u	u	u	u	u	u
356.		.	.	.	.	.	.	.	.	.	.	u	u	u	u	
389.		e	.	.	.	.	.	.	.	.	.	.	o	o	o	
424.		e	.	.	.	ə	ə	ə	ə	.	.	o	o	o	o	
462.		e	e	.	ə	ə	ə	ə	ə	.	.	o	o	o	o	
504.		e	e	e	.	ə	ə	ə	.	.	.	o	o	o		
550.			e	e	.	ə	ə	ə	ə	.	.	o	o	o		
599.			e	e	ə	ə	ə	ə	ə	.	.	o				
654.			e	e	e	.	.	.	.	.	.	o				
713.				e	e	.	.	.	.	.	.	o				
777.				e	e	a	a	ɑ	ɑ	a	a					
848.					e	a	a	ɑ	ɑ	ɑ						
924.						ɑ	ɑ	ɑ	ɑ							
1008.						ɑ	ɑ	ɑ	ɑ							

Figure 7b.
Javanese (5 listeners)

SECOND FORMANT (Hz)

FIRST FORMANT (Hz)	.2614	.2397	.2198	.2016	.1849	.1695	.1554	.1425	.1307	.1199	.1099	.1008	.924	.848	.777	.713
252.	I	I	i	.									u	u	.	.
275.	I	I	i	.	.	.	.	.	.	.	.	.	u	u	u	.
300.	I	I	i	.	.	.	.	.	.	.	.	.	u	U	u	u
327.	I	i	i	.	.	.	.	.	.	.	.	u	u	u	U	u
356.		.	.	.	.	.	.	.	.	.	.	u	u	u	.	
389.		.	.	.	.	e	.	e	e	e	.	.	.	o	.	
424.		e	e	.	.	Ə	Ə	e	e	.	.	.	o	o	o	
462.		e	E	e	.	Ə	Ə	Ə	Ə	.	e	o	O	O	O	
504.		E	E	e	.	Ə	Ə	Ə	Ə	.	.	O	O	O		
550.			e	E	e	e	Ə	Ə	e	o	o	O	O	O		
599.			E	E	E	.	e	e	.	.	.	O				
654.			e	E	E	.	.	.	.	.	.	o				
713.				E	E	.	.	.	.	.	.	.				
777.				e	.	.	.	a	A	.	.					
848.					.	a	a	A	A	A						
924.						a	A	A	A							
1008.						A	A	A	A							

Figure 7c.
Sundanese (4 listeners).

Figures 7a-c.
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).

There are several conspicuous differences across the three listener groups in the locations and sizes of the preferred vowel areas, and 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, intermediate for the Javanese, and largest for the Sundanese. Conversely, the area associated with /u/ is large for the Toba Batak, intermediate for the Javanese, and smallest for the Sundanese.

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 category for the Toba Batak. 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 of a Javanese background, a /ə/ dialect, the preferred area for /ə/is appreciably larger, and the responses are far more normally distributed. For the Sundanese group, having a background dialect with two central vowels, the preferred /ə/ area is larger still, and, perhaps more importantly, the distribution of especially /u/ here 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% (over against 14 and 17 for the Toba 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 of the Sundanese.

Comparison of production and perception data

We are not aware of the existence of other published vowel quality studies in which production and perception data are presented which were collected from the same individuals. More generally, very little is known about the precise relationship between the production and perception of vowel quality in phonetic theory. A preliminary inspection of our own production and perception data reveals considerable disparity in the configurations of vowel means between the two modes (cf. our Figures 3 and 4 versus 5, 6, and 7). Therefore it seems hazardous to compare production and perception vowels with each other in terms of their absolute frequencies.

Nevertheless, there are a number of interesting parallels between the results of the two experiments. In both experiments the vowel diagram of the Toba Batak is 'slimmer' than the other two, especially in the mid area. This can be related to the fact that Toba Batak speakers have no central vowel in their substrate dialect. The large central area which the Sundanese need to accommodate their two central vowels apparently causes them to produce fairly extreme /i/ and /u/ realizations. In the production experiment we can observe this extreme character in the F1 as well as in the F2 dimension, but in the perception experiment only in the F2 dimension (cf. also Figure 7c).

Conclusions

The Indonesian Vowel System

According to our data (cf. Figures 4c and 6c), Indonesian /e/ and /o/ can be classified as mid, not as higher-mid as is done by Crothers (cf. Crothers 1978:141; see also Figure 1). The central vowel as perceived and produced in isolation by our informants is mid or higher-mid and certainly not high. It is possible, however, that it is more closed when pronounced in context (cf. van Zanten and van Heuven 1983:74). Here /ə/ uttered in context was practically in line with the closed vowels /i/ and /u/, whereas produced in isolation it seemed fairly centralized.

The Toba Batak, Javanese and Sundanese Vowel Systems

Assuming that the production and perception of the Indonesian vowels by our three groups of informants tells us something about the vowels of their substrate languages, we suggest that – contrary to the information given by Crothers, see Figure 1 – Toba Batak /e/ and /o/ are both mid or higher-mid, and that Sundanese as well as Javanese /e/ and /o/ might be described as mid. In the case of the Javanese there was marked allophonic variation for these two phonemes.

Vowel Typology

Crothers (1978) does not give us any indication of the formant values associated with his categories 'high', 'low', etc. Thus it is not clear whether Toba Batak, Javanese and Sundanese /i/, for instance, should all be labelled 'high front', or whether a distinction should be drawn between 'high front' and 'lower-high front'. It would seem that to compare the vowel systems of different languages we need well-defined categories.

If our analysis holds good, there is a relation between the number of (central) vowels and the size of the formant space in a given system: Sundanese, a seven-vowel language, has a larger vowel diagram than Toba Batak, which has five vowels. This is contrary to the Liljencrants and Lindblom model, which assumes that all languages have an identical vowel space (Liljencrants and Lindblom 1972:839-862). Crothers' solution to this problem is to define vowels as equal circular areas rather than as points (the more vowels in a system, the smaller each individual vowel diameter), and to 'shrink' the free vowel space around these until the smallest possible vowel space is achieved. It would seem that here also phonetic research into the exact positions and shapes of vowel areas of the various systems could provide a firmer basis for the discussion. Here our perceptual method of charting a vowel system may prove very useful under conditions where sophisticated laboratory equipment is not available for on the spot analysis.

One final point may be of interest as regards the Javanese speakers. From the results of the perception experiment it appears that the Javanese informants reached higher agreement in identifying vowels than the other groups of informants (cf. Table III).

The Javanese informants were also more consistent in their labelling. The total number of times that a stimulus was classified as the same vowel phoneme by the individual listeners on both presentations (disregarding weight factors) was computed per group of listeners. The resultant figure was divided by the total number of times the vowel phoneme was chosen as such by the group of speakers in order to give some measure of consistency. For example, a total of 190 stimuli was identified as /a/ by the Toba Batak informants. 74 of these stimuli were labelled /a/ on both presentations. The consistency measure for Toba Batak /a/ then is (2 × 74)/190 = 0.78 (see Table IV). The consistency is considerably higher for the Javanese group than for the other two groups of informants. These data seem to demonstrate that the Javanese, as speakers of a dialect that is similar to the standard language with respect to its vowel system, are at an advantage here. Predictably, the consistency of Toba Batak /ə/ is low: the Toba Batak informants did not have a stable mental picture of Indonesian /ə/; cf. also the irregular area of dispersion for /ə/ in Figure 7a.

	/a/	/e/	/i/	/o/	/u/	/ə/	/x̄/
Toba Batak	0.78	0.66	0.71	0.77	0.84	0.53	0.71
Javanese	0.87	0.79	0.89	0.82	0.86	0.72	0.83
Sundanese	0.83	0.68	0.74	0.73	0.78	0.72	0.75

Table IV.
Consistency in vowel identification; number of stimulus points that were classified by individuals within a group as the same vowel on both presentations, as a fraction of the total number of times that this vowel was chosen by that group of listeners (4 Toba Batak, 5 Javanese, 4 Sundanese).
Example: Toba Batak: classif. as /a/ : 190 stimuli
classif. as /a/ on both present.: 74 stimuli
fraction: (2 × 74)/190 = 0.78.

APPENDIX

The stimulus material of the production experiment

The 6 monophthongs /i, e, a, o, u, ə/ were presented to the listeners in the carrier sentence *Dalam kata . . . (titik, tetes, . . .) terdapat bunyi . . . (i, e, . . .)*, 'In the word . . . (*titik, tetes*, . . .) the sound (i, e, . . .) is found'.

The vowels uttered in isolation were analysed in sentences with the following stimulus words:

titik	/titik/	'dot', 'drop'
tetes	/tetes/	'drop'
bebe	/bebe/	'dress'
tatap	/tatap/	'to peer, observe intently'
totok	/totok/	'full-blooded; newcomer'
toto	/toto/	'sweepstakes'
tutup	/tutup/	'closed'
tetep	/tətəp/	'fixed, definite' (Jakartan for Standard Indonesian *tetap* /tətap/

NOTES

1 The LPC analysis was carried out at the Institute for Perception Research (IPO) at Eindhoven We are grateful to Dr Ir L L M Vogten for his assistance
2 We wish to acknowledge the cooperation of the Dept of Phonetics at the University of Utrecht in putting this equipment at our disposal

REFERENCES CITED

Crothers, J
1978 'Typology and universals of vowel systems', in J H Greenberg (ed), *Universals of human language, Vol 2 Phonology,* pp 93-152, Stanford, California Stanford University Press
Flanagan, J L
1972 *Speech Analysis, Synthesis and Perception,* Chapter 7, Berlin-Heidelberg-New York Springer-Verlag
Halim, Amran
1974 *Intonation in relation to Syntax in Bahasa Indonesia,* Jakarta Djambatan
Hombert, J -M
1979 'Universals of vowel systems The case of centralized vowels', in E Fischer-Jørgensen, J Rischel and N Thorsen (eds), *Proceedings of the Ninth International Congress of Phonetic Sciences, Vol 2,* pp 27-32, Kopenhagen
Liljencrants, J , and B Lindblom
1972 'Numerical simulation of vowel quality systems The role of perceptual contrast', *Language* 48-4 839-862
Paliwal, K K , D Lindsay and W A Ainsworth
1983 'Correlation between production and perception of English vowels', *Journal of Phonetics* 11 77-83
Ras, J J
1982 *Inleiding tot het Modern Javaans,* Chapter 1 Den Haag M Nijhoff
Schouten, M E H
1975 *Native-language interference in the perception of second-language vowels,* Ph D thesis, Utrecht University
Stokhof, W A L
1975 'On the phonology of Bahasa Indonesia', *Bijdragen tot de Taal-, Land- en Volkenkunde* 131-2/3 254-269
Teeuw, A
1978 *Leerboek Bahasa Indonesia,* Chapter 1, Groningen Wolters Noordhoff
Trubetzkoy, N
1929 'Zur allgemeinen Theorie der Phonologischen Vokalsysteme' *Travaux du Cercle Linguistique de Prague, 1 Melanges Linguistiques dedies au Premier Congres des Philologues Slaves* 39-67
Van Zanten, E , and V J van Heuven
1983 'A phonetic analysis of the Indonesian vowel system A preliminary acoustic study', *NUSA, Linguistic Studies of Indonesian and Other Languages in Indonesia,* 15 70-80