Effects of substrate language on the localization and perceptual evaluation of pitch movements in Indonesian

Heuven, V.J.J.P. van; Zanten-Wervelman, E.A. van; Odé C., Stokhof W.

EFFECTS OF SUBSTRATE LANGUAGE ON THE LOCALIZATION AND PERCEPTUAL EVALUATION OF PITCH MOVEMENTS IN INDONESIAN

Vincent J. van Heuven & Ellen van Zanten

1 Introduction

The phonetics and phonology of the Indonesian stress system have recently attracted a good deal of attention in the literature (cf. Cohn 1989; Odé & van Heuven 1994, and references cited there). In spite of numerous claims to the contrary (cf. Odé 1994 for a literature survey), we would maintain the position that there is possibly only one restriction on stress placement in Indonesian: there is a fairly strong tendency for stress to appear on the prefinal syllable if it is heavy, i.e. closed by at least one consonant. In all other cases tokens of the same words may be heard with stress in any position (including schwa). Since closed syllables are a-typical in Indonesian, the exceptional situation does not arise very often, so that, in the final analysis, word stress in Indonesian is free.

As a logical consequence of this, Indonesian stress serves no contrastive functions. There is not a single minimal stress pair in Indonesian, i.e. there are no words that contain exactly the same sequence of vowels and consonants, but crucially differ in their stress patterns, as do exist in languages such as English (e.g. *'forbear* 'ancestor' ~ *for'bear* 'endure', *'trusty* 'trustworthy' ~ *trus'tee* 'board member', cf. Cutler 1986) or Dutch (e.g. *'kanon* 'canon' ~ *ka'non* 'cannon', cf. van Heuven 1994).

Whether or not Indonesian has free stress (which is tantamount to having no stress at all, van Heuven 1994), it certainly does have accent. It is possible, just as in languages such as Dutch and English, to highlight (or focus on) one particular word over another in a sentence, indicating its communicative importance for the listener, by executing an accent-lending pitch movement on it (cf. van Zanten 1994). However, in contrast to the European languages, the position of this accent-lending pitch movement is not tied to the stressed syllable, but may occur anywhere within the word.

As a result, we would predict that native listeners of Indonesian would be relatively insensitive to variations in the timing of an accent-lending pitch movement as long as the movement remains within the bounds of a word domain. When, however, an accent-lending pitch movement would be displaced in time so that it would fall within the bounds of an adjacent word, the effect should be relatively more noticeable.

It is important to realize that Indonesian is very much a lingua franca that is spoken on a variety of substrate languages. This means that the pronunciation of Indonesian differs considerably depending on the origins of the speakers. As a case in point we have shown that the acoustical realization of, and perceptual norms for, the six standard Indonesian vowels differed systematically and predictably for speakers with a Javanese (with 6 vowels), Sundanese (with 7 vowels) and Toba Batak (with 5 vowels) substrate language (van Zanten & van Heuven 1983, 1984; van Zanten 1989a,b). With respect to the problem at hand, i.e. stress position and stress perception, speakers of Indonesian with a Toba Batak substrate find themselves in an exceptional position in that Toba Batak, in contrast to other Indonesian languages, has deterministic stress rules, and marks stressed syllables with clear phonetic means. Nababan (1981) shows that stress in Toba Batak functions contrastively, both morphologically (e.g. *'tibbo* 'height' ~ *tib'bo* 'high') and lexically (e.g. *'itom* 'black dye' ~ *i'tom* 'your sibling'). In van Zanten & van Heuven (1997) it was found that stressed syllables in Indonesian with a Toba substrate were some 40% longer than their unstressed counterparts, whilst the mean lengthening effect of stress was smaller (25%) for the Sundanese and Javanese speakers.

In the present study, therefore, we predict that Indonesian speakers with a Toba substrate will be as sensitive as other speakers of Indonesian to time shifts in the position of an accent-lending pitch movement when the movement is displaced into another word. However, when a pitch movement is shifted along the time axis by the same time interval without crossing a word boundary, Toba listeners will be more sensitive to the change (since they will perceive the shift as a change of stress pattern) than other listeners of Indonesian (for whom the shift is functionally inconsequential).

In the research reported in the following sections, we aimed to experimentally verify the predictions voiced above. Should we indeed obtain the results as predicted, our admittedly controversial analysis of Indonesian as a predominantly free-stress language would be supported on new grounds.

2 *Methods*

Stimuli. It is obviously impossible for a human speaker to produce two (let alone multiple) versions of a sentence that are exactly the same in all respects but for a single difference of, say 100 ms, in the temporal alignment of a pitch movement. When one is interested in studying the perceptual effects of changes in the position of an accent-lending pitch movement, one therefore has to resort to synthesized speech, or at least to resynthesized human speech. In the present experiment we constructed a single Indonesian utterance by stringing together three Indonesian words, and manipulating in it the position of a single pitch movement. The basic sentence was generated by concatenating the three words *kemarin* 'yesterday', *Indonesia* 'id.', and *bagus* 'beautiful'. Each of these three words had been spoken with a single, narrow-focus accent (cf. van Heuven 1994) in three recordings of the sentence *Kemarin Indonesia bagus* 'Yesterday Indonesia (was) beautiful' by a female speaker of standard Indonesian. As a result we may assume that each of the three target words bore the same accentuation cues, so that perceiving an accent on one word over another after concatenation and pitch elimination could not be influenced by any remaining cues in the original recordings. Also, this procedure ensures the absence of prosodic boundary cues after the first two words. The three target words were analog-to-digital converted (10 KHz, 12 bit, 4.5 KHz LP) and submitted to automatic parameter extraction using the Split-Levinson robust formant tracking algorithm (Willems 1987; 25.6 ms window, 10 ms window shift, lowest four formants and associated bandwidths F_1 through F_4, and B_1 through B_4). Fundamental frequency (the acoustic determinant of pitch) was determined by the method of subharmonic summation (Hermes 1988) followed by an automatic tracking procedure. The three target words were excised from their original context and concatenated. The mean pitch declination in the three original utterances was estimated and superposed onto the concatenated utterance.

Also, the excursion size and steepness of the rise and fall pitch movement on the three target words in their original contexts was measured (in semitones and semitones per second, respectively) and averaged. Nineteen versions of the concatenated utterance were then produced by shifting the position of a single rise-fall configuration (conforming to the specification of the average size and steepness determined above, viz. a 3.5 st rise during 160 ms followed by a 7.6 st fall during 220 ms) along the time axis in steps of 100 ms. These specifications conform to the 231 contour mentioned by Halim (1981) and correspond closely to Ebing & van Heuven's (1997) standardized rise 2 (3 st during 140 ms) and fall A (7 st during 180 ms). It would seem to be the prototypical realization of an accent in Indonesian. In the first version the rise started exactly at the onset of the utterance, and in the nineteenth version the termination of fall coincided with the offset of the sentence-final [s]. Figure 1 specifies the relevant parameters for the nineteen versions: in the top panel the intensity of the utterance is shown (determined on the basis of a constant 100 Hz pitch), the middle panel shows the waveform plus segmental transcription; the bottom panel shows the nineteen rise-fall pitch configurations used in the experiment. Notice that the rise of the first accent is virtual since it coincides with the silent/voiceless /k/; likewise the fall of the nineteenth accent is largely absent since it overlaps with the voiceless final /s/. Approximate syllable boundaries have been indicated by vertical segmentation lines.

Note that among the nineteen versions there were three in which the segmental alignment of the pitch configuration was (virtually) identical to that in the original recordings of the three target words.[1]

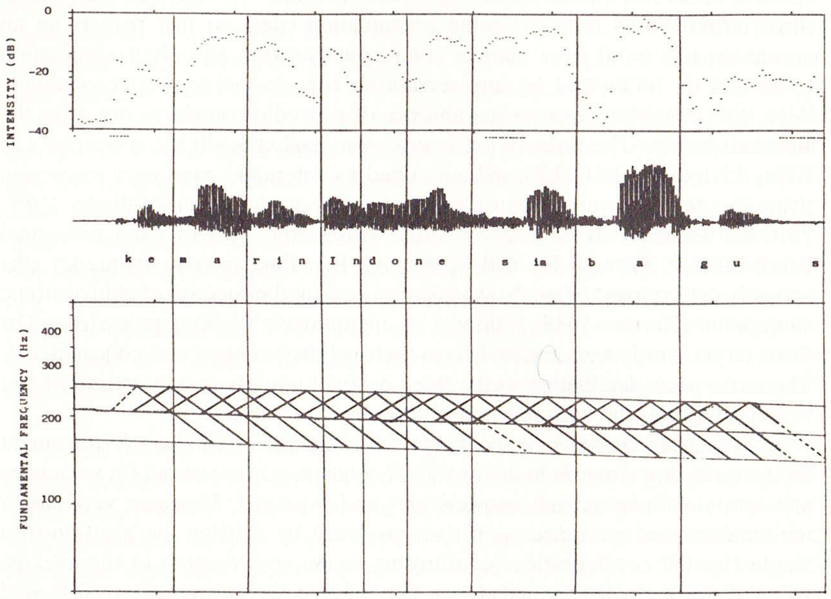

Figure 1: Overview of stimuli used in experiment. Intensity (in decibels) in the top panel; each dot represents a time interval of 10 ms; dots below the -40 dB baseline indicate voicelessness. The middle panel shows the waveform of the basic stimulus. The bottom panel shows the positions and shape of the nineteen rise-fall configurations. Each configuration starts at the middle declination line and each fall ends at the bottom declination line. The time interval between two adjacent tick marks along the horizontal axis equals 227 ms.

Any remaining discrepancies between the synchronization of the synthesized pitch rises and the originals were certainly below threshold: even highly

[1] No provisions could be made to align accents so as to optimally coincide with unstressed syllables as well. Since we have no clear timing parameters for the rise-fall configuration, the timing would have to be copied from human originals; however, it is virtually impossible to get speakers of Indonesian to shift the accent to different syllables in a word in a controlled fashion, e.g. through manipulation of narrow focus on the syllable level (cf. van Heuven 1994b; Sluijter & van Heuven 1995). Therefore we resorted to fixed time shifts of 100 ms.

experienced listeners need at least a 50-ms displacement of a pitch rise along the time axis before an audible difference can be detected (Caspers 1994).[2]

The nineteen versions were resynthesized, converted back to sound, and recorded on audio cassette in two counterbalanced random orders with 5s intervals between versions, yielding a single tape with 38 exemplars, preceded by 10 practice items.

Listeners. Two groups of listeners participated in the experiment. The first was a group of 23 speakers of Indonesian with a free-stress substrate language. The second group comprised 14 Indonesian speakers with a Toba Batak (distinctive stress) substrate. Subjects were students and/or staff at the Universitas Indonesia in Jakarta.[3]

Procedure. The tape was played to the listeners in small groups over good quality earphones in a quiet lecture room. Subjects were instructed (in Indonesian) to indicate on their answer sheets for each exemplar on the tape on which of the nine syllables they perceived the greatest prominence, i.e. the sentence accent. It was made clear to them that they had to make a single choice in all cases; on no account were they allowed to leave an item blank, or assign tied ranks. On a second run of the tape, after the issuing of new answer sheets, subjects had to indicate for each exemplar on the tape the acceptability of its melody along an 11-point scale ranging from 0 'highly unacceptable' to 10 'highly acceptable'.

3 Results

We will first analyse the listeners' responses in terms of the accuracy with which the position of the pitch rise could be localized in the utterance. In a later section we will evaluate the acceptability judgments obtained for the nineteen versions of our stimulus sentence.

[2] Very little is known about the resolution of the human hearing system for synchronization phenomena. Recent research (Verhoeven 1994) shows that Dutch listeners (distinctive stress) need an alignment shift in excess of 60 ms in order to hear that a single pitch rise differs in position, and a shift of more than 100 ms when the movement is a single fall. Caspers's (1994) results indicate a somewhat better resolution when the listeners were experienced intonologists.

[3] More recently, the tape was also played to a small group of native speakers of Dutch who were highly proficient in Indonesian (professors and postgraduate students at the Centre for Non-Western Studies at Leiden University). The results obtained with these listeners will be presented and discussed briefly in appendix I.

3.1 *Localization performance*

Nominally 1406 responses were collected (2 presentations of 19 stimulus types to 37 listeners). Subjects had left an item on their answer sheet blank in 28 cases, so that the actual number of responses was 1378. The raw data are presented in table 1. For each physical position of the stimulus pitch rise (listed vertically) the distribution of responses over the 9 available syllable positions is given in percentages, for the two groups of listeners separately. The row marginals (100%) are nominally 28 cases for the Toba Batak group (distinctive stress substrate) and 46 for the other listeners (free stress substrates). For each stimulus the single syllable that attracted the largest number of reported accent positions is printed in bold face.

Table 1: Judged position of syllable bearing pitch accent for 19 different pitch accent positions, broken down by listener group. Nominal number of responses per row is 46 for the free stress listener group, and 28 for the distinctive stress group.

	FREE STRESS LISTENERS									DISTINCTIVE STRESS LISTENERS								
	judged syllable position bearing F_0-rise									*judged syllable position bearing F_0-rise*								
position F_0-rise	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
1	**58**	16	7	9	2	2	4		2	**57**	29	4	7				4	
2	**51**	33	7	2			2	4		**57**	21	11	4				7	
3	33	**42**	13		2			4	4	29	**61**	7			4			
4	18	**60**	20			2				7	**82**	7		4				
5	13	**50**	30	4	2					4	**61**	36						
6	21	9	**33**	12	5	7	2	9	2	7		**59**	7		7	7	11	
7	7	9	9	**22**	11	11	4	20	7	4		**32**	25	14	18	4	4	
8	7	2	4	**36**	16	16	2	16	2		4	4	**50**	21	11	4	7	
9	7	2	2	7	11	**41**	9	14	7		7		11	32	**39**	7	4	
10	9	7	2	7	2	**50**	14	5	5				4	11	**71**	14		
11	7	2	2	9	9	**39**	16	16		4			4		**79**	11	4	
12	5			5	7	30	**30**	23	2	4		4	7		**50**	18	18	
13	2		2	5		5	**77**	9					4		14	**75**	7	
14	9		2	2		7	**40**	36	4				4		14	**43**	36	4
15	7	2	4	2		2	11	**67**	4			4				18	**79**	
16	4			4		2	4	**83**	2				4				**96**	
17	4			2			4	**76**	13							4	**96**	
18	5	2		7				**70**	16							4	**68**	29
19	2	5				2		**55**	36				7				43	**50**

Impressionistically speaking, the two groups of listeners concur to a large extent in their choice of the syllable that bears the pitch accent. However, the choices made by the Toba group generally seem to be more outspoken, and the distributions of the minority responses (indicated in normal type face) are more tightly clustered around the (bold-faced) preferred response categories.

In order to be able to express these facts in more quantitative measures we analysed two statistics for these data:

1. The mean reported accent position expresses the average syllable position judged to bear the accent-lending pitch configuration. This value should approach the syllable position that shows up as the majority response (bold-faced) in table 1. The difference between the mean reported accent position and the actual location of the pitch configuration, of course, serves as a measure of how far off the listeners' perception is from the physical reality.
2. The standard deviation of the judged position of the pitch accent is taken as a measure of the uncertainty in the listeners' responses. The tighter the clustering of the listeners' responses around the mean judged position of the accent, the smaller the standard deviation.

Mean reported accent position. Figure 2 presents the mean reported linguistic position of the pitch accent (ranging between syllable 1 and syllable 9) as a function of the physical location of the configuration along the time axis (19 steps of 100 ms along the physical time dimension). The results have been broken down by the linguistic background of the listeners, i.e. Toba Batak versus other Indonesian listeners.

Figure 2 shows that, overall, both groups of listeners localized the pitch accent more or less in the position where it physically occurred. A classical analysis of variance with accent location (19 levels) and listener substrate (2 levels) as fixed factors with repeated measures over the two presentations of each stimulus shows that the overall effect of position accounts for 98% of the variance, $F(18,57)=164.8$ ($p<<.001$). There is virtually no difference between Toba and other Indonesian listeners in the mean syllable position they attribute to the pitch accents (mean position is 0.08 syllable later for the Toba group); the difference is completely insignificant, $F(11,74)<1$. There is, however, a small interaction between the physical position of the accent in the sentence and its judged position, that just fails statistical significance, $F(18,38)=1.7$ ($p=.091$). The interaction is caused by the tendency for Toba listeners to perceive the pitch accent in a later position (0.35 syllables on average) than the other Indonesians in the second half of the sentence, viz. for accent positions 10 and up. This difference is significant by paired t-test, $t(19)=3.6$ ($p=.002$, two-tailed). The reverse effect is found in the first half of the sentence (rise positions 1 through 9): here the Toba listeners tend to hear the pitch accent in an earlier position (0.23 syllable) than the other Indonesians do, $t(17)=-1.9$ ($p=.067$, two-tailed). The difference in accent localization between Toba and other listeners in first and second half of the sentence is highly significant by a grouped t-test, $t(36)=3.9$ ($p<.001$).

These results mean that, by and large, both the Toba listeners and the other Indonesians, when each group is considered as a collective measurement instrument, are roughly equally capable of localizing the syllable that bears a pitch accent. However, there is a slight polarization tendency for the Toba listeners: they tend to hear accents at earlier positions than other Indonesians do in the first half of the sentence and at later positions in the second half. This indicates that Toba listeners tend to perceive the pitch configurations a little closer to their actual physical location. In this perspective the data bear out that, as predicted, the perception of accent-lending pitch configurations is more accurate than of other Indonesian listeners.

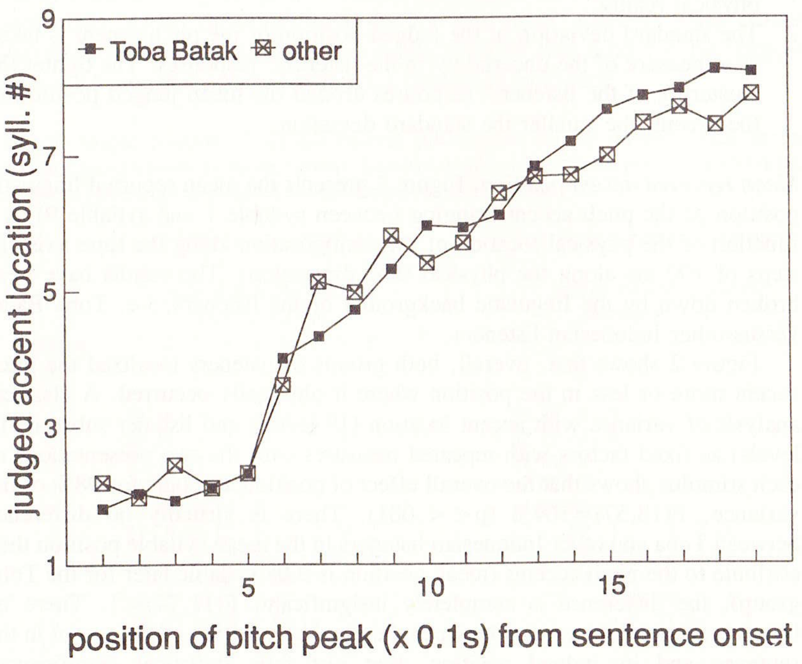

Figure 2: Perceived accent location (ordinal number of syllable) as a function of physical position of pitch accent (F_0-peak in 0.1s steps from sentence onset) broken down by distinctive stress (Toba Batak) and free stress (other) listeners.

Standard deviation of reported accent position. In our informal discussion of table 1 we noted that there is more to our data than just a small difference in average group behaviour. It seemed to be the case that the listener groups differ considerably within themselves in their ability to determine the location of a pitch rise. Let us therefore examine the spread (standard deviation) in the responses around the mean position judgments for each accent position,

separately for Toba and other listeners. This breakdown is presented in figure 3.

This time the results differentiate quite clearly between the two listener groups. The Toba listeners, on average, show much less variation in their judgments of the position of the syllable bearing the pitch configuration than the other listeners do. For each of the 19 stimuli the standard deviation of the judged accent location is smaller for the Toba listeners (1.0 syllables on average) than for the other Indonesians (1.8 syllables on average).

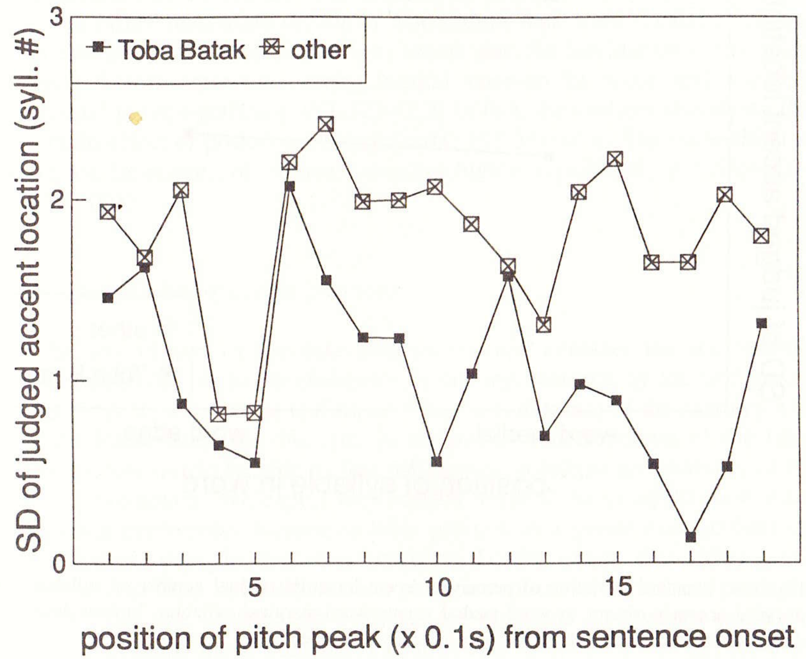

Figure 3: Standard deviation of perceived accent location (ordinal number of syllable) as a function of physical position of pitch accent (F_0-peak in 0.1s steps from sentence onset) broken down by distinctive stress (Toba Batak) and free stress (other) listeners.

The analysis of variance this time shows both an effect of the physical position of the pitch configuration, $F(18,57)=3.7$ ($p<.001$, 32% of the variance accounted for) and for substrate language of the listener, $F(1,74)=69.3$ ($p<<.001$, 35% of the variance accounted for). The accent position by substrate language interaction is insignificant, $F(18,38)=1,6$ (ins.).

We provisionally conclude, therefore, that individual Toba speakers of Indonesian are about twice as consistent in localizing the physical position of an accent-lending pitch configuration as are other speakers of Indonesian.

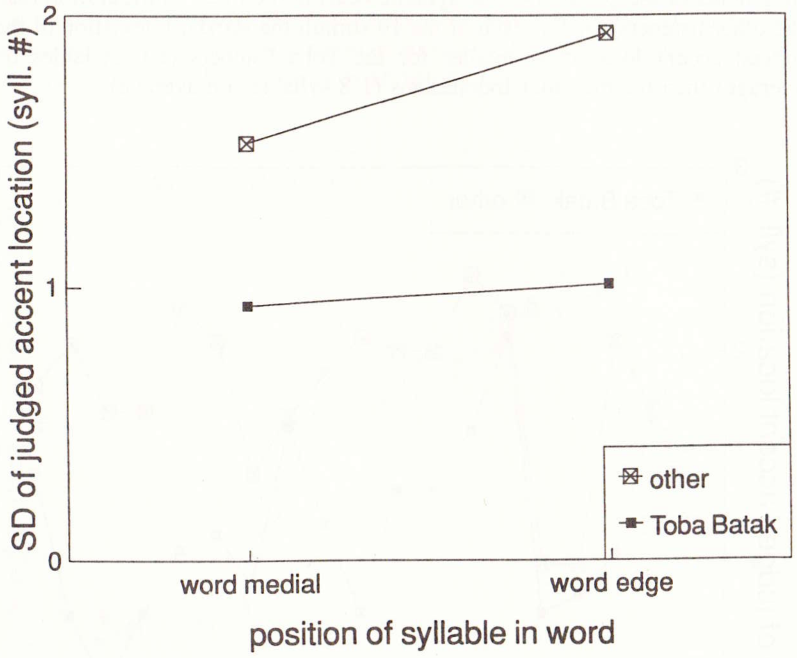

Figure 4: Standard deviation of perceived accent location (ordinal number of syllable) for physical accent positions in word-medial versus word-marginal syllables, broken down by distinctive stress (Toba Batak) and free stress (other) listeners.

Effect of word boundary. So far, the results show that Toba listeners are indeed more effective than other listeners of Indonesian in determining by ear which syllable in an utterance bears an accent-lending pitch configuration. However, in our introductory section we predicted that free stress Indonesian listeners within themselves would be more accurate in localizing an accent-lending pitch movement in the vicinity of a word boundary than in word-medial positions. Distinctive stress Toba Batak listeners, on the other hand, were predicted to be equally accurate with word-marginal and word-medial pitch movements. In order to test this prediction figure 4 presents the uncertainty in localization performance (in terms of standard deviation around the mean judgments) for Toba Batak and other Indonesian listeners broken

down by physical accent positions in word-medial syllables and in syllables at word edges.

In this analysis accents in the first and last syllable of the sentence were excluded, since it makes no sense to perceive an accent on a word before the beginning or after the termination of an isolated sentence. Accents in the syllables *ma*, *do* and *ne* were considered as word-medial, whereas all other syllables were analysed as word-marginal.

The breakdown shows that Toba Batak listeners are roughly as certain in their localization of accents word-medially as they are at word edges. Other Indonesians differ from the Toba Bataks but, counter to our expectation, they are less rather than more certain at word edges than word-medially. A two-way analysis of variance, however, shows that the interaction between the (fixed) factors (marginal versus medial position in word and substrate language) is not significant, $F(1,52)=1.2$. In fact, the analysis also shows that the main effect of position is insignificant, $F(1,54)=2.6$. The main effect of substrate language, of course, remains highly significant, $F(1,56)=28.7$ ($p<<.001$).

3.2 *Acceptability of accent positions*

In the second part of the data analysis we will consider the acceptability judgments given to the 19 exemplars of our test sentence by the two groups of listeners. If word stress is distinctive in the conception of the listeners with a Toba Batak substrate, but free in the mental representation of the other listeners, we should be able to find differences in judged acceptability of the various exemplars. We expect the listeners in the Toba group to show more outspoken preferences for one syllable position in a given word to bear the pitch accent, than do the other listeners. To the extent that either group exhibits clear likes and dislikes for a pitch movement to occur on a particular syllable, this can be taken as evidence for preferred word stress positions. Figure 5 presents mean acceptability judgments (between 0 and 10) for pitch accents on each of the syllable positions in each of the three words, broken down by listener group.

The data in figure 5 were submitted to a classical analysis of variance with substrate language and syllable position as fixed factors and with repeated measures over the two repetitions per listener. The effect of syllable position was highly significant, $F(8,67)=11.5$, $p<<.001$. Also, there was a significant but smaller effect of substrate language, $F(1,74)= 10.2$, $p=.002$, but the crucial interaction was insignificant, $F(8,58)=1.5$ (ins.). This means that Toba listeners are somewhat more lenient in their acceptability ratings throughout, and that some syllable positions are preferred over others as the bearers of an accent-lending pitch configuration, irrespective of listener substrate. It also means that any differences that we may see between Toba

Batak and other listeners in likes and dislikes for accent positions are due to chance.

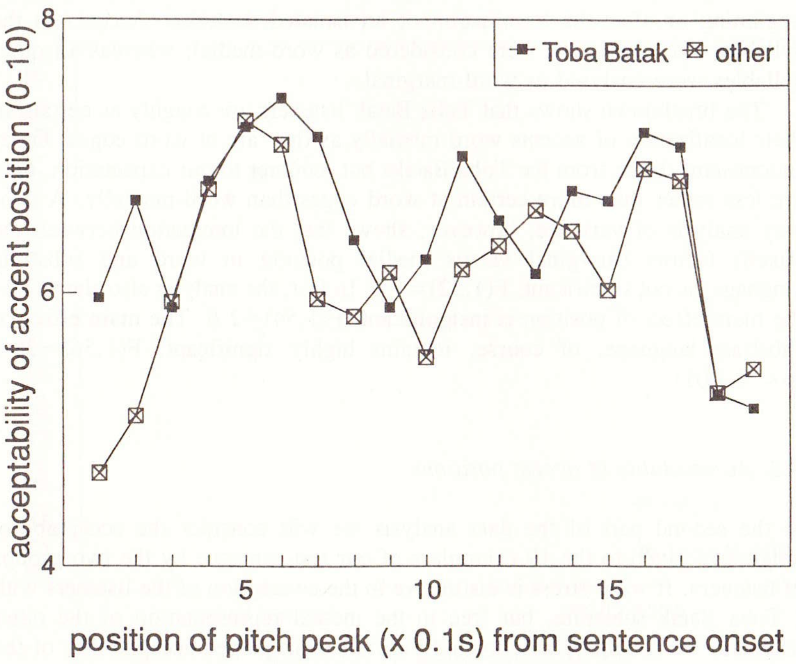

Figure 5: Acceptability of accent location (along a scale from 0 'highly unacceptable' to 10 'highly acceptable') as a function of the physical position of the rise-fall pitch configuration (in steps of 0.1s from the beginning of the sentence), broken down by distinctive stress (Toba Batak) and free stress (other) listeners.

These results of the analysis do not change when we break the statistics down to the level of individual words. However, we are now in a position to be more specific about the (dis)favoured stress locations. Post-hoc tests for contrasts (Scheffé procedure with an alpha of .05) for the sentence-initial word *kemarin* show that a pitch accent on the first syllable is disfavoured, but that accenting any non-initial syllable is equally acceptable. An accent-lending pitch configuration on any of the four syllables in the medial word *Indonesia* is equally acceptable. A pitch accent on the final syllable of *bagus*, however, is clearly disfavoured. We must reiterate, once again, that these preferences are shared by both groups of Indonesian listeners, whether Toba Batak or other.

4 Conclusions and discussion

In this paper we tested the prediction that speakers of Indonesian with a substrate language with lexical, distinctive stress will be more accurate in determining the physical position of an accent-lending pitch rise-fall configuration than other Indonesians, for whom stress position is free and non-distinctive. More specifically, we expected listeners with a distinctive stress substrate to be more accurate than listeners with free stress substrates in discriminating between pitch configurations within the same word. The two groups of listeners should be equally successful when discriminating between pitch configurations that are located in different words.

The results bear out that listeners with the distinctive stress Toba Batak substrate performed the accent localization task closer to the mark and with considerably better accuracy than the listener group with free stress substrates. This shows that specific aspects of an individual's perception can be shaped through linguistic experience.

However, the results provide no indication that the more specific prediction is also true. Distinctive stress listeners (Toba Batak) were accurate in determining the position of the target pitch configuration both within words and at word boundaries. Free stress listeners were inaccurate when localizing pitch movements whether in the vicinity of a word boundary or word-medially. The performance gap between the Toba and the other listener groups did not narrow when only pitch accents in the vicinity of word boundaries are considered. Experience with the use of (contrastive) sentence accent has not made Indonesian listeners more sensitive to the exact position of a pitch configuration in the vicinity of a word boundary than to word medial pitch movements.

We may offer two reasons for the absence of the predicted effect. One is that the exact location of a pitch movement is not crucial, in fact redundant, in marking focus on a particular word in Indonesian, since focus is generally expressed by lexico-syntactic means as well, i.e. through particles and word order. The second reason might be that we were wrong in our initial assumption that listeners could accurately assign prominence to a higher-order unit such as the word even if they were inaccurate in assigning prominence at a lower level, i.e. to a syllable. Further research is needed to settle this issue.

Our results provide no really useful information on the issue of the preferred stress position in Indonesian words. The conclusion seems reasonable that stress (marked by an accent-lending rise-fall configuration) is acceptable in any position in a word as long as it is not on the first or the last syllable of a sentence.

Finally, there is no evidence in our data that Toba listeners would entertain more stringent norms as to the canonical stress position in Indonesian words than other Indonesian listeners. Apparently there is no transfer of the substrate stress norms to the evaluation of stress patterns in the standard language.

Appendix I: Expert Dutch listeners

Using the same materials and procedures as described in section 2, the experiment was repeated with a group of 12 native speakers of Dutch who had extensive experience with Indonesian (highly advanced L2 speakers). The subjects were staff or graduate students at the Centre for Non-Western Studies of Leiden University; all were highly fluent in Indonesian and had lived in Indonesia for one to several years during some stage of their professional careers. Dutch is a language that uses stress distinctively with numerous examples of minimal pairs featuring lexical or morphological stress (Kager 1989; van Heuven & Menert 1996). We therefore predict an accuracy in the localization of accent that is at least as good as that found for the Toba Batak listeners. The results are presented in table 2.

Table 2: Distribution (in per cent) of judged position of syllable bearing pitch rise for 19 different pitch rise positions. Nominal number of responses per row is 24.

LEXICAL STRESS DUTCH LISTENERS

position F_0-rise	*judged syllable position bearing F_0-rise* 1	2	3	4	5	6	7	8	9
1	45	40			5			10	
2	44	**52**						4	
3	8	**79**						4	8
4	4	**92**	4						
5	4	**88**	8						
6		**33**	29	10		5	5	19	
7		4	26	**52**	4	10		4	
8		4		**74**	4	9	9		
9			8	4	17	**63**	4		4
10						**92**	8		
11						**96**	4		
12	5					**77**	5	10	5
13		4				8	**67**	21	
14						4	46	**50**	
15			4				8	**88**	
16							4	**96**	
17								**96**	4
18						4		**83**	13
19	4						4	**61**	31
overall	6	**21**	4	7	2	**19**	9	**29**	3

A summary graph plots the standard deviation of the judged accent location as a function of the position of the F_0-peak, with separate curves for the Dutch, Toba Batak, and 'other' listeners.

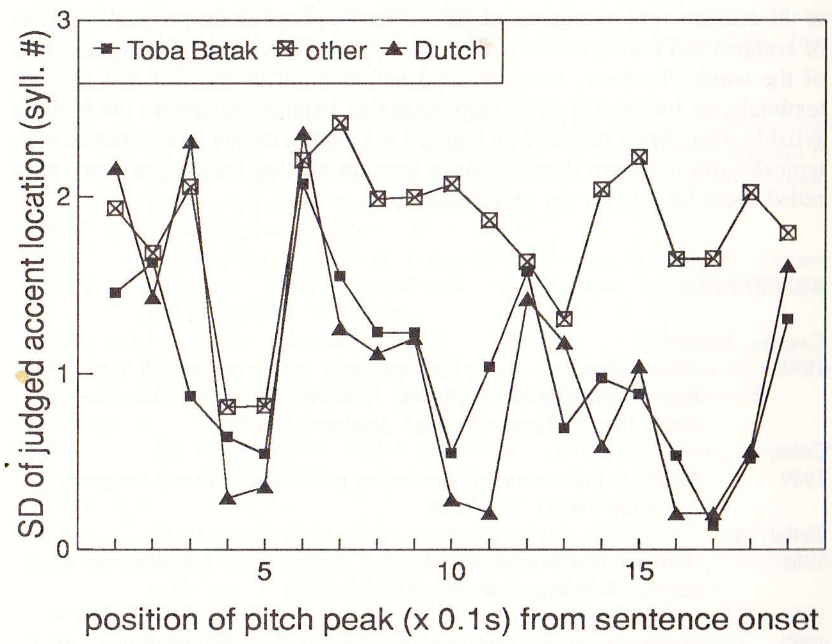

Figure 6: Standard deviation of perceived accent location (ordinal number of syllable) as a function of physical position of pitch accent (F_0-peak in 0.1s steps from sentence onset) broken down by listener group: Dutch, Toba Batak, and 'other' listeners.

The results show that the overall task performance of the Dutch listeners is quite similar to that of other distinctive stress (Toba) listeners in terms of clustering of responses around the means. The Dutch responses indicate a strong response bias: prefinal syllables function as magnets, even more so than for L1 speakers of Indonesian.

There are at least two ways to explain the strength difference in response bias between Dutch and Toba Batak listeners. First of all, notice that the standard deviations are small for the prefinal syllables of the words *kemarin* and *Indonesia*. The stress rules of Dutch are quantity-sensitive; the great majority of the words receive their stress according to the following rule: stress goes to a heavy prefinal syllable (i.e. containing a two segment rhyme) but to the final syllable if the latter is superheavy (i.e. containing a three segment rhyme; cf. Neijt & van Heuven 1992 and references given there). Therefore, the words *kemarin* and *Indonesia* but not *bagus* (the rhyme /us/ would be superheavy in Dutch), had originally been pronounced with the accent on the position indicated by the Dutch default stress rules. Although F_0

had been manipulated systematically in our resynthesis, the duration effects of the accents were left unaltered, so that the F_0-peaks on the prefinal syllables of *kemarin* and *Indonesia* were reinforced by the proper temporal organization of the word. However, a second explanation could be the fact that (Dutch) textbooks on Indonesian stress emphasize that Indonesian stresses the prefinal syllable. Our Dutch listeners, in contrast to native speakers of Indonesian who typically did not learn their language through explicit instruction, may have relied more heavily on the prescriptive rule.

REFERENCES

Caspers, Johanneke
1994 *Pitch movements under time pressure: effects of speech rate on the melodic marking of accents and boundaries in Dutch*. HIL dissertation series 10. The Hague: Holland Academic Graphics.

Cohn, A.
1989 'Stress in Indonesian and bracketing paradoxes', *Natural Language and Linguistic Theory* 7:167-216.

Cutler, A.
1986 'Forbear is a homophone: Lexical prosody does not constrain lexical access', *Language and Speech* 29:201-220.

Ebing, E.F. & V.J. van Heuven
1997 'Some formal and functional aspects of Indonesian intonation'. In the present volume, 45-61.

Halim, Amran
1981 *Intonation in relation to syntax in Indonesian*. Materials in Languages of Indonesia 5. Pacific Linguistics D 36.

Hermes, D.J.
1988 'Measurement of pitch by subharmonic summation', *Journal of the Acoustical Society of America* 83:257-264.

Heuven, V.J. van
1994 'Introducing prosodic phonetics', in: C. Odé & V.J. van Heuven (eds.) *Experimental studies of Indonesian prosody*. Semaian 9. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit te Leiden, 1-26.

Heuven, V.J. van & L. Menert
1996 'Why stress position bias?', *Journal of the Acoustical Society of America* 100:2439-2451.

Kager, R.W.J.
1989 *A metrical theory of stress and destressing in English and Dutch*. Dordrecht: Foris.

Nababan, P.W.J.
1981 *A grammar of Toba Batak*. Materials in languages of Indonesia 6. Pacific Linguistics D 37.

Neijt, A.H. & V.J. van Heuven
1992 'Rules and exceptions in Dutch word stress', in: R. Bok-Bennema, R. van Hout (eds.) *Linguistics in the Netherlands 1992*. Amsterdam: John Benjamins, 185-196.

Odé, C.
1994 'On the perception of prominence in Indonesian', in: C. Odé & V.J. van Heuven (eds.) *Experimental studies of Indonesian prosody*. Semaian 9. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit te Leiden, 27-107.

Odé, C. & V.J. van Heuven (eds.)
1994 *Experimental studies of Indonesian prosody*. Semaian 9. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit te Leiden.

Sluijter, A.M.C. & V.J. van Heuven
1995 'Effects of focus distribution, pitch accent and lexical stress on the temporal organisation of syllables in Dutch', *Phonetica* 52:71-89.

Verhoeven, J.
1994 'The discrimination of pitch movement alignment for Dutch', *Journal of Phonetics* 22:65-85.

Willems, L.F.
1987 'Robust formant analysis for speech synthesis applications', in: J. Laver & M.A. Jack (eds.) *Proceedings of the European conference on speech technology* 2. Edinburgh: CEP Consultants, 250-253.

Zanten, E. van
1989a *The Indonesian vowels; acoustic and perceptual explorations*. Leiden University, unpublished Doct. Diss.
1989b *Vokal-vokal Bahasa Indonesia; penelitian akustik dan perseptual*. Balai Pustaka, Seri ILDEP, edited by W.A.L. Stokhof (Indonesian translation of van Zanten 1989a). Jakarta: Balai Pustaka.
1994 'The effect of sentence position and accent on the duration of Indonesian words: a pilot study', in: C. Odé & V.J. van Heuven (eds.) *Experimental studies of Indonesian prosody*. Semaian 9. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit te Leiden, 140-180.

Zanten, E. van & 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.
1984 'The Indonesian vowels as pronounced and perceived by Toba Batak, Sundanese and Javanese speakers', *Bijdragen tot de Taal-, Land- en Volkenkunde* 140:497-521.
1997 'Effects of word length and substrate language on the temporal organization of words in Indonesian'. In the present volume, 201-213.

Proceedings of the Seventh International Conference on *Austronesian Linguistics*

Leiden
22-27 August 1994

Cecilia Odé & Wim Stokhof
Editors

Amsterdam - Atlanta, GA 1997