

Chapter three

Stress and accent in Indonesian*

Rob Goedemans & Ellen van Zanten

Leiden University Centre for Linguistics

3.1 Introduction

For many phonologists, the odd feature with respect to Indonesian stress is the initial-dactyl effect, reported by Cohn (1989). This effect describes the distribution of secondary stresses in words of more than three syllables, which reportedly always occur on the first syllable and every odd syllable thereafter, but never adjacent to the main stress, which is penultimate (unless the penultimate syllable contains a schwa). Such rigidity in the location of secondary stress means that in words with an odd number of syllables, like the five-syllable word *pascasarjana* ‘postgraduate’, just one secondary stress appears, on the first syllable. The observed pattern is *pàscasarjána* (Cohn 1989). In stress languages, stressed and unstressed syllables usually alternate regularly. Should we apply the default metrical rules for such languages to our example we would derive *pascâsarjána*. Hence, Indonesian seems to belong to a group of languages that are exceptional in this respect.

In this article, we will claim that this so-called initial-dactyl effect is not the most remarkable aspect of Indonesian stress. Rather, the location and even the identity of the main stress itself are open to discussion. The view that main stress in Indonesian is predominantly penultimate has been challenged on many occasions in the past, as we will see below. The confusing picture that emerges after a thorough survey of the literature is one in which stress may occur almost anywhere in the Indonesian word. Since Indonesian is a language spoken on a great variety of substrate languages the confusion with respect to stress location may have been partly caused by regional differences in pronunciation. Therefore, we must include the substrate factor in our investigations, and determine whether variability in stress location exists in the

* This research is supported by the Royal Netherlands Academy for the Arts and Sciences (KNAW) under project number 95-CS-05. We would like to thank the staff of the Erasmus House in Jakarta, in particular Herman Poelman, as well as Wim van Zanten, for running the perception experiments. Moreover, we thank Johanneke Caspers, Laura Downing and Vincent van Heuven for their constructive comments and Lilie Roosman for providing data and judgements on Indonesian intonation. This article will also appear in D. Gil (ed.), *Malay/Indonesian linguistics*. London: Curzon Press.

Indonesian spoken by these substrate groups. If there is only variation *between* substrate groups the confusing picture may have been caused by different people describing the Indonesian of speakers with different substrates. However, if there is variation *within* substrate groups, we must conclude that Indonesian stress is *free* (i.e. unpredictable by phonological rules, and variable *within* words).

After a short digression on the background of this research, we will discuss several experiments we conducted to determine the location of main stress in Indonesian. Section 3.2 discusses the details of a production experiment we conducted to gather data on the phonetic properties of prominent syllables in the Indonesian spoken by a Toba Batak and a Javanese speaker. In sections 3.3 and 3.4, we present two perception experiments in which we asked Toba Batak as well as Javanese listeners to judge the acceptability of stimuli derived from the words in the production experiment. We created these stimuli through manipulation of the original words, such that the phonetic properties we found to signal prominence occurred on a different syllable in each stimulus. On the results of these perception experiments we will base the conclusion that Indonesian does not have word stress at all. In section 3.5, we will discuss the consequences of this conclusion for the prosodic system of Indonesian. Section 3.6 sums up and briefly discusses the main conclusions.

3.1.1 Background

When listening to speech, we often perceive some syllables as more prominent than others. In many languages, the patterns of prominence we observe are not random or coincidental. They reflect organization at a more abstract level. Per word, for instance, there is often one syllable that is the most prominent, whilst the other syllables are perceived as “weak” or “strong” in an alternating rhythmical pattern. The abstract linguistic phenomenon that governs these word-level prominence patterns is called *stress*. Phonological rules express the hierarchical (weak-strong) relations between the syllables of a given word, in most cases selecting one specific syllable as the most prominent one: the *main stress* (see, among others, Hayes 1995). In some languages the rules are more straightforward than in others. In Czech, for instance, main stress is uniformly initial, but in Dutch we must use a more complicated rule to determine which of the three syllables at the right word edge will bear main stress. What these two languages have in common is that they employ phonological rules predicting stress locations. However, for some Dutch monomorphemic words (15% according to Langeweg 1988), the location of stress is not predictable. This group includes cases in which stress serves to contrast between lexical items that are segmentally identical but differ in meaning, like *voórkomen* ‘to appear on trial’ and *voorkómen* ‘to prevent’. It will be clear that only one of these two adheres to the stress rules of Dutch, the other being an exception. These exceptional cases are said to be *lexically specified* for stress; the location of stress must be learned from the dictionary, so to speak. In other languages, like Russian, stress location is lexical in 100% of the words. Traditionally however, Russian is said to have *fixed* stress, because the location of stress in any particular monomorphemic word is always the same. In that sense, the location of stress in

Russian is also predictable. Given this definition of fixed stress one might expect a label to exist for languages in which stress may occur in a different location each time a word is pronounced. These languages are sometimes said to have *free* stress. However, it is unclear what that label signifies. If stress is truly unpredictable, even on a lexical level, there is no place for it in the phonology of the language. We follow van Heuven (1994: 18) in his claim that having free stress is tantamount to having no stress at all.

From a phonetic point of view, the notion of word-level prominence is much less clear. It is very difficult to determine a unique set of acoustic properties that convey stress related prominence. The phonetic manifestation of stress depends on several factors, and is not uniquely defined across languages. For English, Fry (1958) determined that stressed syllables can be characterized by higher pitch, longer duration, greater loudness and more clearly pronounced vowels. These phonetic *correlates* of stress may have different values depending on sentence position, communicative relevance and some phonetic considerations we will not go into here. Moreover, the absolute values of the correlates may differ across languages, even to the point where one or more correlates are completely absent (Dogil 1999).

To complicate matters further, word-level prominence is not the only source for the prominence of syllables we observe when listening to speech. At the level of the phrase, the speaker marks one word (or word group) as communicatively relevant; he places this constituent in *focus* by locating an *accent* on its prosodic head (see, among others, Baart 1987, Ladd 1996). As a result, (one of the syllables of) the prosodic head becomes prominent. Usually this prominence takes the form of a conspicuous pitch movement. In the question-answer pair in (1a), the constituent between square brackets is placed in focus by an accent on the first syllable of ‘coffin’ (accent in (1) is indicated by capitals). The word ‘coffin’ is the head of the focus domain. The stressed syllable is the head of ‘coffin’ at the word level (Bolinger 1958) and is accented.

- (1) a. Q: What did John make?
 A: John made [a wooden COffin]_{+F}
 b. Q: Did you say coffer or coffin?
 A: I said cof[FIN]_{+F}
 c. Q: Did he make a wooden coffin or an iron one?
 A: He made a [WOODen]_{+F} coffin

In cases like (1b), the accent falls on some other syllable than the stressed one, narrowing down the focus domain to only that syllable. Constituents other than the head of a constituent group can also be placed in narrow focus, as is illustrated in (1c). Thus, when a word is placed in focus, the syllable that carries main stress is normally accented. Hence, in languages like Dutch and English, which use both accent and stress, the default accent location within a focus domain is predictable because the location of stress is predictable. The difference between the two phenomena is that accent location may vary depending on the speaker’s intentions, while stress location may not. At the phonetic (perceptual) level, words in focus bear properties of both word and phrase-level prominence. Part of the problem to define stress phonetically is caused by such compounding of word- and phrase-level

prominence. It is, however, imperative that we separate the acoustic correlates of stress and accent to facilitate separate investigations. We follow Bolinger (1958) and Sluijter & van Heuven (1996) in their claim that a prominence-lending pitch movement is the main correlate of accent, while duration, loudness and vowel quality are the main correlates of stress.¹ Note, however, that the link between these correlates and either word- or phrase-level prominence is not absolute. Exceptions probably abound. A more reliable (phonological) difference is that word-level prominence (the result of stressing) is fixed on a particular syllable for each word, while the location of phrase-level prominence (the result of accentuation) may vary. We will use this distinction later in our argumentation.

A further problem with the phonetic correlates of stress that is relevant to our discussion is formed by the apparent existence of languages that do not phonetically mark stressed syllables at all, i.e. the stressed syllables are not prominent at the word-level. In these cases, the location of the syllable that carries main stress is only revealed when the word is in focus, because in that case it will be prominent at the phrase level. Non-prominent stressed syllables are labeled rhythmically or metrically strong (Goedemans, van der Hulst & Visch 1996, Ladd 1996). Ladd claims that French is a language for which metrically strong syllables are needed. In French, he claims, the word-final syllable is phonetically non-distinct, yet it invariably serves as the docking site for phrase-level prominence. Goedemans et al. (1996) use rhythmically strong syllables in the analysis of languages that have no overt rhythm, but which do use rhythmical feet to calculate main stress location.

After this introduction of the theoretical notions we will use, it is now time to take a closer look at Indonesian.

3.1.2 Indonesian stress

There is an ongoing discussion on stress position in Indonesian.² Traditionally, most authors claim that the penultimate syllable is stressed, unless this syllable contains a schwa, in which case stress is final (Alieva, Arakin, Ogloblin & Sirk 1991: 63; Teeuw 1984: 9). However, Laksman (1994) found evidence that schwa can be stressed as well as any other vowel. Working in a current metrical framework, Cohn (1989) and Cohn & McCarthy (1994) present a set of rules by which the patterns of main stresses in Indonesian can be derived. They also describe secondary stresses, which fall on the first syllable and every odd syllable thereafter (but never on the syllable abutting the main stress) in words of four or more syllables, as in (2).

- | | | | | |
|-----|--|----------------|--|------------------|
| (2) | sòlida [́] rita [́] s | ‘solidarity’ | mà [́] syarà [́] ka [́] t | ‘society’ |
| | pà [́] scasa [́] rja [́] na | ‘postgraduate’ | sà [́] ndi [́] wà [́] ra | ‘theatre, drama’ |

¹ Though Cohn & McCarthy (1994: note 3) realise that the relationship between stress and intonation in Indonesian warrants further investigation, they use words in isolation to impressionistically determine the location of stress. In such citation forms, the accent and stress correlates cannot be separated.

² We offer only a summary here. For a more elaborate literature survey on Indonesian stress the reader is referred to Odé (1994: 39–41).

Opposite to these apparently iron-clad rules we find the opinion that main stress is on the final syllable of the word (Samsuri 1971) or that Indonesian has no word stress at all (Zubkova 1966, Halim 1974). According to Halim (1974: 111–113), prominence depends on the position of the word in the sentence: before a sentence-internal boundary it falls on the final syllable of the word preceding the boundary, whereas sentence-final prominence falls on the penultimate syllable of the last word of the sentence. Recent investigations reveal a general preference for speakers to stress the pre-final syllable (van Zanten & van Heuven 2004), but free variation of stress position is commonly observed, especially in longer words (van Zanten 1994: 161–163).

Most authors state that complex words (base plus one or more suffixes; prefixes are generally agreed not to influence the stress pattern) have the stress on the penultimate syllable regardless of word-internal structure (e.g. Lapoliwa 1981: 127–131; Cohn & McCarthy 1994), but de Hollander (1984: 27–28) and Alieva et al. (1991: 64) claim that in some cases stress is maintained on the penultimate syllable of the base when a suffix is attached to it. Prentice (1994: 417) proposes a solution to this controversy that is based on the fact that Indonesian is spoken on a variety of substrate languages. Prentice divides the Indonesian speaking world into two regions: A ‘Western’ region (Kalimantan, Sumatra), where suffixation does not induce a rightward stress-shift, and an ‘Eastern’ region (Java, Sulawesi and eastward), where stress falls on the penultimate syllable of the word, regardless of its internal structure. We follow Prentice in the assumption that the substrate language of the speaker is of crucial importance for the realization of stress in Indonesian. For this reason, we decided to include the substrate factor in our experiments.

Indonesian stress is phonetically only weakly marked (Teeuw 1984: 9). Nonetheless, deviations from the correct pronunciation “sound awkward” (Moeliono & Dardjowidjojo 1988: 73). No phonological rules or structural differences based on stress are observed. Finally, stress does not seem to be communicatively relevant; it at least serves no contrastive functions (van Zanten & van Heuven 1998). The inconsistent stressing, and the apparent absence of clear phonetic correlates for stress, prompt Ladd (1996: 58) to say that the penultimate syllable in Indonesian is not overtly stressed (i.e. not prominent at the word level), but that it is metrically strong, since it serves as the docking site for phrase-level prominence. We feel that this is by no means a clear-cut case. Many of the cases of *non*-penultimate prominence reported above come from research on words in focus. It seems that, at least in those cases, the penultimate syllable is not the metrically strong anchor point for phrase-level prominence.

With respect to the location of prominence within the phrase there is general consensus. Generally speaking, the final word of the phrase is the head and carries the accent (it is prominent at the phrase-level). Samsuri (1978) claims that any constituent of a sentence can be put in focus by making it prominent with a pitch movement. After this accented constituent, he claims, a final constituent spoken on a low pitch may follow.

3.1.3 The experiments

The combined reports on Indonesian stress and accent sketch a confusing picture. The first aim of this study is to resolve this confusion by determining the location of Indonesian word stress through experimental research. We keep open the possibility that Indonesian has a stressed (or metrically strong) penultimate syllable, as advocated in the traditional literature. We will measure the acoustic properties that are relevant for prominence for all the syllables in six four-syllable words in a production experiment. We will try to determine whether the most prominent syllable is in a fixed (penultimate) position or not, and whether the source of this prominence is a word- or a phrase-level phenomenon, i.e. whether it is due to stress, accent, or both. As stated above, we will include regional background as a variable in the experiment. We will look at the speech of one speaker who speaks a substrate language in which we find clearly defined (sometimes lexical) stresses, Toba Batak, and one speaker of whom the substrate language, Javanese, is said to contain only weak stresses, the location of which is not undisputed. We expect these differences in substrate language to be reflected in the Indonesian of these two speakers.

In the speech of these two subjects, we will investigate duration, intensity and pitch, leaving aside vowel quality because vowel reduction is not a very important phenomenon in Indonesian. The data we will gather in this production experiment will also serve to create models of syllables with prototypical prominence properties in both varieties of Indonesian. These models will be used in two perception experiments, which we will now briefly introduce.

The differences that languages and dialects show in their realization of stress force us to be cautious in drawing hard and fast conclusions based on impressionistic data, or only from a production experiment. In case the stress rule in the researcher's mother tongue differs from the rule in the language he investigates (which may even have no stress rule at all), the perception of the linguist may be colored by the stress rule in his own language. In this light it seems appropriate to perceptually test the native speakers' intuition on stress position. Such intuitions have to result from carefully controlled perception experiments in which judgments are indirectly obtained (cf. Berinstein 1979). In the perception experiments presented in this article we use two experimental paradigms in which we obtain such indirect native speaker judgments from Indonesian listeners for a variety of prominence patterns.

In the first test, Indonesian listeners with the same regional backgrounds as our two speakers are asked to indicate which of two prominence patterns they prefer in the speech of both speakers. The second part is an evaluation test in which the same listeners are asked to rate the acceptability of different prominence patterns, again for both speakers. If Indonesian words are preferred and judged more acceptable when prominence is realized on the pre-final syllable we will conclude that the traditional rule (stress occurs on the pre-final syllable) holds true. If, on the other hand, acceptability and preference are not significantly influenced by the location of the prominent syllable, we will conclude that Indonesian has no word stress.

3.2 Production data

3.2.1 Method

To investigate stress position and its relevance in Indonesian, a speech-production experiment was devised for which the words in (3) were selected.

(3) masyarakat	‘society’	kacamata	‘spectacles’
laksamana	‘admiral’	dikatakan	‘it was said’
perempuan	‘woman’	cendekia	‘clever’

The target words were embedded in the carrier sentence *Dia mengucapkan kata (masyarakat)*, ‘He pronounces the word (*masyarakat*)’. Target words are thus in sentence-final (focus) position and expected to receive an accent-lending pitch movement on the stressed syllable (van Heuven 1994: 15, Samsuri 1978). Secondly, to be able to measure the properties of word-level prominence without the confusing influence of phrase-level prominence, the targets were embedded in non-final (non-focus) position in the carrier sentence: *Kata (masyarakat) itu tepat*, ‘The word (*masyarakat*) is correct’.

The target words in their carrier sentences were each read twice by two male Indonesian speakers. One of the speakers had a Javanese background; he originated from Klaten (Central Java), and had come to The Netherlands quite recently. *Javanese* is considered to be the most influential regional language of Indonesia (Poedjosoedarmo 1982, Steinhauer 1980). In 1990, over one-third of the Indonesian population spoke Javanese as a first language (Steinhauer 1994: 781). Like some of the sources claim for Indonesian, the penultimate syllable is weakly stressed in Javanese, unless this syllable contains a schwa, in which case stress is shifted to the final syllable (Ras 1982: 13). Poedjosoedarmo, on the other hand, seems to hold the view that stress is final in Javanese (Poedjosoedarmo 1982: 49; footnote 45), and Horne (1961: 26) claims that “it does not matter which syllable in a Javanese word gets the loudest stress”. Our second speaker was a Toba Batak who had lived in The Netherlands for some years but spoke Toba Batak as well as Indonesian frequently. *Toba Batak* differs crucially from Javanese in that stress can be contrastive, and in that the stressed syllable – usually the penult – is clearly marked by prosodic means (Nababan 1981: 27, 135, Percival 1981: 42–44; Roosman this volume and literature there).

All material was recorded on DAT with a Sennheiser MKH 416 unidirectional condenser microphone and transferred to a Silicon Graphics workstation (downsampled to 16 kHz), and stylized and resynthesized (’t Hart, Collier & Cohen 1990), after which the relevant pivotal points in the pitch contour were marked and stored in a database with their frequency and time coordinates. For all segments of the target words duration was measured, while peak intensity was measured for all syllables in the targets. The next section contains a summary of the results of the production experiment.

3.2.2 Results and discussion

Turning to the production data of our Toba Batak speaker we observed that, on average, the pre-final syllable was significantly longer and louder – both in and out of focus – than the other syllables. Target words in the [–focus] condition were spoken on a level pitch. When in focus, vowels in prominent syllables were spoken on a higher pitch than the rest of the word. According to van Heuven (1994: 19) pitch can only be the auditory correlate of accent if the pitch movement is steep (and minimally 3 semitones; a semitone (st) is a 6% difference between two frequencies) and occurs in a specific position within the (stressed) syllable. The stylized pitch movements of the Toba Batak speaker fit this description: the average pitch movement consisted of a steep rise of approximately 3.5 st which started (mid level) at the end of the preceding vowel. A high pitch plateau followed, which lasted for the full length of the prominent vowel. A pitch fall of around 9 st closed the contour. Moreover, the average duration of vowels in prominent syllables was approximately 50% longer than the average duration of non-prominent vowels. For consonants, lengthening was around 25%. Such lengthening effects were also attested for Toba Batak speakers by van Zanten & van Heuven (1997: 210–211); they fit in well with data on (stress) languages like English and Dutch (cf. Nooteboom 1972, Eefting 1991). Finally, peak intensity in the prominent syllable was 2.5 dB (decibel) higher, on average, than in the non-prominent vowels in the Toba Batak speech data. In our targets it was always the penultimate syllable that was made prominent in this way. In Figure 1 we present the oscillogram and original (not stylized) intonation contour of *kacamata* in and out of focus, as an example.

The speech data for the Javanese speaker were quite different. Acoustically, no duration or intensity differences between syllables were found that could be related to a pattern with penultimate prominence. On the contrary, the second syllable, considered to be unstressed in all the reports we know, was often longer than the other syllables in the [+focus] condition. As regards pitch in [+focus] condition, a small pitch rise (around 2 st) was often found on the first syllable of the target, and a relatively large pitch fall (of approximately 8 st) which started somewhere near the boundary between the pre-final and the final syllable (cf. also Ebing 1997); this apparently common Indonesian pitch contour is also reported in van Heuven & van Zanten (1997). Neither pitch rise nor fall meet van Heuven's requirements for accent perception in stress languages. The rise is well below the threshold excursion size of around 3 st, and the fall is not in a specific position in the syllable. However, impressionistically, we noticed a tendency for the fall to lend prominence to the final or penultimate syllable, depending on the syllable in which it started. In the [–focus] condition no correlates of word-level prominence were found. The only salient feature we found was a considerable rise on the pre-final syllable. This pitch movement is part of a pre-boundary rise which continues up to the phrase-level boundary between *itu* and *tepat*. Figure 2 contains oscillograms and intonation contours for *kacamata* as spoken by the Javanese speaker in and out of focus, as an example.

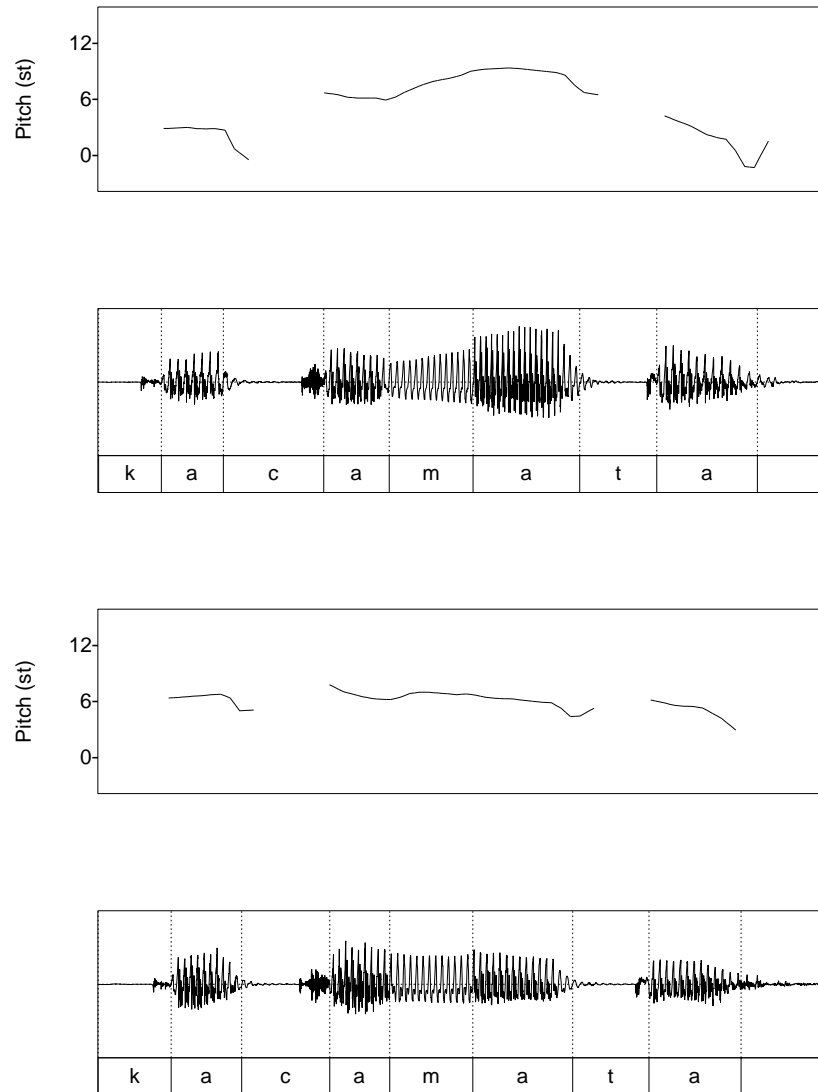


Figure 1: Oscillograms and original intonation contours for *kacamata* as produced by the Toba Batak speaker, in focus, (top) and out of focus (bottom).

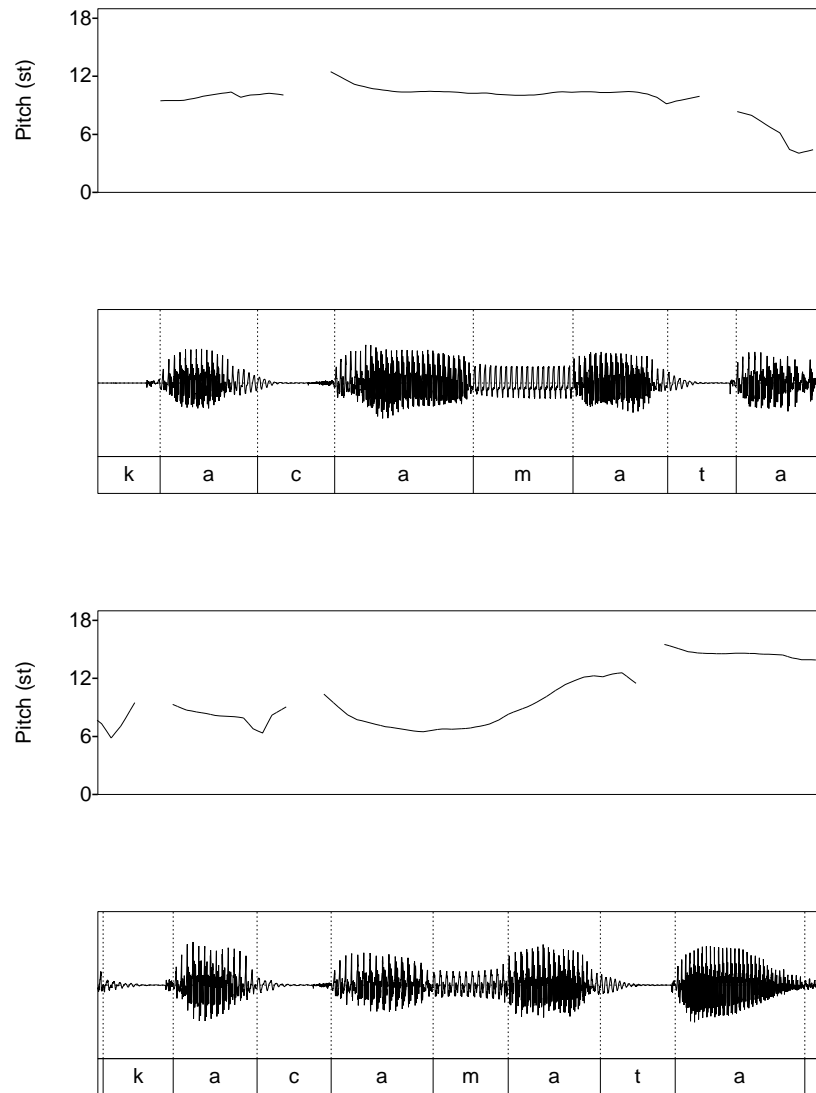


Figure 2: Oscillograms and original intonation contours for *kacamata* as produced by the Javanese speaker, in focus, (top) and out of focus (bottom).

The speech data we collected from the two speakers thus differ fundamentally with respect to the acoustic properties of prominent syllables. The evidence suggests that this difference may be caused by different abstract sources for the prominence. For the Toba Batak speaker we measured steep and sharply defined pitch movements as well as duration and loudness effects which could be related to prominence both at the word and the phrase level. For the Javanese speaker, on the other hand, we only found evidence for prominence at the phrase level (in the form of pitch movements). This can be related to the background languages of the speakers, viz. Toba Batak, a (lexical) stress language, and Javanese, a language for which stress is described as weak, and for which there is some debate concerning its location (see section 3.2.1). We will not speculate on what this means for the status of stress in Javanese, but we suspect that the “weak” stress reported by Ras (1982) could be indicative of Javanese prominence patterns being much like what we found for Indonesian.

We provisionally conclude that the Indonesian spoken by the Toba Batak speaker shows unmistakable reflexes of the clearly defined word stress in his substrate language, while in the Indonesian produced by the Javanese speaker, on the other hand, stress is absent, since no acoustic properties of word-level prominence were found. However, as we stated in section 3.1, the absence of acoustic correlates of stress (word-level prominence) in a particular language does not necessarily mean that this language has no stress. It could use non-overt stresses as alignment points for phrasal prominence.

In the perception experiments we will use a location test to verify that prominence in Toba Batak Indonesian really reflects word stress, and to determine whether there are underlying stresses in Javanese Indonesian to which accents (the source of phrase-level prominence) align. As we mentioned in the introduction, the default position for the focus-marking accent is the stressed syllable of the head of the focus domain. So, if Indonesian listeners prefer the accent to be located in one particular position per word, we will have found the stressed syllable. If we find no such preferred syllable for the accent, as we expect to be the case in Javanese-based Indonesian, the last reason to adopt stress vanishes. In that case no abstractly defined (stress) location for accent alignment will be needed. We already know that there are no phonetic correlates of stress and that no phonological or phonotactic rules that are based on stress exist in this variant.

Since we can only find stress positions in Javanese Indonesian through accent location, there is no point in considering the [-focus] condition in our perception experiments. Only in the [+focus] condition will the stressed syllable be accented. Hence, in the perception experiments, which will be discussed in the next section, we will place all the target words in final [+focus] position.

3.3 Perception experiments

3.3.1 Stimuli

The perception tests were carried out with the six four-syllable words in (3). These target words were manipulated in accordance with our findings in the production data. We constructed two sets of stimuli, one based on the production of the target

words by the Toba Batak speaker, the other on the utterances of the Javanese speaker, spoken in the final [+focus] position of the carrier sentence (*Dia mengucapkan kata*).

In order to investigate the acceptability of variable stress positions in Indonesian words we decided to compare the judgments of native speakers on stimuli with a prominent pre-final syllable (i.e. stressed according to the traditional rule; e.g. *masyarákat*) with stimuli in which the acoustic properties of prominence we found in the production experiment were transferred to one of the other syllables (*másyarakat*, *masyáarakat* and *masyarakát*). This means that four such stimuli were generated for each word. In addition, we devised one stimulus per target word, in which none of the syllables carried prominence properties, but in which the first syllable of the preceding word *kata* did. We expected this “0” stimulus to score low in the tests. Altogether five stimuli were created for each word.

For the Toba Batak speaker we varied prominence location by manipulating pitch, duration and intensity in accordance with the mean values for accented syllables we found in the Toba Batak speech data. First, the relevant acoustical cues of prominent syllables were lowered to the value they had in non-prominent syllables in each target word. For each stimulus, one syllable then received new prominence cues in the following way. Vowels were lengthened by 50%, consonants by 25%, and intensity was raised by 3 dB. Pitch movements were adapted as indicated in the top panel in Figure 3, which schematically represents the pitch contour for the Toba Batak stimuli with prominent penultimate syllables.³

Similarly, the Javanese stimuli were based on the Javanese speech data. There was no variation in durational structure or intensity that could be attributed to stress position for the Javanese speaker. The (start of the) fall was the only possible prominence lending cue. We decided to vary its position, leaving the inconspicuous rise in a fixed position. Thus, the pitch contour consisted of a 2-st rise on the first syllable of the target word, followed by a high pitch plateau of variable duration, and a steep 8-st fall. An example of these manipulations of the pitch contour is schematically represented in Figure 3 (bottom panel). In the Javanese speech data, the pitch fall often occurred somewhere in the border region between the penult and the final syllable. To sharpen the contrast between accent positions, we devised the stimuli such that the fall started within one specific syllable in each stimulus, exactly in the middle of the vowel. Finally, we included an exact (stylized) copy “S” of the original pronunciation of each of the six target words in which the fall occurred in the border region between the pre-final and final syllables.

All stimuli were superimposed on the mean pitch declination of the original utterances: a downtrend of 1 st per second for both speakers. To reduce the workload on the listeners, the first part of the carrier sentences (*Dia mengucapkan*) was deleted.

³ In the prominent final syllable, the rise was immediately followed by the fall. Only in this way could the low pitch indicating the end of the utterance be audible. We copied this contour from previously recorded speech data (cf. van Zanten, Goedemans & Pacilly 2003).

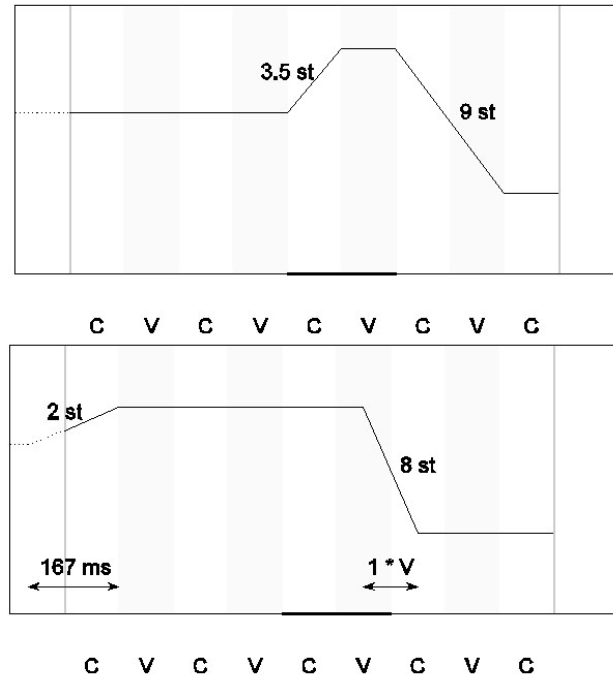


Figure 3: Schematic representation of the pitch contours for the Toba Batak (top panel) and Javanese (bottom panel) target words with prominent penultimate syllables (indicated by the bold line on the x-axis).

3.3.2 Experimental paradigms

Two types of listening experiments were devised. The first was a pairwise-comparison experiment in which subjects are requested to choose between two members of a stimulus pair. In our case, the pairs consisted of a reference stimulus (with prominent pre-final syllable) and any one version of the same target word. For the Toba Batak-based speech this amounted to 60 pairs of stimuli (that is, $6 \times 5 = 30$ stimulus pairs, in both orders). The Javanese-based stimuli consisted of 72 stimulus pairs (30 + 6 stylized versions “S”, in both orders). Stimuli were recorded on audio tape in quasi-random order.

Secondly, we devised an evaluation test in which subjects were to judge the acceptability of the individual stimulus words. Each individual stimulus was copied twice on tape in counter-balanced random orders. Consequently, the number of judgments asked was equal to that of the previous experiment: 60 for the Toba Batak-based stimuli, and 72 for the Javanese-based stimuli.

3.3.3 Listeners and procedure

Two groups of Indonesians took part in the listening experiments. These groups were selected to match the substrate languages of the original speakers, i.e. a group of 20 speakers of Indonesian who had Javanese as a substrate language, and a group of 13 Indonesian speakers with a Toba Batak substrate.⁴

The tape was played to the listeners over good-quality earphones at a comfortable listening level. Eleven listeners were tested individually at the Phonetics Laboratory of Leiden University and the remainder was tested in two groups in a language laboratory in Jakarta. Subjects were told that they were going to listen to the final parts of declarative sentences and that these had different intonations (*lagu kalimat*). They were not informed about the actual purpose of the experiments, i.e. to compare the acceptability of different stress patterns. For the paired comparison test, listeners were asked to indicate on their answer sheets which of each pair they preferred. It was made clear to them that they had to make a single choice in all cases; blanks were not allowed. For the acceptability test, subjects were instructed to rate the acceptability of each phrase on a ten point scale, ranging from 1 (“very bad”) to 10 (“very good”). They were requested to encircle the appropriate mark on the answer sheets for each stimulus phrase. Each test was preceded by three practice items. After this, the tape was stopped to answer any questions raised by the listeners. After every ten items a short beep was inserted to help the listeners keep track of the stimuli on their answer sheets. All instructions were in Indonesian. Approximately half of the subjects (i.e. half of each listener group) listened to the “Javanese” stimuli first (first the pairwise-comparison test and then the acceptability test) and then to the “Toba Batak” pairwise-comparison and acceptability tests, respectively. The other subjects were presented with the “Toba Batak” stimuli before listening to the “Javanese” stimuli; again, the pairwise-comparison test was followed by the acceptability test.

3.4 Results and discussion

3.4.1 Javanese-based stimuli

3.4.1.1 Pairwise-comparison experiment

In the pairwise-comparison experiment, each stimulus pair contained two instances of the same word; one of the comparison stimuli and the reference stimulus with prominence in penultimate position. For each substrate listener group, and all possible prominence positions, we calculated the percentage of responses in which the subjects chose the comparison stimulus as the better one of the pair. In case the reference and comparison stimuli are identical (i.e. the penultimate syllable is

⁴ A report on the reactions of a group of Jakartan listeners to these stimuli can be found in van Zanten, Goedemans & Pacilly (2003). The results for the Jakartan listeners largely coincide with the results for the Javanese listeners presented below.

prominent in both stimuli), the outcome should be exactly 50%, because it is impossible to select one stimulus as better than the other in that case. In practice, however, subjects tend to choose the first member of the pair when they are unable to make a motivated choice. This bias for the left-hand member of a stimulus pair is known as the Time Order Error (cf. Woodrow 1951, van Heuven & van den Broecke 1982). In our experiments we tried to eliminate the TOE effect by presenting the stimuli to the subjects in both orders.⁵ If Indonesian does indeed have penultimate stress, all the percentages representing responses for non-identical stimulus pairs should lie well below this 50% mark. In these cases, the comparison stimulus does not show prominence on the penultimate syllable, and should, therefore, not be selected as the better one of the pair.

In this article, we will disregard any differences between the six individual words. In Figure 4, we present the results for the two groups of listeners and the Javanese-based stimuli. In this figure, the percentage score for the comparison stimulus (Scomp) is plotted along the y-axis, and the prominence locations are placed on the x-axis (remember that in the “0” case, the first syllable of *káta* is the prominent one; “S” represents the stylized version of the original utterance). The scores for each of the two substrate groups are connected by lines.

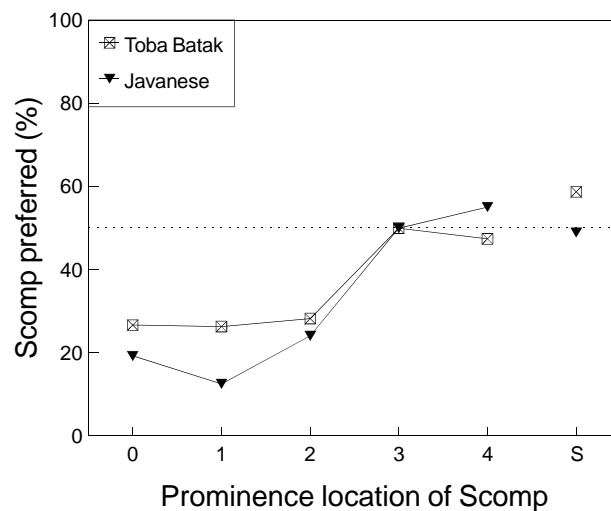


Figure 4: Javanese-based stimuli. Percentage of cases in which the comparison stimulus (Scomp) was judged better than the reference stimulus. Broken down by prominence location and substrate listener group.

⁵ Eliminating the TOE effect for the identical reference stimulus pairs proved to be impossible. We opted instead to calculate the percentage of choices for the left-hand member of the pair to obtain a measure for the size of the TOE effect in our experiment, which proved to be 9.5%. In the figures below, the score is set at the theoretically motivated 50%, which is also used in the interpretation of the statistics.

We observe that the data points for the *Javanese* listeners do not follow the pattern for penultimate prominence, in which position 3 (the penult) gets a 50% score while the rest remains well below the 50% line. Prominence on the final syllable is judged as acceptable as prominence on the penultimate syllable. A one-way analysis of variance (ANOVA, $\alpha = .05$) shows that there is a significant difference in the percentage scores; $F(5,1290) = 38.0$, $p < .001$. A post-hoc SNK analysis shows that this is attributable to a difference between a “0”, “1” and “2” group on the one hand, and a “3”, “4” and “S” group on the other. We interpret this as the difference between acceptable and unacceptable prominence locations. We postpone the discussion of the status of prominence in Javanese Indonesian to section 3.5.1.

We find the same pattern for the *Toba Batak* listeners. Both final and pre-final prominence are acceptable, but prominence on the first or second syllable (or none at all) is not. A one-way ANOVA again reveals significant differences in the percentage scores: $F(5,930) = 20.8$, $p < .001$, and a post-hoc SNK analysis again separates “0”, “1” and “2” from “3”, “4” and “S”.

Within the “unacceptable” group, “2” seems to be judged better than “1” by the Javanese listeners. We do not yet know whether this (statistically insignificant) tendency is linguistically relevant. Within the “acceptable” group, the Javanese listeners, as opposed to the Toba Bataks, seem to prefer stress on the final syllable over stress on the penultimate syllable. This tendency goes against the claim that stress is predominantly penultimate in Indonesian. Finally, the two groups of listeners both judge the stimuli with penultimate or final prominence to be approximately as acceptable as our close-copy version of the Javanese pitch contour. We take this as an indication that these manipulated stimuli sounded sufficiently natural to Indonesian ears.

So far, the evidence we have found argues against the claim that stress is penultimate in Indonesian. Let us now look at the data from our evaluation experiment to see whether these point in the same direction.

3.4.1.2 Evaluation experiment

Figure 5 shows the mean evaluation scores, represented on a scale from 1 to 10 on the y-axis, for all prominence locations (x-axis), again broken down by substrate group, as in Figure 4.

The lines connecting the scores for each group largely follow the same pattern as those in Figure 4. Prominence locations “3”, “4” and “S” again form one group which is significantly different from the combination “0”, “1” and “2” (one-way ANOVA with post-hoc SNK test: Javanese $F(5,1577) = 106.1$, $p < .001$; Toba Batak $F(5,642) = 28.3$, $p < .001$).

The results from our evaluation experiment corroborate the findings of the pairwise-comparison experiment: prominence in Indonesian, as spoken by the Javanese speaker, is acceptable on either the final or the penultimate syllable. Such variation suggests that Javanese Indonesian has free stress, which is tantamount to having no stress at all (van Heuven 1994: 18). Let us consider this possibility. Surely, the absence of stress in their own form of Indonesian should influence their perception of it in other varieties. It would be interesting to test their reactions to a

form of Indonesian that does have “real” stresses. We have determined in the production experiment that the Toba Batak speaker realizes stressed syllables in [+focus] condition with a pitch movement, a rise in intensity and lengthening. Considering the speaker’s substrate language (which has word-based stress), and his phonetic realization of the word stresses, we are positive that our manipulations of the Toba Batak speech involved the manipulation of “real” stresses. To confirm this, we included the Toba Batak listeners (who may be looked upon as the expert judges of Toba Batak stress) in this part of the experiment as well. Let us see how the subjects reacted to these stimuli.

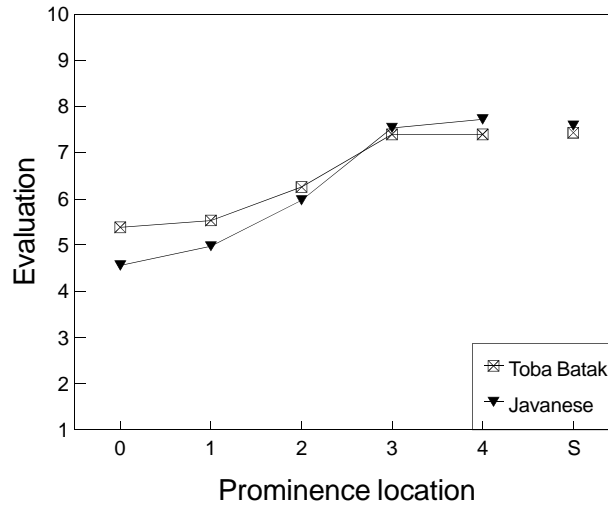


Figure 5: Javanese-based stimuli. Evaluation scores for all prominence locations, broken down by substrate listener group.

3.4.2 Toba Batak-based stimuli

3.4.2.1 Pairwise-comparison experiment

Figure 6 presents the pairwise-comparison data for the Toba Batak-based stimuli in the fashion of Figure 4. Let us first consider the *Toba Batak* listeners. The figure shows that they clearly prefer prominent penultimate syllables. The percentage score for prominence on the penultimate syllable, which is slightly, though not crucially, flattered by the bias (TOE, see section 3.4.1.1), is shown to be different from all the other scores in a one-way ANOVA with post-hoc SNK test: $F(4,775) = 14.4$, $p < .001$.

We interpret this result as a reflection of the Toba Batak stress rule in the Indonesian spoken by the Toba speaker. It seems that the Toba listeners prefer penultimate stress when listening to a Toba Batak speaker’s Indonesian. We will come back to this issue in the general discussion.

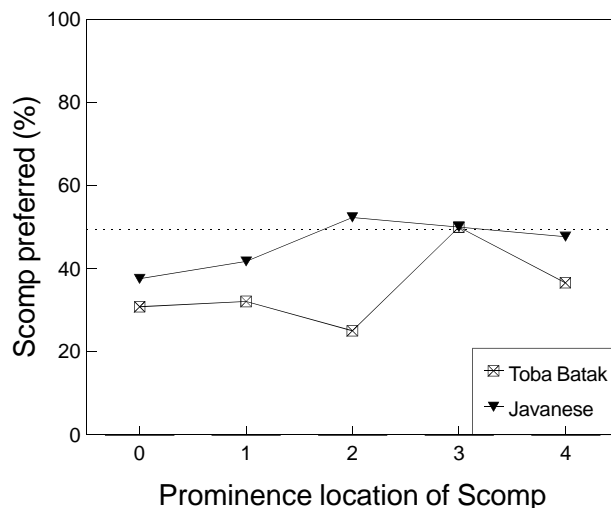


Figure 6: Toba Batak-based stimuli. Percentage of cases in which the comparison stimulus (Scomp) was judged better than the reference stimulus. Broken down by prominence location and substrate listener group.

The *Javanese* listeners react crucially different to the Toba Batak stimuli. The percentage scores for all prominence positions are much closer together than for the Toba listeners. The only significant difference is found between the scores for prominence on the second syllable of the target word and prominence on the first syllable of *káta*, as a one-way ANOVA (with SNK test) reveals: $F(4,1075) = 3.1$ ($p < .001$).⁶ There is no significant difference between any of the scores for prominence *within* the target word, $F(3,860) = 1.7$ (ins.). Apparently, a Toba-style stress realization is equally acceptable to Javanese ears on any syllable of a four-syllable word. Let us now see whether the evaluation data corroborate the pairwise-comparison data for the Toba Batak-based stimuli, as they did for the Javanese-based stimuli.

3.4.2.2 Evaluation experiment

The evaluation scores for the Toba Batak-based stimuli by our two groups of listeners are shown in Figure 7. The preference for the pre-final syllable by the *Toba Batak* listeners, which we found in the pairwise-comparison test is confirmed by the

⁶ The slight preference for the second syllable reminds us of the fact that the second syllable was also longer in the Javanese production data. Perhaps the two observations are related. As we noted in the discussion of the production data, we do not believe the phenomenon has anything to do with stress.

high score for this syllable (7.5). Statistically however, “3” differs only from “0” and “4”, $F(4,544) = 6.8$ ($p < .001$).

The evaluations of the *Javanese* listeners resemble the results for the Javanese listeners found in the pairwise-comparison experiment quite closely. The rating for prominence on the second syllable is again somewhat higher than the others, but this difference is not significant. The only significant difference is that between all the possible prominence locations in the target word on the one hand, and prominence on *káta* on the other, $F(4,1331) = 12.11$ ($p < .001$), indicating that, for Javanese listeners, Toba Batak style stresses are equally acceptable on all syllables in the word.

As in the pairwise-comparison experiment, the Javanese listeners seem indifferent to where the stress falls in Indonesian four-syllable words pronounced by the Toba Batak speaker. We suppose that the difference between the prominence patterns of the Javanese and the Toba Batak speaker is responsible for this difference in perception.

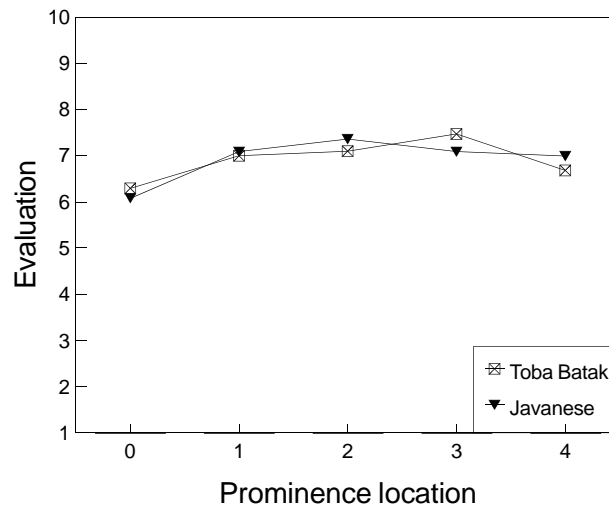


Figure 7: Toba Batak-based stimuli. Evaluation scores for all prominence locations, broken down by substrate listener group.

3.5 Discussion

3.5.1 The status of word stress in Indonesian

Recapitulating the previous sections, we briefly consider the importance of the presented results. Firstly, as Figures 4 and 5 clearly show, listeners from two different backgrounds judge the “Javanese-style” prominence on either the final or the penultimate syllable to be acceptable. Figures 6 and 7 show that the Javanese

listeners are indifferent to the location of “Toba Batak-style” prominence. The Toba Batak listeners themselves clearly prefer penultimate prominence in this case.

As was noted above, stress in the Javanese variant, if it exists at all, could only be found through the accent that normally aligns to it (Ladd 1996). It appears, however, that accent location in Indonesian as spoken by Javanese speakers is not associated to any particular syllable. Consequently, we must conclude that this variant has no need for word stress. There are no phonetic correlates for it, neither is it used at the abstract phonological level. We claim, then, that the Javanese variant does not have word stress, and will, from now on, consider prominence in this variant to be the result of *accent* only. The Toba Batak speech contained rather canonical stresses, realized by greater intensity and longer duration as well as pitch movement. These properties suggest that we must look upon the Toba Batak stress as a “real” one.

Turning our attention, then, to the influence of substrate-listener group on prominence perception, we observe that the reactions of the two substrate groups to each other’s prominence patterns support the claim that the Javanese variant does not have stress. While keeping in mind the split between accent and stress that we introduced above, we are able to explain some of the differences in reactions that we elicited with our two different types of stimuli. Stress locations other than penultimate in the Toba Batak-based stimuli are correctly rejected by the Toba Batak listeners themselves, because they are used to hearing stresses that occur in strictly defined positions (mostly penultimate) in their substrate language, and would no doubt also reject stimuli with alternative stress locations in that language. The Javanese listeners, however, do not differentiate between the different Toba Batak stress locations. They either cannot hear, or do not care, on which syllable the Toba Batak stress properties are realized. We can simply say that stress has no meaning to them, which is all the more reason to assume that word stress is neither a feature of Javanese nor of the Javanese variant of Indonesian.

When we consider their reactions to the prominence differences in the Javanese-based stimuli, we find that the Javanese listeners are not totally insensitive to intonation movements. Instead, they prefer the accent to occur at the right edge of the word and judge stimuli to gradually worsen as the accent moves further to the left edge of the word. Remarkable, in this respect, is the behavior of the Toba Batak listeners. We expected them to interpret the Javanese accent as a stress, and consequently, allow it to occur only on the penultimate syllable. The fact that they copy the behavior of the Javanese listeners in this case, and do accept final as well as pre-final prominence, indicates to us that they enter a different “mode” when listening to the group of speakers that defines (through group size, but also through greater influence in politics and the media) the most common form of Indonesian. We should mention here that all but one of the Toba Batak listeners lived in Jakarta.

We note that, with respect to stress (or prominence) there is no uniform rule for Indonesian. A long history of debates on the exact location of stress is indicative of the absence of such uniformity. We add to this evidence the large differences in the reactions of listeners with different substrate backgrounds to variants of Indonesian stress produced by speakers with other substrate backgrounds. However, if we are forced to choose one particular stress or prominence rule for Indonesian, it would be the rule used by the influential Javanese speakers. This choice is motivated by the

fact that the Toba Batak listeners react to the Javanese variant of Indonesian in the same way the Javanese themselves do.

As a final remark concerning stress location in Indonesian, we note that the absence of word stress and the relatively unrestricted accent location, have one predictable consequence with respect to focus placement. Since the accent can, in principle, occur on any syllable of the word on any given occasion, it should be impossible for speakers of Javanese Indonesian to narrow down the focus to one syllable, as in (1b) which is repeated (without the focus-domain information) in (4a) for convenience. They should not be able to make contrastive accents like the one in (4a), simply because there is no default position for the accent; all alternatives are allowed and place the entire constituent in focus. In our example that would mean that the answer to the question in (4a) given by an Indonesian speaker could just as well have the accent on the first syllable of the word *coffin* as on the second. For Indonesian speakers the initial accent on *coffin* is ambiguous. It may signify a contrast with *coffer* as well as, for instance, *muffin*. Ebing (1997: 92–95) indeed finds that such contrasts are not possible in Indonesian, in spite of the fact that Halim (1974: 77–79) reports the occurrence of such contrastive accents.⁷ Ebing found that Indonesian listeners could not correctly perceive the intended focus structure of utterances like those in (4b).

- (4) a. Q: Did you say coffer or coffin?
 A: I said coffIN
 b. Maksud saya caRI, bukan caTAT (Indonesian)
 mean I search not write.down
 ‘I meant search, not write down’
 c. *I said deFLECT, not inFLECT
 *I_k zag een koNIJN, niet een toNIJN (Dutch)
 I saw a rabbit not a tuna.fish

Also, we expect Indonesian speakers to have difficulties in the production of contrastive accents in languages that do use them, like English and Dutch. Judging from our own observations, and that of our informant (Roosman p.c.), mistakes of the type in (4c) abound. In our view, these observations support the claim that the variant of Indonesian these speakers use has no fixed stress (and accent) rule.

In conclusion, we state that neither the indifference of Javanese listeners to Toba Batak word-stress location, nor the acceptance of penultimate and final Javanese-style prominence by both groups of listeners, point in the direction of a penultimate stress. On the contrary, the data we have gathered in these experiments forcefully refute the claim that stress is predominantly penultimate in Indonesian. We did, however, find it to be penultimate in one of its variants, in which Toba Batak listeners were subjected to the Indonesian of a Toba Batak speaker, but we can hardly claim that the pattern of this variant can be generalized over the entire Indonesian community. As we have shown, there is no “Indonesian” with respect to

⁷ Halim worked with a variant of Indonesian from Sumatra. It is possible that this particular variant has stress, and that its speakers can realise contrastive accents. Our informants, however, assured us that such accents were impossible in Javanese Indonesian.

stress; other substrate groups might speak Indonesian with other stress patterns yet again. However, even if we would take the term Indonesian to cover its most common variant (Javanese Indonesian), the claim that stress is penultimate cannot be maintained. On the contrary, we have found compelling evidence for the claim that this variant does not have stress at all. However, since it does contain prominent pitch movements, we tentatively adopted the view that this variant does make use of accents. In the light of many theories on the phonology of intonation, in which accents are invariably linked to stressed syllables, Indonesian poses a problem. In the next section we will discuss that problem in greater depth.

3.5.2 Accent in the Indonesian phrase

With the adoption of the claim that Javanese Indonesian prominence represents accent rather than stress, we have left the domain of the word-level phenomena. The domain of the accent is, by default, the phrase. Only in special circumstances can that domain be smaller than the phrase, as we have seen in the introduction.

In this section we will present some speculative proposals, on the assumption that the Javanese Indonesian prominence patterns are phrase-level phenomena. We do not present this as the final argument that closes the case, but would rather look upon it as an incentive for much needed further research on the accentual system of Indonesian (and prosodically similar languages). Evidence for the claim that Javanese Indonesian prominence patterns are indeed phrase-based can easily be found. The acceptability of prominence, for instance, tends to rise as it occurs further towards the right edge of the phrase-final word. Such a gradual rise in acceptability cannot be linked to a word-level phenomenon, but it fits in well with the notion of an accent that is not bound to a particular syllable, but should occur somewhere near the right edge of the phrase.⁸ An important observation in this respect is that the location of the pitch movement is not exactly aligned to the syllable positions. In our production experiment we found many cases in which the pitch fall occurred somewhere in between the penultimate and the final syllable. Some of these cases formed the basis for the stylized versions (“S”) of the words we included in the stimuli of the experiments. These stylized versions were invariably

⁸ Alternatively (Gil p.c.) we might claim that accent is foot-based, and must occur on either the penultimate or the final syllable of the phrase-final word. Though we found no evidence for metrical feet in the Javanese variant of Indonesian, the fact that the canonical monomorphemic word in Indonesian is disyllabic constitutes evidence for foot-sized units at the morphological level. It might be that these units form the domain within which accents can occur. Halim’s (1974: 111 note 27) claim that accent may fall on either one of the two final syllables but not on any other certainly points in that direction. However, since we have found no sharp drop in the acceptability of stimuli in which stress/accents occurred on the antepenultimate syllable (as was also found by van Zanten & van Heuven 2004), we are inclined to adopt the phrase-based option until further research has resolved this issue. We note that adopting a foot-sized domain for accentuation at the right word edge would only marginally influence our argumentation. Our claims crucially hinge on the fact that accent location is unpredictable within whatever domain we choose.

judged as good as the stimuli in which the pitch fall was clearly located on either the penultimate or the final syllable.⁹

Furthermore, the adoption of a phrase-based accent straightforwardly explains the behavior of our listener groups. The judgments of the Toba Batak listeners changed when they were listening to the Javanese-based stimuli. In that mode, they were apparently judging phrase-level accents, a task that should not be beyond them, since they are well acquainted with the Javanese-style prominence patterns and they have proven to be quite sensitive to differences in pitch movements (van Heuven & van Zanten 1997). The Javanese listeners, on the other hand, cannot judge the Toba Batak word-level stresses because they are used to hearing phrase-level prominence patterns only (they considered the stresses “too harsh”).

In unequivocal stress languages, the phrase-level accent phenomenon is tied to the word-level phenomenon of stress because an accent always aligns to the stressed syllable of the phrasal head (Ladd 1996, Pierrehumbert & Hirschberg 1990). In this respect, Indonesian presents a problem to current theories on accentuation, or, in a wider sense, intonation. These theories are all developed with stress languages, like English and Dutch, in mind, as evidenced by Cruttenden’s remark that “rhythmical patterns are the backbone of intonation” (1986: 6). But what happens if there is no word-level stress for the accent to align to?

In order to answer this question we must first consider the phenomenon of *boundary marking*, which may be more intimately related to accent in Indonesian than in the stress languages we mentioned. We already noted that the accent in Indonesian is typically located on the last word of the phrase. This means that it always occurs quite close to the pitch movements that mark the end of the phrase: the phrase accent and the boundary tone (Pierrehumbert & Hirschberg 1990). It seems that the distinction between accent lending and boundary marking intonation movements is very difficult to make in Indonesian. Looking at the IPO-method of intonation description for Dutch (’t Hart et al. 1990), we observe that the only difference between the full accent-lending fall A and the full boundary marking fall B is one of timing. As we have seen above, timing is completely irrelevant in Indonesian, a result that replicates perceptual findings of van Zanten & van Heuven (2004). Moreover, Ebing (1997), who compared discrimination of accent and boundary marking by Indonesian and Dutch subjects, notes that “crucially, there was a substantial interdependence between accent and boundary perception” and “boundary-marking and accent-lending functions are less distinct in Indonesian than in Dutch” suggesting that “this difference reflects a typological difference between languages with a phrasal accent rather than lexical stress – here represented by Indonesian – on the one hand, and languages with both lexical stress and accent – here represented by Dutch – on the other” (Ebing 1997: 111–113). In the same vein,

⁹ According to Suparno (1993: 70–71) there is a difference in meaning depending on the position of the accent in the word. When the accent is on the final syllable there is a causal relation with another utterance, but when the accent is on the penultimate syllable there is no such causal relation. In a recent perception experiment Laksman (1996; cf. also Laksman & van Heuven 1999) found a correlation between accented final syllables and the perception of anger. Unfortunately, we have no data on the exact position of the accent-lending pitch movement in these cases.

Beaugendre (1994: 118) mentions that it is difficult to distinguish between accent lending and boundary marking pitch movements in French. He claims that the accent (“accent fixe”) in principle has a demarcative function.

In the light of the evidence above, we should consider the possibility that accent and boundary marker are the same thing in Indonesian. It is important to note that accent-lending pitch movements are not necessary to place words in focus. In principle, the word that is in focus is predictable: it is the final word in the phrase.¹⁰ The way in which the utterance is divided into phrases may be marked by syntactic means. Suparno (1993: 72–79), for instance, mentions nine linguistic categories which can indicate phrase structure in Indonesian as spoken in Malang (‘Konstruksi tema rema’ in Suparno’s terminology). For a survey of the literature on the relation between sentence structure and intonation in Indonesian, cf. Suparno (1993: 39–67).

Hence, the Indonesian listener can usually infer focus from sentence structure. No focus marking by means of pitch seems to be needed.¹¹ Accent does not seem to have a well defined function in Indonesian, while boundary marking is clearly of crucial importance. In languages like Indonesian, focus cannot be used to contrast between non-phrase final words, as is done in (1c), repeated in (5) for convenience. In Indonesian, the accent must occur phrase finally, on *coffin*. Hence, we predict that Indonesian speakers cannot correctly interpret such sentences, which is reminiscent of their problems with focus at syllable level presented in (4). Judging from our own impressions of Indonesian speech, and the type of mistakes Indonesian speakers make in Dutch and English, we firmly believe that contrastive accents on the phrase level are impossible.¹²

- (5) Q: Did he make a wooden coffin or an iron one?
A: He made a [WOODen]_{+F} coffin

The clear difference between the functionality of accent and boundary markers prompt us to look upon the whole intonation contour at the end of the Indonesian phrase as a boundary-marking pitch movement.

Finally, we note that our views on Indonesian intonation as primarily signaling boundaries are compatible with our observation that the boundary marking pitch movement is mostly, but not necessarily, on the penultimate syllable.

First note that the final part of the intonation contour we observed in our data consists of a high pitch level followed by a fall, and ending on a low pitch. It is the change in pitch (from level to falling) that we perceive as an accent (and some have

¹⁰ After the focalized constituent a reduced contour may follow which contains a defocalized (‘retracted’) constituent (Halim 1974: 115–117, 125). Such a reduced contour – elsewhere called *ekor* ‘tail’, cf. Suparno (1993: 71–73, 80–83) – is spoken on a low pitch (no accents are permitted in the tail). Cf. also Stoel (2005, this volume) on Manado Malay.

¹¹ However, we often found a slight rise in pitch on the first syllable of the phrase-final word which seemed to enhance its prominence, thus alerting the listener to the presence of a prosodic head at an early stage.

¹² Ebing (p.c.), for instance, reports mistakes like ‘I do not have black LAbel, only red LAbel’. However, we will refrain from drawing any firm conclusions until experimental evidence proves our intuition to be correct.

interpreted as stress in the past). Remember that, whatever we may call them, these elements are all part of the Indonesian phrase-final boundary marker we postulate. If we take that into account, we may explain the predominance of penultimate-syllable prominence as a statistical effect. In a language where “accent” alignment is in principle free, the end of the intonation contour is the only fixed point. It aligns to the end of the utterance-final word. It is to be expected that, by default, the contour aligns such that most of the fall occurs on the final syllable, while it starts on the pre-final syllable; the simple reason being that the fall needs space to be expressed but normally does not take more space than necessary. The point is that this is only a tendency. The start of the fall (high pitch point) that would indicate the pitch accent in English and Dutch (and must align to the stressed syllable of the prosodic head in those languages, however far from the word edge this syllable occurs) may in Indonesian freely occur earlier or later than the penultimate syllable. In those cases, the duration of the fall is simply lengthened, or shortened, respectively. Such statistical considerations might also explain the observations of van Zanten & van Heuven (2004), who note that pre-final closed syllables “seem to attract stress”. That might simply be so because of the longer duration of such pre-final syllables, which makes it more likely that the starting point of the fall occurs there.

3.6 Conclusion

The most important conclusion we draw from the results of our experiments is that there is no reason whatsoever to assume that stress in Indonesian always falls on the penultimate syllable if it contains a full vowel. We have shown that speakers with different substrate languages behave differently with respect to stress realization and perception. Even if we set this caveat aside, however, and concentrate on the variety spoken by the most dominant substrate group (Javanese), we conclude that there is good reason to exclude the penultimate stress hypothesis. In our view, the rule that drives prominence patterns in the influential Javanese variety of Indonesian is phrasal. Possibly, the only phonological rule that is relevant for accent location in Indonesian states that it must occur somewhere at the right edge of the phrase.

Since there is no evidence for strict patterning of the main stress, we think any proposal that describes the patterning of secondary stresses with respect to the main stress must be received with great caution. The initial-dactyl effect that we mentioned in the introduction as the most striking feature of Indonesian stress according to many phonologists, might well occur in Indonesian if prominence in a five-syllable word is penultimate and the slight rise in pitch on the first syllable is interpreted as a secondary accent. However, this initial dactyl is by no means the result of strict patterning according to metrical rules.

As we have seen in the introduction, some alternative stress rules have also been proposed for Indonesian in the past. We believe that this variation is caused by the differences in prominence patterns that we observe in speakers with different substrate languages, but in some cases also by the fact that the linguists in question have tried to construct stress rules for something that is not stress at all. The impression to the non-native linguist may have been that there should be a stress on

the penultimate syllable, but native listeners appear to perceive things quite differently.

This research constitutes one example of the crucial importance of careful phonetic experimentation, which provides a basis for phonological claims, and can be instrumental in the resolution of long-standing phonological debates.

References

- Alieva, N.F., V.D. Arakin, A.K. Ogloblin & Yu.H. Sirk (1991) *Bahasa Indonesia: Deskripsi dan teori* [The Indonesian language: Description and theory]. Yogyakarta: Penerbit Kanisius.
- Baart, J.L.G. (1987) *Focus, syntax and accent placement: Towards a rule system for the derivation of pitch accent patterns in Dutch as spoken by humans and machines*. PhD dissertation, Leiden University.
- Beaugendre, F. (1994) *Une étude perceptive de l'intonation du Français* [A perceptual study of French intonation]. PhD dissertation, Université Paris XI.
- Bernstein, A.E. (1979) A cross-linguistic study on the perception and production of stress. *UCLA Working Papers in Phonetics* 47, 1–59.
- Bolinger, D. (1958) A theory of pitch accents in English. *Word* 14, 109–149.
- Cohn, A. (1989) Stress in Indonesian and bracketing paradoxes. *Natural Language and Linguistic Theory* 7, 167–216.
- Cohn, A. & J.J. McCarthy (1994) Alignment and parallelism in Indonesian phonology. *Rutgers Optimality Archive* 25.
- Cruttenden, A. (1986) *Intonation*. Cambridge: Cambridge University Press.
- Dogil, G. (1999) The phonetic manifestation of word stress. In H. van der Hulst (ed) *Word prosodic systems in the languages of Europe*. Mouton de Gruyter: Berlin, 273–334.
- Ebing, E.F. (1997) *Form and function of pitch movements in Indonesian*. Leiden: CNWS Publications 55.
- Eefting, W.Z.F. (1991) The effect of information value and accentuation on the duration of Dutch words, syllables and segments. *Journal of the Acoustical Society of America* 89, 412–414.
- Fry, D.B. (1958) Experiments in the perception of stress. *Language and Speech* 1, 126–152.
- Goedemans, R.W.N., H.G. van der Hulst & E.A.M. Visch (1996) *Stress patterns of the world: Part 1, Background*. HIL Publications 2. Leiden: Holland Institute of Linguistics/The Hague: Holland Academic Graphics.
- Halim, A. (1974) *Intonation in relation to syntax in Bahasa Indonesia*. Jakarta: Djambatan.
- Hart, J. 't, R. Collier & A. Cohen (1990) *A perceptual study of intonation: An experimental-phonetic approach to speech perception*. Cambridge: Cambridge University Press.
- Hayes, B.P. (1995) *Metrical stress theory: Principles and case studies*. Chicago: The University of Chicago Press.
- 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, Leiden University, 1–26.
- Heuven, V.J. van & M.P.R. van den Broecke (1982) Interactions of spectral distribution and decay time in the auditory perception of duration and their relevance to speech. *Progress Report of the Institute of Phonetics Utrecht* 7:1, 21–41.
- Heuven, V.J. van & E. van Zanten (1997) Effects of substrate language on the localization and perceptual evaluation of pitch movements in Indonesian. In C. Odé & W. Stokhof

- (eds) *Proceedings of the 7th International Conference on Austronesian Linguistics*. Amsterdam/Atlanta: Rodopi, 63–79.
- Hollander, J.J. de (1984) *Pedoman bahasa dan sastra Melayu [Guide to the Malay language and literature]*. Jakarta: Balai Pustaka.
- Horne, E.C. (1961) *Beginning Javanese*. New Haven: Yale University Press.
- Ladd, R.D. (1996) *Intonational Phonology*. Cambridge: Cambridge University Press.
- Laksman, M. (1994) Location of stress in Indonesian words and sentences. 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, Leiden University, 108–139.
- Laksman, M. (1996) Pengaruh perbedaan sistem prosodi dua bahasa terhadap persepsi emosi [Influences of the difference in prosodic system of two languages as regards emotion perception]. Depok: Research Report Fakultas Sastra Universitas Indonesia.
- Laksman, M. & V.J. van Heuven (1999) *Perception of emotion in Indonesian by native and foreign listeners*. Paper presented at the 3rd Symposium on Malay/Indonesian Linguistics, Amsterdam, 24–25 August 1999.
- Langeweg, S.J. (1988) *The stress system of Dutch*. PhD dissertation, Leiden University.
- Lapoliwa, H. (1981) *A generative approach to the phonology of Bahasa Indonesia*. Pacific Linguistics D-36. Canberra: RSPAS.
- Moeliono, A.M. & S. Dardjowidjojo (1988) *Tata bahasa baku Bahasa Indonesia [Indonesian reference grammar]*. Jakarta: Balai Pustaka.
- Nababan, P.W.J. (1981) *A grammar of Toba Batak*. Pacific Linguistics D-37. Canberra: SPAS.
- Nooteboom, S.G. (1972) *Production and perception of vowel duration: A study of durational properties of vowels in Dutch*. PhD dissertation, Utrecht University.
- 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, Leiden University, 27–107.
- Percival, W.K. (1981) *A grammar of the urbanised Toba-Batak of Medan*. Pacific Linguistics B-76. Canberra: RSPAS.
- Pierrehumbert, J. & J. Hirschberg (1990) The meaning of intonational contours in the interpretation of discourse. In P.R. Cohen, J. Morgan & M.E. Pollack (eds) *Intentions in communication*. System development foundation benchmark series. Cambridge: The MIT press, 271–311.
- Poedjosoedarmo, S. (1982) *Javanese influence on Indonesian*. Pacific Linguistics D-38. Canberra: RSPAS.
- Prentice, J. (1994) Manado Malay: product and agent of language change. In T. Dutton & D.T. Tryon (eds) *Language Contact and Language Change in the Austronesian World*. Trends in Linguistics 7. Berlin/New York: Mouton de Gruyter, 411–441.
- Ras, J.J. (1982) *Inleiding tot het modern Javaans [Introduction to modern Javanese]*. M. Nijhoff: Den Haag.
- Samsuri (1971) *Ciri-ciri prosodi dalam kalimat Bahasa Indonesia [Prosodic features in the Indonesian sentence]*. Flores: Nusa Indah.
- Samsuri (1978). Fokus dan alat-alat pembentukannya dalam Bahasa Indonesia [Focus and the tools to form it in Indonesian]. Unpublished ms.
- Sluijter, A.M.C. & V.J. van Heuven (1996) Spectral balance as an acoustic correlate for linguistic stress. *Journal of the Acoustical Society of America* 100, 2471–2485.
- Steinhauer, H. (1980) On the history of Indonesian. In A.A. Barentsen, B.M. Groen & R. Sprenger (eds) *Studies in Slavic and general linguistics* 1. Amsterdam: Rodopi, 349–375.
- Steinhauer, H. (1994) The Indonesian language situation and linguistics: Prospects and possibilities. *Journal of the Humanities and Social Sciences of Southeast Asia and Oceania [Bijdragen tot de Taal-, Land- and Volkenkunde]* 150 IV, 755–784.

- Stoel, R.B. (2005) *Focus in Manado Malay: Grammar, particles, and intonation*. Leiden: CNWS Publications.
- Suparno (1993) *Konstruksi tema rema dalam Bahasa Indonesia lisan tidak resmi masyarakat kotamadya Malang [Thema-rhema construction in Indonesian as spoken in Malang]*. Jakarta: Pusat Pembinaan dan Pengembangan Bahasa.
- Teeuw, A. (1984) *Leerboek Bahasa Indonesia [Indonesian textbook]*. Groningen: Wolters-Noordhoff (4th ed).
- Woodrow, H. (1951) Time perception. In S. Stevens (ed) *Handbook of experimental phonology*. New York: Wiley & Sons.
- Zanten, E. van (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, Leiden University, 140–180.
- Zanten E. van, R.W.N. Goedemans & J.J.A. Pacilly (2003) The status of word stress in Indonesian. In J. van de Weijer, V.J. van Heuven & H.G. van der Hulst (eds) *The phonological spectrum, Volume II: Suprasegmental structure*. Amsterdam/Philadelphia: John Benjamins, 151–175.
- Zanten E. van & V.J. van Heuven (1997) Effects of word length and substrate language on the temporal organisation of words in Indonesian. In C. Odé & W. Stokhof (eds) *Proceedings of the 7th International Conference on Austronesian Linguistics*. Amsterdam/Atlanta: Rodopi, 201–213.
- Zanten, E. van & V.J. van Heuven (1998) Word stress in Indonesian: Its communicative relevance. *Journal of the Humanities and Social Sciences of Southeast Asia and Oceania [Bijdragen tot de Taal-, Land- and Volkenkunde]* 154, 129–147.
- Zanten, E. van & V.J. van Heuven (2004) Word stress in Indonesian: Fixed or free? *NUSA Linguistic Studies of Indonesian and other Languages in Indonesia* 53, 1–20.
- Zubkova, L.G. (1966) *Vokalizm Indonezijskogo jazyka [Indonesian vowels]*. PhD dissertation, Leningrad University.