Open Library of Humanities

The perception of word stress cues in Papuan Malay: A typological perspective and experimental investigation

Constantijn Kaland, Institute of Linguistics, University of Cologne, Germany, c.c.l.kaland@outlook.com

Analyses of word prosody have shown that in some Indonesian languages listeners do not make use of word stress cues. The outcomes have contributed to the conclusion that these languages do not have word stress. The current study revisits this conclusion and investigates to what extent speakers of Papuan Malay, a language of Eastern Indonesia, use suprasegmental stress cues to recognize words. Acoustically, this language exhibits predictable word level prominence patterns, which could facilitate word recognition. However, the literature lacks a crucial perceptual verification, and related languages among the Trade Malay varieties have been analyzed as stressless. This could be indicative of either regional variation or different criteria to diagnose word stress. To investigate this issue, the current study reviews the literature on which criteria were decisive to diagnose (the absence of) word stress in Indonesian and Trade Malay. An acoustic analysis and a gating task investigate the usefulness of Papuan Malay suprasegmental stress parameters for word recognition. Results show that Papuan Malay listeners are indeed able to use stress cues to identify words. The outcomes are discussed in a typological perspective to shed light on how production and perception studies contribute to stress diagnosis cross-linguistically.

1. Introduction

Word stress (henceforth also 'stress') refers to the presence of a single acoustically most prominent syllable in a word (Hyman, 2006). Languages vary on several aspects of word stress, such as the degree to which the location of the stressed syllable in the word can be predicted by rules, which acoustic correlates are used to make the stressed syllable prominent, and the extent to which stress patterns are useful for listeners. Although the presence of word stress is well established for some languages of the world, in other languages it is controversial (e.g., Indonesian, French, Korean). Examples of the latter group of languages are found in Indonesia, an area with a considerable amount of linguistic diversity. Apart from this diversity, the limited number of well documented languages and (consequently) empirical investigations have contributed to diverging claims over the last decades. To date, there are still new studies that counter earlier work in fundamental ways. The results from a small number of perception studies play a key role in (resolving) this controversy. To further add to this research, the goals of the current study are two-fold. First, this paper reviews the quantitative studies on Indonesian languages in order to provide structure in the arguments in favour or against the presence of word stress, and to reveal where more research is needed. Second, a perception experiment is carried out to further complete the current research on Papuan Malay word stress, in particular concerning its potential function in word identification.

The next sections summarize three key aspects of word stress across different languages before turning to how these aspects have been covered in Indonesian languages (Section 2); its acoustic realization (Section 1.1), its role in speech perception (Section 1.2), and potential communicative functions (Section 1.3). Section 1.4 summarizes the state of the art in word stress research and based on the current gaps in the literature defines the goals of the current study in more detail.

1.1. Acoustic realization

When a language has word stress, prosodic parameters such as duration, (spectrally weighed) intensity, vowel quality, and f0 contribute to a more or lesser degree to its acoustic realization (e.g., van Heuven, 2018 for an overview). These make stressed syllables stand out in acoustic prominence compared to unstressed syllables. It should be noted that there are two main reasons why the acoustic cues are not all equally important for stress realizations. First, some correlates have properties that make them intrinsically less suitable for signalling stress. For example, f0 has been claimed to be a primary correlate of phrase level prosody, with limited or only indirect contributions to the realization of word stress (Gordon, 2014; Gordon & Roettger, 2017; but see Vogel, Athanasopoulou, & Pincus, 2016). Second, languages differ in how they deploy the available correlates. The Functional Load Hypothesis (FLH; e.g., Hockett, 1955; Berinstein, 1979) explains why in some languages not all correlates are available to signal word stress.

That is, if one acoustic correlate serves a prosodic function other than word stress, such as f0 in lexical tone (Potisuk, Gandour, & Harper, 1996; Remijsen, 2002) or duration in final lengthening (McDonnell, 2016), this correlate is not or only limitedly available for word stress. At a more general level the FLH also holds true. It was found that in languages with a fixed position of the stressed syllable the acoustic differences between stressed and unstressed syllables are smaller (i.e., stress is weakly realized), than in languages with more positional variation of the stressed syllable (e.g., Dogil, 1999). Given the higher functional load on stress in the latter type of languages, the FLH can also be taken as an explanation of the strength differences of the acoustic correlates as observed crosslinguistically.

1.2. Perception

The extent to which listeners perceive the differences between stressed and unstressed syllables not only depends on their acoustic realization. It has been shown that the way word stress patterns are distributed in the lexicon determines how sensitive listeners are to stress cues (Peperkamp, Vendelin, & Dupoux, 2010). With a high number of exceptions to the phonological rules, there is a higher need to store prosodic information of the words in the lexicon. Spanish is an example of this type of language, and its listeners are therefore highly sensitive to the acoustic realization of stress in order to successfully identify words. It differs per language to which cues listeners attend exactly, with notable differences among closely related languages. It was found that vowel quality was the strongest perceptual cue to stress in English, Mandarin, and Russian (Chrabaszcz, Winn, Lin, & Idsardi, 2014), while duration was the strongest in Dutch and German (Sluijter, van Heuven, & Pacilly, 1997; Mengel, 2000). When the stressed syllable in the word is highly predictable by phonological rules, listeners do not need to store the stress information in their lexicon. This is the case in French, for which listeners have been reported as being 'stress-deaf' (Dupoux, Pallier, Sebastian, & Mehler, 1997). It should be noted that 'deafness' referring to the insensitivity of (French) listeners to word stress cues is a relative notion. Experiments also showed that French listeners do hear the acoustic cues, but they do not process them at an abstract phonological level as these cues do not have a function there (Dupoux et al., 1997). In Polish, a language with largely predictable stress and a small number of exceptions, listeners were found to be mainly sensitive to the exceptional stress pattern (Domahs, Knaus, Orzechowska, & Wiese, 2012). Similarly, Italian listeners recognized the most dominant stress pattern by default and showed sensitivity to particular cues associated with the non-default pattern (Sulpizio & McQueen, 2012).

1.3. Functions

Then, what types of functions could word stress have? In this brief overview, three main functions are distinguished; lexical contrast, word segmentation, and word identification (but see Cutler,

2005 for a more fine-grained overview). A well-known function is that stress parameters alone can distinguish one word from the other. This is often referred to as the lexically contrastive function of word stress, as illustrated with the Dutch word pair /ˈkaː.nɔn/ and /kaː.ˈnɔn/, translating to 'canon' (music) and 'cannon' (military), respectively. The number of minimal stress pairs in the lexicon differs per language and negatively correlates with the predictability of the stress patterns. With a high number of minimal stress pairs there is a low degree of predictability and vice versa. In languages without minimal stress pairs, patterns can be mostly predicted by phonological rules. Word stress can also help listeners to detect word boundaries and thus help them to segment the incoming speech signal (Cutler, 2012). This holds—unsurprisingly—for languages where stress has a largely fixed position, such as the initial syllable in Slovak (Hanulíková, McQueen, & Mitterer, 2010). However, in languages with more variability in the stress position, such as in English and Dutch, segmentation is facilitated as well (Cutler & Butterfield, 1992; Vroomen, van Zon, & de Gelder, 1996). The function of word stress central to the current study concerns the facilitation of word identification. Studies have shown that listeners can correctly discriminate words based on segmentally identical first syllables, which only differed in stress such as *admiral* and *admiration* in English (Cooper, Cutler, & Wales, 2002) and /ɔr.ˈkɛst/ (orchestra) and /ˈɔr.gəl/ (organ) in Dutch (van Heuven, 1988). The word identification function of stress has mainly been shown for English and Dutch; languages with more fixed stress positions lack experimental investigation (Cutler, 2005). Studies have shown how the facilitation effect on word identification originates from how stress is distributed in the lexicon. These lexical analyses investigated the occurrence of word embeddings, such as *bee*, which counts as embedding in *belay* and *beanie*, due to the fact that their initial syllables match for their segmental make-up. It appeared that the number of embedded words is reduced when taking into account stress information. Thus considering stress, *bee* (stressed) only counts as embedding in *beanie* (first syllable stressed) but no longer in *belay* (second syllable stressed). Lexically stored stress information could therefore help listeners to reject alternative word candidates that would otherwise be activated (and compete) during processing. The amount of reduction differed per language, with Spanish showing the largest degree of reduction, followed by Dutch and German, and English showing the smallest degree of reduction (Cutler, Norris, & Sebastián-Gallés, 2004; Cutler & Pasveer, 2006).

1.4. Current study

Generalizing from the brief overview of acoustic, perceptual, and functional aspects of word stress crosslinguistically, one factor stands out as determining all these aspects to some extent: the degree of stress predictability by phonological rules. The lower the degree of stress predictability, the more need for clear acoustic correlates and sensitive listeners, and the more important stress is for word processing. As the overview has shown, there are several fine-grained differences between languages. It appears that neither the ability to perceive stress patterns (e.g., French,

Polish) nor the presence of the lexically contrastive function alone can be sufficient to conclude whether a language has word stress or not. Studies on the perception or functions of word stress are generally limited to a small number of languages, suggesting that our knowledge of stress perception and stress processing might be far from generalizable to underdescribed languages. This stands in large contrast to the extensive list of languages for which acoustic correlates to stress were reported (see the overview in Gordon & Roettger, 2017). The state of the art in word stress research therefore allows for two observations. First, more research needs to be done to complement acoustic studies, in particular with regard to perception and the functions of word stress. This type of research, as already mentioned, lacks investigations of more diverse languages. Second and more theoretically, with a growing body of perception research on word stress, more attention should be given to its interpretation relative to the existing acoustic work. A central issue addressed in the current study concerns the extent to which perception studies contribute to the question of whether a language could be analyzed as a stress-language. A conservative answer would be that acoustic evidence is sufficient to show whether systematic alternations between stressed and unstressed syllables exist in the speech signal. However, perception research sheds important light on the many different types and the communicative functions of word stress attested in languages of the world, as already briefly illustrated in the overview above. The goal of the current study is therefore two-fold. The primary goal is illustrating how exactly word stress has been diagnosed in a small number of Indonesian languages (Section 2). The linguistic diversity in this area is large and there are diverging claims on the status of word stress in some of its languages. Perception research has played a key role in resolving some of this controversy. The secondary goal is to add more perception research using more diverse languages. This is done by an experimental investigation of the word identification function of stress in Papuan Malay (Section 3). Recent studies have found a number of indications for the existence of word stress in this language, crucially still lacking perceptual verification. The results of the word identification experiment are discussed by evaluating how they contribute to the diagnosis of word stress in Papuan Malay and crosslinguistically.

2. Experimental research on stress in Indonesian languages

The literature overview in this section focuses on languages of Indonesia for which quantitative analyses on word stress have been carried out relatively recently. Much of the quantitative research has experimentally tested impressionistic claims originating from a limited number of grammar sketches with sometimes little coverage of phonology. The latter type of work is therefore excluded from the current overview. The reader is referred to Odé (1994) for an extensive review of stress claims from mainly non-experimental work on Indonesian languages. The overview in this study is furthermore limited to nine languages in order to obtain a linguistically diverse and yet relevant impression of the research. Language diversity is revealed by the diverging

results in the studies and the different areas where the languages are spoken. The relevance of the selected languages for the current study is shown by the inclusion of several (Trade) Malay languages (Paauw, 2009). Papuan Malay, investigated in the current study, belongs to this language group. It is to date one of the few Indonesian languages for which word stress has been experimentally studied for its acoustic realization, perception, and function. Note that Papuan Malay is a regional language spoken in the provinces Papua and West-Papua, across different major urban areas (Kluge, 2017, p. xxiv), which are the home of many smaller local languages as well. The research conducted on this language so far involved participants from the Sarmi region (Kluge, 2017; Kaland, Himmelmann, & Kluge, 2019; Kaland & van Heuven, 2020), Manokwari (Kaland, 2019, 2020; Kaland & Baumann, 2020), and Sentani (this study). Although it should be noted that these regions exhibit dialectal differences, Papuan Malay is a distinct language due to its "structural uniqueness, limited or nonexistent inherent intelligibility, and the lack of shared ethnolinguistic identity with other Malay varieties" (Kluge, 2017, p. 9). In addition, studies on word stress conducted so far show a considerable amount of consistency in their results (see discussion in Section 2).

The overview furthermore includes Javanese accented Indonesian, Toba Batak accented Indonesian, Betawi Malay, Besemah, Ambonese Malay, Manado Malay, as well as two related standard languages: Indonesian (as spoken in Jakarta) and Malay (as spoken in Malaysia). It should be noted that the standard varieties included in this study are distinguished from regional varieties, which appeared a crucial distinction in earlier stress research (e.g., Goedemans & van Zanten, 2007). Thus, Javanese and Toba Batak accented Indonesian refer to Standard Indonesian as spoken by speakers with either Javanese or Toba Batak as their first language. Note that although Betawi Malay is spoken in Jakarta, it is distinguished from Standard (Jakartan) Indonesian because it is spoken by a homogeneous group (the Betawi; see also van Heuven, Roosman, & van Zanten, 2008). The overview furthermore includes Manado Malay, for which only minimal experimental data on stress was reported (Stoel, 2005; 2007). This language is still included to complement the other Trade Malay varieties (Ambonese Malay and Papuan Malay), and because its phonological description is elaborate and has been carried out systematically.

The literature overview is structured according to the three aspects of word stress discussed in Section 1 (acoustic realisation, perception, function). Not all aspects are covered in the available literature, given the current lack of research. Note that word stress in this overview is separated from phrase prosodic events such as pitch accents and boundary tones. Some studies have made claims on word stress based on phrase prosodic data only (Odé, 1994 for Standard Indonesian; Riesberg, Kalbertodt, Baumann, & Himmelmann, 2018; 2020 for Papuan Malay) and are therefore excluded from this overview. The importance of separating word-level from phrase-level prosody in Malayo-Polynesian languages of South East Asia has been pointed out in recent work (e.g., Kaufman & Himmelmann, in press, p. 12), in particular because the absence

of phrase-level pitch accents does not imply the absence of word stress (e.g., Gordon, 2014; Lindström & Remijsen, 2005). **Table 1** gives a schematic overview of the three stress aspects and lists additional information about the reported stress distribution and vowel inventories for each language. The stress distributions are notated following the coding system in the StressTyp database (Goedemans, Heinz, & van der Hulst, 2020; codes explained in van der Hulst et al., 2010). The distributions relevant for the current study concern penultimate (P) and ultimate (U) stress, with P being the dominant pattern on heavy syllables (P/U), or with some variability between the two positions (P;U), being either lexically contrastive (LEX) or not, or no main stress at all (NMS). Note that the StressTyp codes were derived from the available reports in the

| Language	ISO	V inventory	Distribution	Acoustic corr.	Perception	Function
Standard Indonesian	IND	/i,e,a,o,u,ə/	P(/U)	f0/i > d	preference for P if heavy (f0)	no minimal pairs, no word identification
Standard Malay	ZSM	/i,e,ɐ,o,u,ə/	NMS	–		
Javanese accented Indonesian	*JAV*	/i,e,a,o,u,ə/	NMS (P/U)	–	no preference for P or U (f0)	
Betawi Malay	BEW	/i,e,a,o,u,ə/	NMS	–		
Toba Batak accented Indonesian	*BBC*	/i,e,a,o,u/	P(;U LEX)	d/i > f0	preference for P (d/i/f0)	minimal pairs
Besemah	PSE	/i,a,u/	U	f0/i > d		
Ambonese Malay	ABS	/i,e,a,aᶜ,o,u/	NMS (P;U LEX)	– > s		no minimal pairs
Manado Malay	XMM	/i,e,a,o,u,ə/	(P;U LEX)			minimal pairs
Papuan Malay	PMY	/i,e,a,o,u/	P/P	d/v/s > f0	sensitive for U over P stress	no m. pairs, potential w. identification

Table 1: Schematic overview of stress research for nine Indonesian/Malay languages. *ISO* lists the ISO 639 three letter language codes (non-Malay L1 in italics). *Vowel inventories* are taken from the experimental studies (alternatively from the most recent literature source that provides explicit listing). *Distributions* are as reported in the reviewed literature and notated using StressTyp codes (Goedemans et al., 2020; van der Hulst et al., 2010), with claims unconfirmed by acoustic research between brackets. Acoustic correlates list d(uration), i(ntensity), v(owel quality), s(pectral tilt), and/or f0 if measured and found as correlate (> indicating strength differences). Shaded cells indicate lack of research.

reviewed literature, sometimes lacking precise descriptions of the stress distribution. **Figure 1** shows the geographical area for each language in the overview.

2.1. Acoustic realization

After a series of impressionistic claims on word stress in Standard Indonesian (Odé, 1994 for an overview), Laksman (1994) found f0 to be the strongest stress correlate, on the basis of data from a single speaker from Jakarta. This study concluded that stress always falls on the penultimate syllable (P) and that schwa in that position can be stressed as any other vowel. This claim does not appear to hold in basic descriptions of the IPA where schwa is part of the vowel inventory and causes stress to shift to the ultimate syllable (P/U; Soderberg & Olson, 2008).

Acoustic work on Standard Malay agrees on the lack of word stress, as duration, intensity, and f0 did not show effects of stress (Mohd Don, Knowles, & Yong, 2008; Wan, 2012). Given the lack of stable f0 alignment with syllables, Mohd Don et al. (2008) concluded that the syllable is not a relevant unit in the prosody of Malay (see Odé, 1994, p.63, for a similar conclusion based on the perception of phrase prosody in Standard Indonesian). It is important to note that neither of the studies on Standard Malay systematically compared alleged stressed syllables with alleged unstressed syllables, thus lacking the test of a specific stress hypothesis.
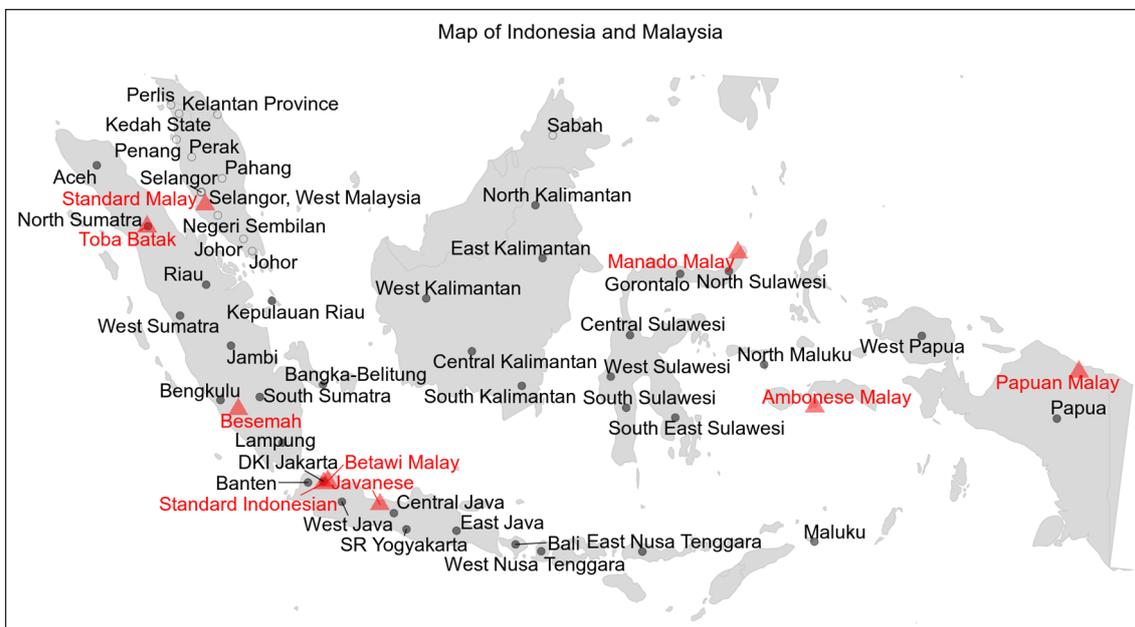
Figure 1: Map showing provinces and major governmental districts of Indonesia (filled dots) and Malaysia (open dots), and the geographical location (coordinates from Hammarström, Forkel, & Haspelmath, 2021) of the nine languages in the overview (red triangles).

For Javanese accented Indonesian, Goedemans and van Zanten (2007) found no stress related differences in duration and intensity that could constitute evidence for the P/U stress claim (duration results confirmed the ones in van Zanten & van Heuven, 1997). F0 showed a rise on the penultimate syllable, which was claimed to be a pre-boundary phrase prosodic phenomenon.

Betawi Malay (vowel inventory in Ikranagara, 1975) was investigated for f0 using words obtained in and out of focus in phrase final and phrase medial positions (van Heuven et al., 2008). No systematic f0 alignment with allegedly stressed penultimate syllables was found, but rather a large degree of variability in f0 movements. This result led to the conclusion that stress is absent in Betawi Malay. Note that correlates which have been reported traditionally as stronger indicators of word stress (duration, intensity, or spectral tilt) were not investigated for this language.

Support for penultimate stress (P) was found for Toba Batak accented Indonesian in duration and intensity (Goedemans & van Zanten, 2007, see also van Zanten & van Heuven, 1997). F0 was also measured and correlated with focus rather than with word stress (i.e., increased f0 in focus compared to non-focus).

Besemah (also Pasemah or Central Malay) is a Malay variety with a different stress distribution compared to the other languages in this overview. The strongest correlates were f0 and intensity, supporting the claim that stress is always ultimate in this language (McDonnell, 2016). Duration showed minimal effects due to its co-occurrence with final lengthening.

From the Trade Malay varieties, Ambonese Malay was shown to lack word stress (Maskikit-Essed & Gussenhoven, 2016), counter to the P;U (LEX) claim in van Minde (1997). A small effect of spectral tilt could be found, but no systematic acoustic differences due to stress in duration or f0. The re-analysis of Ambonese Malay stress also led to a different claim with regard to its vowel inventory. Given the lack of acoustic effects, alleged stress differences in van Minde (1997) were re-analyzed as segmental differences between /a/ and a slightly raised/centralized /a/, which was termed 'a-caduc' (a^c). Note that both /a/ variants have highly similar vowel quality.

The analysis of Manado Malay concerns an impressionistic interpretation of elicited material (Stoel, 2005) and led to the claim of regular penultimate stress. An elicitation task was carried out to obtain additional impressions on variable stress, i.e., words that were sometimes produced with P stress and sometimes with U stress (i.e., P;U, Stoel, 2005; p. 16).

A series of duration, intensity, spectral tilt, f0, and vowel quality measures taken from spontaneous Papuan Malay data revealed that duration, formant displacement (vowel quality), and spectral tilt were the strongest stress correlates (Kaland, 2019). F0 alignment correlated strongly with word stress, although the direction of the effects was different for P and U stress, which could be explained as originating from duration differences. Importantly, f0 excursion size was among the weakest stress correlates. Overall, these results confirmed the impressionistic

claim in Kluge (2017) and showed that Papuan Malay word stress falls on the penultimate by default, shifts to ultimate when /ɛ/ is in the penultimate, and remains penultimate when the final two syllables have /ɛ/ (P/P; see also Kaland et al., 2019).

Before turning to the perception and functions of word stress, some remarks should be made on the relevance of acoustic verification of stress claims. The literature has claimed that stress distributions in Indonesian/Malay languages follow a geographical division (Prentice, 1994). That is, in the Western part (Borneo, Sumatra, and the Malay Peninsula) P stress can be realized on schwa, e.g., /ˈdɛ.ŋar/ ('to hear'), whereas in the Eastern part (Java, Celebes, and the islands to the east thereof) schwa in the penultimate syllable causes stress to move to the ultimate syllable (P/U distribution), /dɛ.ˈŋar/ ('to hear'). It should be noted that the claim that schwa can be stressed in Standard Indonesian as spoken in Jakarta was supported in Laksman (1994). Geographical divisions have been widely adopted in typological accounts of prosody in Austronesian languages (e.g., Goedemans & van Zanten, 2014; Kaufman & Himmelmann, in press). The relatively small overview of languages discussed here already shows that geographical generalizations on word stress do not always hold. Acoustic measures supported the P/U distribution in Toba Batak accented Indonesian (Goedemans & van Zanten, 2007), even though it would classify as a Western variant according to a purely geographical division. And for Ambonese Malay, an Eastern variant, the traditional P/U claim did not find any acoustic support (Maskikit-Essed & Gussenhoven, 2016). Another closely related factor that has been claimed to affect the stress distributions in Malay languages is the presence of schwa in the vowel inventory. Trade Malay languages (except Larantuka Malay spoken in East-Flores) were reported to have lost schwa, which should have led to the development of word stress (Paauw, 2009). Originally, though, the lack of stress was hypothesized to be a feature of all Trade Malay varieties (Goedemans & van Zanten, 2014). Again, the overview presented here shows that the schwa-claim does not hold for Ambonese Malay, which has no schwa in the inventory and is analyzed as stressless (Maskikit-Essed & Gussenhoven, 2016). Manado Malay, claimed to have both stress and schwa (Stoel, 2005; 2007) also counters this generalization. Whether these languages are midway in the development of acquiring word stress remains—at the moment—fruitless speculation due to the lack of empirical data. Nevertheless, it is clear that more regional variation among these languages could be expected when more detailed prosodic investigations are carried out. It seems that the grouping of languages according to geographical location or according to shared traits as currently done in typological accounts are often too coarse grained for this area. This observation should not come as a surprise, given the vast archipelago where these languages are spoken and given the lack of research on prosody in this area. The importance of quantitative research on stress claims will furthermore show from the next sections on the perception and functions of word stress, which shed a new light on some of the acoustic results.

2.2. Perception

In van Zanten and van Heuven (2004) Indonesian recordings of three trisyllabic target words embedded in a carrier sentence were manipulated for f0. That is, the position and shape of the f0 movement on the target word (a rise-fall) was varied systematically between the three syllables, such that the different onsets of the rise and fall generate six positions and twelve shapes. Listeners were presented the manipulated stimuli and indicated which syllable they perceived as stressed. Results showed significantly more indications for the penultimate syllable (compared to other syllables) as being stressed for one of the three target words (*anaknya,* /a.ˈnak.ɲa/, 'his child'). This word was the only one among the target words with a heavy (closed) penultimate syllable, plausibly attracting acoustic prominence. Overall, most stress positions were acceptable to listeners, which was taken as an indication that stress is not bound to a specific syllable (i.e., free stress) and therefore not present in Indonesian. It should be noted that other cues than f0 were not tested perceptually in this study.

Javanese and Toba Batak accented Indonesian were both tested in the same perception study (Goedemans & van Zanten, 2007). Listeners chose the preferred word from a pair of acoustically manipulated words (comparison task) and rated the acceptability of these words (acceptability task). The words were embedded in a phrase. Manipulation concerned the position of the stressed syllable using the relevant acoustic correlates as found in each language, which was the alignment of the f0 fall for Javanese and f0, duration, and intensity for Toba Batak. All stimuli were presented to both Javanese and Toba Batak listeners. Results were similar for the comparison task and the acceptability task. They showed that Javanese listeners accepted different locations of the stressed syllable, in particular the final two syllables. Toba Batak listeners, however, had a clear preference for stress on the penultimate syllable. The results were interpreted as an indication for the presence of word stress in Toba Batak accented Indonesian and for the absence of word stress in Javanese. The study furthermore showed the importance of taking regional variation into account.

For Papuan Malay, listeners were presented a carrier phrase in which a bisyllabic target word was replaced by an acoustically manipulated sequence of hummed speech (Kaland, 2020). Duration, f0, intensity, and spectral tilt were manipulated such that either the first or second syllable stood out as more prominent. Listeners chose one out of two words that matched the manipulated sequence best. One word had alleged P stress, the other word had alleged U stress. Results showed overall low correctness scores, although f0 predicted the outcomes best. In a subsequent perception experiment with the same stimuli presented in isolation, the effect of f0 disappeared. Although no manipulated cue alone was strong enough to affect listeners' choices, the outcomes of both experiments showed overall higher correctness scores for U stress (above chance level) than for P stress. It was therefore concluded that listeners were sensitive mainly to the irregular stress pattern in Papuan Malay.

2.3. Function

The lexically contrastive function of word stress was reported for the Toba Batak language (Nababan, 1981, p.23) in minimal pairs such as /ˈi.tɔm/ ('black dye') and /i.ˈtɔm/ ('your brother/ sister'). These descriptions would fit a P;U (LEX) stress claim. Although minimal stress pairs have not been directly investigated in later acoustic studies on Toba Batak and Toba Batak accented Indonesian (Roosman, 2006; Goedemans & van Zanten, 2007), it is possible that this function exists in this variant as well, given the acoustic support for a P distribution in these studies. However, these results would need further support given that Goedemans and van Zanten (2007) investigated the production by one speaker only. The situation is different from the one for Standard Indonesian and Besemah, for which acoustic support mainly confirmed a P or U stress distribution respectively (Laksman, 1994; McDonnell, 2016). Thus without alternating patterns, no contrastive function can be attributed to word stress in these languages. As for the Trade Malay varieties in the overview, each of them shows a different state concerning the contrastive function. For Ambonese Malay minimal pairs were reported in van Minde (1997), but rejected in Maskikit-Essed and Gussenhoven (2016). For Manado Malay minimal pairs such as /ˈla.la/ (girl's name) and /la.ˈla/ ('tired'), as well as /ˈse.pi/ ('edge,' 'brink') and /se.ˈpi/ ('silent,' 'lonely') were reported without acoustic support (Stoel, 2005; Prentice, 1994). For Papuan Malay no minimal pairs were reported (Kluge, 2017).

As for word identification, a gating task (van Zanten & van Heuven, 1998) presented Indonesian listeners with (parts of) words that were identical in segmental content and supposedly different with respect to the location of the stressed syllable. That is, word triplets with alleged stress on the first, second, or third syllable (e.g., /ˈa.nak/, /a.ˈnak.ɲa/, /a.nak-ˈa.nak/) were presented in such a way that either only the first syllable (gate 1) or only the first and second syllable (gate 2) were audible. The listeners' task was to identify one out of the three words after hearing each gate embedded in a carrier sentence (presented in order of increasing gate length). The hypothesis tested in this experiment was whether stress cues in Standard Indonesian help listeners to identify words. If so, they should identify the words that matched in stress location with the gates, despite the ambiguity in their segmental content. Only one of the six Indonesian listeners correctly identified the words above chance level (only for gate 2), indicating that stress had no function in word identification. An acoustic analysis of the stimulus material for f0 and duration revealed that falling f0 in either the first or second gate predicted the alleged stress location best. The same gating task was done with Dutch listeners (with basic command of Indonesian) and revealed above chance level scores for correct word identification for nearly all participants. It was concluded that despite the presence of acoustic stress correlates in the signal, they were useless for word identification by Indonesian listeners. It should be noted that the Indonesian participants in this study (the speaker of the stimulus material and the listeners in the gating task) had different first languages (Balinese, Sundanese, and Javanese). Although

they were reported to be proficient speakers of Standard Indonesian, it remains unclear whether language background had an effect on the materials and/or results.

In an online word processing task (Kaland, 2020) it was found that Papuan Malay listeners identified bisyllabic words faster when the initial syllable was stressed (P stress) compared to when that syllable was unstressed (U stress). The effect could have been partially caused by a generic processing benefit for word-initial syllables, although the predominance of P stress in Papuan Malay makes it difficult to disentangle a generic effect from one that is exclusively related to word stress. Another study showed that the number of lexical embeddings (see also Section 1.3) is reduced when taking into account stress information in Papuan Malay (Kaland & van Heuven, 2020). This would mean that listeners can successfully reject activated word candidates during processing, if they make use of the stress cues. Although these studies show that Papuan Malay stress patterns have the potential of facilitating word identification, direct experimental evidence is still needed to corroborate these findings.

2.4. Current study: Implications and research questions

In the introduction, three main aspects of word stress were distinguished: acoustic realization, perception, and functions. The studies discussed in this section show that perception experiments have led to a better understanding of word stress in Indonesian and Malay than could be achieved on the basis of acoustic studies alone. In fact, for Standard Indonesian and Javanese accented Indonesian the conclusion that word stress is absent in these languages was mainly based on the results of perception studies. Due to the crucial refinement (or countering) of previous claims in the literature, these perception studies have been described as 'resolving' the discussion on Indonesian stress (e.g., van Zanten, Stoel, & Remijsen, 2010, p. 101). However, on the basis of the current overview a few more remarks should be made to illustrate the extent to which perception studies complement the acoustic studies in diagnosing word stress.

There is a crucial difference between the work on Standard Indonesian and Malay spoken by the Javanese in that studies on the former mainly investigated f0 in production and perception studies (Laksman, 1994; van Zanten & van Heuven, 1998; but duration is measured for the single speaker in van Zanten & van Heuven, 2004). Thus, it remains unclear what exactly the effect of duration, intensity, or any other possible correlate is in the production and perception of Standard Indonesian word stress, in particular because f0 could be a weak correlate (Section 1.1). The lack of acoustic investigations also casts doubt on what perception studies have resolved precisely, as the latter considered only f0 as well. An additional complicating factor is defining what counts as Standard Indonesian given the influence of regional languages (e.g., van Zanten & van Heuven, 1984; van Zanten & van Heuven, 1997). For Malay spoken by the Javanese, the state of the research is crucially different, as the conclusion that it lacks word stress was based

on the acoustic measures of multiple correlates (duration, intensity, and f0) and all of these were taken into account for the design of the perception experiment (Goedemans & van Zanten, 2007).

It should be noted that there is no reason to reject the conclusions on the absence of word stress in Standard Indonesian based on the overview presented here. However, it becomes clear that for perception studies to maximally contribute to the word stress question, these studies crucially depend on the available knowledge of the acoustic correlates. It can be questioned whether perception studies alone can be decisive to diagnose word stress. This is an important issue, in particular given the central question in many studies of whether or not the *language* has stress. Studies generally take acoustic analyses as decisive means to diagnose word stress (Gordon & Roettger, 2017). These acoustic studies are much more frequent than perception studies and have been taken as the basis for claims on the existence of word stress for many languages, including controversial as well as uncontroversial ones (Gordon & Roettger, 2017). Perception studies, on the other hand, primarily describe the (in)ability of listeners to perceive or meaningfully use the available cues. These studies reveal crucial differences in the (functional) contribution of stress distributions to speech perception across languages (e.g., Peperkamp et al., 2010), rather than a single decisive diagnostic of whether a certain stress distribution is present in the (produced) language or not.

The observations above ask for a recall of the perception results on Standard Indonesian (van Zanten & van Heuven, 1998). Listeners in this study did not pick up the acoustic cues to word stress that were present in the speech signal. This is not the only study that revealed a discrepancy between the correlates in the signal and the cues listeners attend to. This outcome reflects what has been shown in comparative studies on Dutch and English. English listeners detect stressed syllables mainly on the basis of (segmental) vowel quality differences, whereas Dutch listeners use mainly suprasegmental cues (duration, spectral tilt; Sluijter et al., 1997) to do so. Interestingly, suprasegmental stress correlates are generally present in the English speech signal. As a consequence, Dutch listeners were shown to outperform English listeners in the detection of English stressed syllables (Cooper et al., 2002; Cutler, Wales, Cooper, & Janssen, 2007). These studies show, therefore, that stress perception is more intricate than listeners processing whatever the speech signal has to offer. It rather seems that listeners attend to what they need to process the signal. In English, listeners attend mainly to vowel quality differences such as in the pair 'subject' (noun) and 'to subject' (verb), as these are generally sufficient to distinguish lexical meanings. Crucially, these studies also show that non-native listeners might pick up acoustic cues as stress, even though native speakers do not. This issue makes a strong case for doing empirical work on word stress (both production and perception), rather than reporting potentially misleading auditory impressions, shaped by mainly the researchers' native language (see Odé, 1994 and Kaland, 2019, for a discussion of this issue in Indonesian and Trade

Malay stress research respectively). And even though stress patterns in Indonesian and Malay might be present, they are generally highly predictable (**Table 1**) and therefore have little or no function in distinguishing lexical meanings (e.g., small number or no minimal stress pairs). The functional load of stress parameters is therefore small and listeners have little need to attend to specific acoustic cues. This is exactly the situation reported for Standard Indonesian, which was analyzed as a language without stress; Dutch listeners (with command of Standard Indonesian) were able to detect stress differences in the same stimuli for which native Standard Indonesian listeners did not (van Zanten & van Heuven, 1998).

A central issue in the current study is therefore whether a language could still be analyzed as having stress when listeners do not use its cues. This question applies in particular to Papuan Malay. The acoustic correlates of word stress in this language (Kaland, 2019) do not warrant that listeners use them, in particular given their low functional load. If Papuan Malay listeners appear to not use these cues for word identification, the challenging question remains whether the acoustic patterns in Kaland (2019) can still be interpreted as word stress or not (see also the discussion in Section 1.4). While this question cannot be answered at this stage, it is an important one that could challenge the crosslinguistic definition of word stress. The presence of correlates in Indonesian (van Zanten & van Heuven, 1998) was not enough to analyze this language as having word stress. The acoustic research on Papuan Malay, however, makes a stronger case for word stress as signalled by multiple cues simultaneously and showing at least some auditory sensitivity in listeners (Kaland, 2020). Although these results could be taken as indication that listeners will use them (H1), the possibility that they do not should be taken into account (H0) and might apply to other (related) languages. It should also be noted that posing the word stress question as a binary one has its limitations. That is, for investigating how meaningful word stress patterns are in a particular language, perception studies are indispensable (see also Section 1.2 and 2.2), and are likely to show different degrees of listener sensitivity for stress patterns (Peperkamp et al., 2010). Lexical analyses have shown that word stress information can help to disambiguate competing lexical candidates (Kaland & van Heuven, 2020; Kaland, Kluge, & van Heuven, 2021). Thus, whether listeners actually make use of the available stress cues if the segments are ambiguous remains to be tested directly. This research question is answered in the current study and the hypotheses are formulated in the following:

H0: Papuan Malay listeners do not use stress cues to identify words.
H1: Papuan Malay listeners use stress cues to identify words.

To investigate the research question, the current study reports an acoustic investigation, and gating task similar to the one in van Zanten and van Heuven (1998). The next section outlines the methodology of these investigations.

3. Methodology

In order to investigate the extent to which stress parameters contribute to word identification a forced choice gating task was carried out. In this task, participants identified one member from a pair of bisyllabic words. Each word in the pair was presented in gated fashion (Cotton & Grosjean, 1984) in final position in the Papuan Malay matrix phrase "Sa blum taw ko pu kata itu, kata [word]" (I don't yet know that word of yours, the word [word]). Although phrase-final words are affected by phrase prosodic phenomena such as boundary marking (Kaland & Baumann, 2020), the choice for these materials was motivated in two ways. First, co-articulatory cues from neighbouring words would have affected phrase-medial words more (both word edges) than phrase-final words (left word edge only), potentially reducing the quality of the gates. Second, the availability of phrase-medial words that would fit the design of the experiment was limited. Phrase-final words were generally more clearly articulated due to their position. The next section provides further details on the stimulus material, including the presence of coarticulation cues and the design.

3.1. Material preparation and design

The phrase-final words were taken including their original matrix phrase from a corpus of recordings (Kluge, Rumaropen, & Aweta, 2014). The recordings differed in recording quality, which were largely overcome in two phases of audio processing. First, noise reduction was applied per wave file in the corpus. This was done using the Noise Reduction function in Audacity (Audacity Team, 2019). Using this function, a profile of the background noise in a part of the recording where there was no speech (e.g., a silent pause) was generated. Thereafter, noise reduction based on that profile was applied across the entire wave file. Second, the intensity of all noise-reduced recordings was scaled using Praat (Boersma & Weenink, 2019) such that the average intensity in each recording was 70 dB SPL.

From the word list that constitutes the corpus (Kluge et al., 2014; Kluge, 2017) word pairs were created such that the words in a pair had identical segmental composition either the first or the second syllable (henceforth *matching syllable*). Crucially, the matching syllable was the stressed one in one word and the unstressed one in the other word. For example, the first syllables match in the pair ['bɛ.bɛk] (duck) and [bɛ.'ban] (burden) and the second syllables match in the pair ['am.pas] (waste) and [lɛ.'pas] (to free). In total, 32 word pairs were created, of which half had a matching first syllable and half had a matching second syllable (Appendix). Then, three gates per word were created (**Table 2**). The first gate consisted of the edge segment (ES) of the word; either the first segment (matching first syllable) or the last segment (matching second syllable). Note that the word pairs were chosen in such a way that the ES gate always consisted of a consonant, which reveals little or no acoustic stress cues compared to vowels. The second gate

Matching syllable	Gating direction	Gates		
		edge segment (ES)	matching syllable (MS)	entire word (EW)
First	Forward	[b]	[ˈbɛ]	[ˈbɛ.bɛk]
		[b]	[bɛ]	[bɛ.ˈban]
Second	Backward	[s]	[pas]	[ˈam.pas]
		[s]	[ˈpas]	[lɛ.ˈpas]

Table 2: Examples of the three gates from two word pairs that match either for their first or second syllable.

consisted of the matching syllable (MS, either the first or the second). The third gate consisted of the entire word (EW).

For the ES and MS gates, the part of the word that was not present in the gate was masked with white noise, such that the position of the gate and the word duration could still be identified. The white noise was added with 30% of the RMS amplitude of the original part of the word. This value was chosen to obtain a white noise intensity that fell well below the intensity of the original speech, such that the unmasked part of the word was still audible. The segment or syllable boundaries in the gates were drawn on the basis of auditory and visual (spectral) information. These boundaries were drawn such that there were only minimal co-articulatory cues in the gate. Nevertheless, the presence of co-articulatory cues could not be entirely avoided. In this respect, it should be noted that the segment directly adjacent to the matching syllable could have been audible in the MS-gate, causing a facilitation effect. To assess this potential facilitation effect, each word pair was categorized for whether or not the segment directly adjacent to the matching syllable was identical for both words in the pair (no co-articulatory cues; e.g., [ˈbɛ.**b**ɛk] versus [bɛ.ˈ**b**an]) or different (possible co-articulatory cues present). Note that the extent to which co-articulatory cues could actually facilitate recognition in the MS-gate varied depending on the type of articulatory difference between the segments directly adjacent to the MS-gate. For example, it can be expected that co-articulatory cues have a smaller effect, if at all present, in e.g., [ˈbɛn.**tʃ**ɔŋ] versus [bɛn.ˈ**t**ʊk] than in [ˈtʃɛ.**b**ɔ] versus [tʃɛ.ˈ**r**ɛj]); see the Appendix.

For each word that needed to be identified (target) participants chose which of the two gates (each corresponding to one member of a word pair) matched with the target. The rationale behind presenting both gates in a forced choice manner was to make the relevant acoustic differences between stressed and unstressed syllables available for participants (in the MS- and EW-gates). Forced choice gating tasks have been applied in previous research on word recognition (e.g., Davis, Marslen-Wilson, & Gaskell, 2002). It is expected that little to no stress information is stored lexically for languages with highly predictable stress (Peperkamp et al., 2010). Papuan Malay has

a small number of exceptional stress patterns that could increase the need for lexical storage of stress information. The availability of the relevant acoustic differences in the experiment ensured that participants, if able to use the cues, could map them readily onto their lexically stored knowledge of the target word.

Due to the forced-choice nature of the gating task participants had a 50% chance of selecting the correct member. Given the segmental material present in the three gates, it is expected that the ES-gates elicit a chance level response, as little to no unique acoustic cue is provided to identify the correct word. As for the MS-gate, participants are expected to score above chance level when they successfully use the acoustic cues to stress to identify the word and around chance level when they do not. As for the EW-gate, responses are expected to yield a 100% correct score as all (supra)segmental material is present to identify the correct word.

The three successive gates were presented in triplets in the order described above (ES–MS–EW). In this way, word pairs with a matching first syllable were forward gated, whereas word pairs with a matching second syllable were backward gated. Backward gating has been applied in previous research on word recognition (Salasoo & Pisoni, 1985; Wingfield, Goodglass, & Lindfield, 1997), indicating that listeners can successfully identify words based on their final segment(s). Both gating directions were applied in the current experiment because the Papuan Malay word pairs with a matching first syllable all consisted of /ɛ/ in that syllable. Thus, to be able to investigate the stress parameters provided by the other Papuan Malay vowels, word pairs that matched for the second syllable were included (Kluge, 2017). The other vowels occurring in the matching second syllables were /a/, /i/, and /u/. Although /ɔ/ also belongs to the Papuan Malay vowel inventory, it did not occur frequently enough in the word lists (Kluge, 2017) to include in word pairs in the current paradigm.

3.2. Acoustic measurements

The matching syllables of the 64 words that were selected for the word pairs (described above) were then acoustically measured in order to confirm that they differed in the presence of stress parameters. The measures were taken from the matching syllables only, in order to assure identical segmental composition of the stressed and unstressed syllables. Four acoustic correlates were measured in the first or second syllable of each word: duration, F0, spectral tilt (H1*-A2*), and vowel quality (F1/F2). These measures were chosen as they appeared the strongest correlates of Papuan Malay word stress in Kaland (2019). F0 did not correlate strongly with word stress in Kaland (2019), however showed effects of alignment with the stressed syllable (also Kaland & Baumann, 2020) and showed effects on stress perception in a phrase-context similar to the one in the current study (Kaland, 2020). Duration of the entire syllable was measured in milliseconds. Given that segmental composition was identical across stressed and unstressed syllables, no further conversion of the duration values was applied. F0 was measured in semitones as the

mean per syllable. Measures of vowel formants (F1/F2) and the first harmonic (H1) were taken to compute the spectral tilt and vowel quality values respectively. These were measured in a subinterval of the syllable. This subinterval was set around the intensity peak of the syllable, where stable formant trajectories were found. The boundaries of the subinterval were set on either side at the points where the intensity level (measured in dB) had dropped 4% relative to the peak intensity. Spectral magnitude correction (Hanson, 1997; p.113–115; Iseli, Shue, & Alwan, 2007) was applied to the spectral tilt measure. Note that the measurement methods were identical to the ones used in Kaland (2019). The vowel nucleus of the matching syllable was exclusively /ɛ/ in first syllables and almost exclusively /a/ in second syllables (/i/ and /u/ appeared once each in matching second syllables; see also the Appendix). For this reason, only the formant measures of /ɛ/ and /a/ in their respective syllable positions are reported here (**Table 3**).

Statistical tests on the acoustic measures were not performed due to the small number of words in the dataset ($N = 64$). It can still be observed that syllables are longer when stressed, with the second syllable being generally longer than the first syllable (**Table 3**). Mean F0 is higher in the stressed syllable than in the unstressed syllable with overall higher values in the first syllable than in the second syllable, possibly due to the declination effect (Breckenridge, 1977). The spectral tilt is more shallow (less intensity roll-off towards the higher frequencies) for stressed syllables than for unstressed syllables. As for the formant measures, the average position of the vowel in the acoustic space as measured by the actual formant values did not vary much due to stress (**Figure 2**). Note, however, that /a/ is somewhat more peripheral when stressed,

Acoustic measure	Stress position	σ1	σ2
Duration (ms)	Penultimate	**275.56 (38.64)**	414.48 (94.89)
	Ultimate	203.83 (50.84)	**512.08 (96.06)**
H1*-A2* (dB)	Penultimate	**3.13 (9.27)**	4.09 (8.67)
	Ultimate	6.99 (10.83)	**4.09 (6.94)**
F0 (ST)	Penultimate	**86.73 (1.19)**	83.62 (1.68)
	Ultimate	84.92 (0.84)	**85.46 (1.65)**
		/ɛ/	/a/
F1 (Bark)	Penultimate	**4.39 (0.59)**	5.63 (1.22)
	Ultimate	4.38 (1.10)	**5.98 (0.28)**
		/ɛ/	/a/
F2 (Bark)	Penultimate	**12.38 (0.98)**	10.41 (1.79)
	Ultimate	12.10 (1.22)	**9.93 (0.60)**

Table 3: Mean values for each of the acoustic measures and their standard deviations in each syllable position (σ1, σ2) and stress position (penultimate, ultimate). Values in bold indicate stressed syllables.
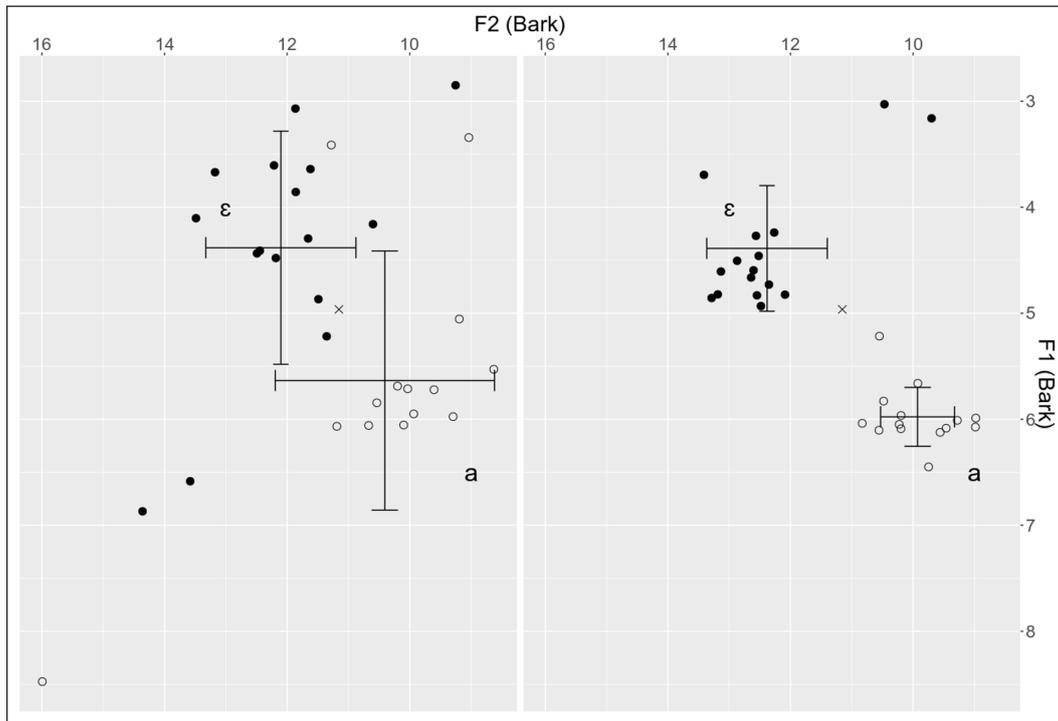
Figure 2: Individual values (dots) and mean F1 and F2 values (in Bark; with error bars) of unstressed (left panel) and stressed (right panel) /ɛ/ in first syllables (filled dots) and /a/ in second syllables (open dots). The cross marks the centre of the vowel space.

whereas the position of /ɛ/ remains almost identical in either stress condition. A clearer effect of stress can be observed in the standard deviations of the formant values of both vowels, indicating more target undershoot (larger SDs) in unstressed syllables (e.g., Lindblom, 1963). Standard deviations of the spectral tilt measures and those of the duration measures in first syllables indicated a similar effect. Final lengthening, applied here on two levels (word and phrase), is likely to cause larger SDs for duration measures in the second syllable, irrespective of stress.

The acoustic results therefore confirm that duration and spectral tilt are indicative of Papuan Malay word stress. A direct effect of stress on vowel quality was mainly found for /a/, as the target undershoot effects can be interpreted as an indirect effect of the shorter durations of unstressed syllables. In sum, the results of the acoustic analysis show that stress parameters are present in the selected words for the stimulus set of the gating experiment, which is further described in the following.

3.3. Participants

A total of 24 participants (21 female/3 male, Mean age: 23.40, age range: 18–35) completed the experiment. They were all native speakers of Papuan Malay without hearing problems living in

the Sentani region. Eleven participants reported to also speak Standard Indonesian in daily life (as native language and/or at home). All participants were remunerated for their participation.

3.4. Procedure

The task was designed and run online (PsyToolkit; Stoet, 2010; 2017). Participants completed the task behind a laptop computer using over-ear headphones (JBL PPT-450) to listen to the stimuli. They were seated in a quiet room with limited to no background noise. For each target word the computer screen displayed a written version of the matrix phrase including that target word (**Figure 3**). The target word was chosen randomly from the word pair. In this way, correct identification of the target required participants' attention to the presence of stress parameters in the gate (matching syllable was stressed in target) for approximately half of the gates and to the absence of stress parameters in the gate (matching syllable was unstressed in target) for the other

Figure 3: Example stimulus screen showing (from top to bottom): a percentage counter indicating the experiment progress, the matrix phrase and target word 'tandang,' a bar indicating which part of the word was gated (here: backward MS-gate 'dang'), two play buttons and corresponding selection boxes (here: gate played with the right play button selected), and a button with an arrow to continue to the next stimulus.

gates. Furthermore, the screen displayed two play buttons corresponding to the gates of each member in the word pair (**Figure 3**). The play buttons needed to be pressed at least once each before participants could make their choice. Each gate was presented auditorily in its original matrix phrase as recorded in Kluge et al. (2014). There was no maximum to the number of times participants could listen to a gate.

In addition, the screen displayed a visual indication of which gate in the triplet was presented (**Figure 3**). This was done by showing a rectangle directly underneath the target word. A red-colored filling of the rectangle starting from either the left side (matching first syllable) or the right side (matching second syllable) indicated which gate was presented; 10% filling (ES-gate), 50% filling (MS-gate), or 100% filling (EW-gate). After completing each triplet, participants received visual feedback on the number of correctly identified words (minimally zero and maximally three). For zero and one correctly identified word a sad-looking emoticon :-(was shown, for two correctly identified words a neutral-looking emoticon was shown :-| and for three correctly identified words a happy-looking emoticon :-) was shown. Emoticons were shown for 2.5 seconds after completing the third gate. The visual feedback (gate indicator and emoticon) were included to guide and to retain the auditory attention of the participant. A triplet of gates corresponded to one word pair, totalling 96 gates. Triplets were selected in random order and the random order was different for each participant to balance out potential learning effects. The order of the gates within each triplet was fixed as described above (ES-MS-EW). Each triplet occurred twice per experiment to balance whether participants needed to identify a target with penultimate stress or a target with ultimate stress. On the top right corner of the screen a counter indicated the percentage of the task participants had completed (**Figure 3**). The entire experiment lasted approximately 45 minutes. Responses were collected in the online experiment environment (PsyToolkit; Stoet, 2010; 2017) as correctness score, counting 1 for each correctly identified target and 0 for each incorrectly identified target.

3.5. Statistical analysis

Generalized linear mixed model (GLMM) analysis was performed using the 'lme4' package (Bates, Mächler, Bolker, & Walker, 2015) in R (R Core Team, R Studio Team, 2019) with *correctness score* as response, with *gate* (three levels: ES, MS, EW) and *gating direction* (two levels: forward, backward) each in interaction with *coarticulation cue* (two levels: present, absent) as predictors, and with random intercepts for participants and stimulus pair (the maximally converging model). Post-hoc pairwise comparisons using Tukey HSD test (Bonferroni corrected) were performed using the package 'multcomp' (Hothorn, Bretz, & Westfall, 2008) for the predictor *gate*.

Table 4 reports the mean correctness scores per gate, split by gating direction and the presence of coarticulation cues. **Table 5** reports the results of the two GLMMs and pairwise

coarticulation	direction	ES	MS	EW
present	forward	0.50 (0.50)	0.93 (0.26)	0.99 (0.11)
	backward	0.71 (0.45)	0.93 (0.26)	0.98 (0.14)
absent	forward	0.50 (0.50)	0.81 (0.39)	0.98 (0.16)
	backward	0.64 (0.48)	0.84 (0.36)	0.97 (0.16)

Table 4: Mean correctness scores (standard deviations) per gate (ES, MS, EW) split by the presence of coarticulation cues and gating direction.

analysis	Predictor	estimate	*SE*	*z*	*p*
GLMM	Intercept	0.08	0.22	0.37	0.71
	gate:ms	1.35	0.15	9.23	<0.001
	gate:ew	3.49	0.29	12.19	<0.001
	coarticulation:present	0.03	0.26	0.14	n.s.
	direction:bwd	0.52	0.27	1.94	0.05
	gate:ms * coarticulation:present	0.91	0.20	4.46	<0.001
	gate:ew * coarticulation:present	0.33	0.39	0.86	n.s.
	direction:bwd * coarticulation:present	0.14	0.34	0.42	n.s.
pairwise comparisons	gate:ms-es	1.35	0.15	9.23	<0.001
	gate:ew-es	3.50	0.29	12.19	<0.001
	gate:ew-ms	2.15	0.29	7.29	<0.001

Table 5: Results of the GLMM and pairwise comparisons on the correctness scores.

comparisons. **Figure 4** shows the mean correctness scores per gate, split by the presence of coarticulation cues in two bar charts.

4. Results

The results of the GLMM show significant effects of the predictor *gate*; the MS-gate elicited higher correctness scores than the ES-gate, and the EW-gate elicited higher responses than the ES-gate. The effect of *gating direction* showed a trend in that higher correctness scores were obtained for backward gating (i.e., when the second syllable was the matching syllable) than for forward gating. The interactions with *coarticulation* showed a significant effect for the MS-gate only in that higher correctness scores were obtained when coarticulation cues were present than when these were absent. The pairwise comparisons revealed that the correctness scores increased significantly with each successive gate.
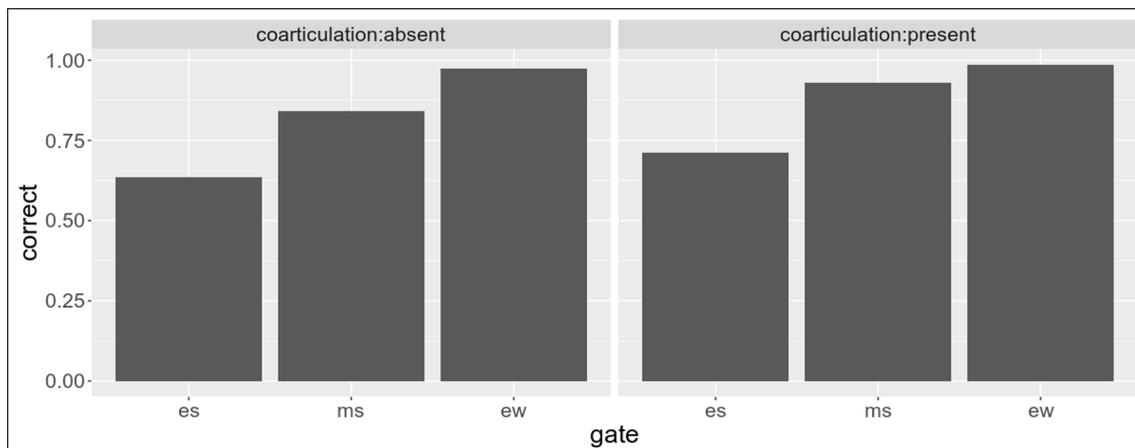
Figure 4: Mean correctness scores per gate (ES, MS, EW) split by the presence of coarticulation cues.

5. Discussion and conclusion

The gating experiment shows that Papuan Malay listeners used the suprasegmental stress cues to identify words. The presence of coarticulation cues increased listeners' correctness scores for the MS-gate, indicating that listeners' choices were not only affected by the stress cues. It should be noted that the correctness scores for the MS-gate was significantly lower than the correctness scores for the EW-gate (pairwise comparisons, **Table 5**). The latter scores were close to one, as expected, indicating that listeners identified the target correctly in virtually all cases in which they heard the entire word. The difference between the correctness scores of the MS-gate and the EW-gate reveals that, although stress cues facilitated word recognition, they were not sufficient to make listeners identify the target in nearly all cases (as in the EW-gate). This result can be explained by the fact that stress parameters in Papuan Malay are not lexically contrastive and have a low functional load (Section 2.3). Listeners are therefore used to primarily rely on segmental cues to identify words. Only for cases in which stress cues are the only ones to reveal word identity, e.g., word disambiguation such as enforced in the current task, listeners show that they are able to use them. This result crucially complements the lexical analysis in Kaland and van Heuven (2020) and Kaland et al. (2021), in which a representative subset of the Papuan Malay lexicon showed that stress information could have a facilitating effect on word disambiguation, in addition to the segmental information. Although this facilitation is smaller in Papuan Malay than in languages with less regular stress patterns such as Spanish, the results showed that the Papuan Malay lexicon leaves room for a facilitating role of stress to a similar extent as found for English (Kaland & van Heuven, 2020; Kaland et al., 2021). The results of the current study confirm that listeners are indeed able to rely on stress cues for word disambiguation.

It is furthermore interesting to observe that the differences between the gating directions showed a trend. **Table 4** reveals that gating direction differences were only found in the ES-gate, with higher correctness scores when the matching syllable was the second (backward gating) than when the matching syllable was the first (forward gating). This outcome indicates that listeners were better in identifying the target when listening to the final segment of that word than when listening to the first segment of that word. Identification scores were not expected to differ as a direct result of the way the gates were presented (Salasoo & Pisoni, 1985; Wingfield et al., 1997), and the current design allows for alternative explanations. That is, gating direction in the current study was confound with position of the matching syllable (and therefore with stress position) and with vowel identity. That is, all forward gates concerned matching first syllables with /ɛ/ as their nucleus, whereas most backward gates concerned matching second syllables with /a/ as their nucleus (Section 3.1). It should be recalled that in Kaland (2019) ultimate stress was realized with larger acoustic differences than penultimate stress. A possible explanation of the gating direction differences in the ES-gate could therefore lie in the stress position, which would match with the formant displacement differences observed in the stimuli (more displacement for /a/ than for /ɛ/; **Table 3**, **Figure 2**). It should also be noted that the forward ES-gate concerned a voiced consonant in 6/16 cases, whereas backward ES-gate concerned a voiced consonant in 11/16 cases. It is therefore likely that if stress realization had a 'spill-over' effect on the voiced edge segments (i.e., in the ES-gate), this effect would be larger for backward gated stimuli than for forward gated stimuli. Thus, in the current study, the acoustic cues to stress were likely to be more salient for listeners in the second (ultimate) syllable than in the first (penultimate) syllable (cf. **Table 3**).

The issue raised in Section 2 concerned the extent to which perception research contributes to the diagnosis of word stress in underresearched languages. The argument put forward on the basis of the literature overview is that perception studies contribute most to the word stress question when sufficient acoustic support is present (see also **Table 1**). The current study relies to a large extent on Kaland (2019), reporting evidence for word stress in Papuan Malay in multiple acoustic correlates. These results were supported by the small acoustic analysis on the stimuli used in the current study. Thus, just on the basis of the speech signal, Papuan Malay could be analyzed as a language with word stress. As the literature has shown for many languages, and in particular the ones spoken in Indonesia, perception studies provide a crucial insight into the functionality of the available stress parameters. On the basis of the current gating experiment it can therefore be concluded that Papuan Malay listeners are indeed able to use these cues when they don't have an alternative. Given that there is a role for Papuan Malay word stress parameters to disambiguate embedded words (Kaland & van Heuven, 2020; Kaland et al., 2021), listeners have an incentive to use them and, given the current results, will do so. It should be noted that word disambiguation does not concern a problem Papuan Malay listeners face

regularly. Embeddings are not frequent and context provides additional facilitation. As already discussed in Section 2.4, stress in Papuan Malay has a low functional load. This suggests that if there would be no stress cues in the Papuan Malay speech signal, it is unlikely that listeners would face perception difficulties that disrupt the communication process.

The above observations bring back the question raised in Section 2.4: Could a language have word stress when listeners don't need to use its acoustic cues? On the basis of both Kaland (2019) and the current study the answer is undoubtedly affirmative for Papuan Malay. The speech signal provides multiple stress correlates and listeners will use them to their advantage in the absence of other cues to word identification. As such, the low functional load of word stress does not justify the conclusion that this language lacks word stress. Rather, it appears to be a type of stress that requires controlled investigations to be revealed. This makes the outcomes of the current study crucially different from the ones found in a similar gating task for Standard Indonesian (van Zanten & van Heuven, 1998). In that study, the stimuli also provided acoustic cues to stress, although Indonesian listeners were not able to use them functionally. The gating task in the current study was different in that participants matched one of two auditory stimuli to a single written target word, whereas in van Zanten and van Heuven (1998) participants listened to a single auditory stimulus and matched it with one of three written words. Thus, in the current study participants were presented the crucial acoustic contrast between stressed and unstressed syllables for each choice they needed to make (see also Section 3.4). This could have made the stress cues more salient in the current study than in van Zanten and van Heuven (1998).

Apart from methodological differences, it is important to point out the diversity among Indonesian languages (see also Section 2). The Trade Malay languages alone reveal two important reasons why the language diversity in Indonesia requires careful investigation and reticence in assuming overall similarities in prosodic structure. First, many empirical investigations are still lacking and allow neither firm conclusions nor generalizations on which features are unique and which are shared among all Trade Malay languages (**Table 1**). Second, the limited empirical work has already hinted at the existence of different prosodic structures among closely related languages (cf. Ambonese Malay). In this respect it is also important to reconsider that the low functional load of stress gave rise to analyses that attribute word level acoustic differences to different phonological domains. Two examples illustrate this type of analyses. For Austronesian languages in general, the alleged word stress patterns have been explained as reflexes of (intermediate) phrase prosody (Goedemans & van Zanten, 2014). For Ambonese Malay in particular, re-analysis of its vowel inventory rendered the minimal differences in word pairs as segmental (Maskikit-Essed & Gussenhoven, 2016) instead of suprasegmental (van Minde, 1997). In impressionistic work on other Trade Malay languages (see **Table 1** in Kaland, 2019 for an overview), the presence of minimal stress pairs has also been taken as the central argument in (binary) stress diagnoses. The literature has shown that the lexically contrastive function is not the only way in which word

stress facilitates word identification (Cutler, 2012 and Section 1.3). Papuan Malay appears to be a language in which stress parameters play a facilitative role for word identification despite the absence of minimal stress pairs. This observation reconfirms what perception studies have already shown: It is more fruitful to consider the presence and functionality of word stress on a scale, with the predictability of stress distributions as its core determinant (Peperkamp et al., 2010 and Section 1), rather than as a strictly categorical feature. In this respect, (re-)analysis of other Trade Malay and Indonesian languages is needed to understand which aspects indicate typological diversity and which aspects indicate typological homogeneity.

Additional File

The additional file for this article can be found as follows:

- **Appendix**: Overview of the word pairs, indicating the matching syllable (σ), stress location (penultimate, ultimate), English gloss, and the possible presence (yes/no) of co-arciculatory cues in the matching syllable due to non-identical segments directly adjacent to the matching syllable. DOI: https://doi.org/10.16995/labphon.6447.s1

Acknowledgements

Competing Interests

The author has no competing interests to declare.

References

Audacity Team. (2019). Audacity: Free, open source, cross-platform audio software for multi-track recording and editing. (2.2.1) [Computer software]. https://www.audacityteam.org/

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. DOI: https://doi.org/10.18637/jss.v067.i01

Berinstein, A. E. (1979). A Cross-linguistic Study on the Perception and Production of Stress (MA Thesis, Unviersity of California). https://escholarship.org/uc/item/0t0699hc#author

Boersma, P., & Weenink, D. (2019). Praat: Doing Phonetics by Computer (6.0.56) [Computer software]. http://www.praat.org/

Breckenridge, J. (1977). The declination effect. *The Journal of the Acoustical Society of America*, *61*(S1), S90–S90. DOI: https://doi.org/10.1121/1.2015971

Chrabaszcz, A., Winn, M., Lin, C. Y., & Idsardi, W. J. (2014). Acoustic Cues to Perception of Word Stress by English, Mandarin, and Russian Speakers. *Journal of Speech, Language, and Hearing Research*, *57*(4), 1468–1479. DOI: https://doi.org/10.1044/2014_JSLHR-L-13-0279

Cooper, N., Cutler, A., & Wales, R. (2002). Constraints of Lexical Stress on Lexical Access in English: Evidence from Native and Non-native Listeners. *Language and Speech*, *45*(3), 207–228. DOI: https://doi.org/10.1177/00238309020450030101

Cotton, S., & Grosjean, F. (1984). The gating paradigm: A comparison of successive and individual presentation formats. *Perception & Psychophysics*, *35*(1), 41–48. DOI: https://doi.org/10.3758/BF03205923

Cutler, A. (2005). Lexical Stress. In D. B. Pisoni & R. E. Remez (Eds.), *The Handbook of Speech Perception* (pp. 264–289). Blackwell Publishing Ltd. DOI: https://doi.org/10.1002/9780470757024.ch11

Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. The MIT Press. DOI: https://doi.org/10.7551/mitpress/9012.001.0001

Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation: Evidence from juncture misperception. *Journal of Memory and Language, 31*(2), 218–236. DOI: https://doi.org/10.1016/0749-596X(92)90012-M

Cutler, A., Norris, D., & Sebastián-Gallés, N. (2004). Phonemic repertoire and similarity within the vocabulary. In S. H. Kin & M. J. Bae (Eds.), *Proceedings of the 8th International Conference on Spoken Language Processing (Interspeech 2004-ICSLP)* (pp. 65–68). Sunjijn Printing Co. DOI: https://doi.org/10.21437/Interspeech.2004-61

Cutler, A., & Pasveer, D. (2006). Explaining cross-linguistic differences in effects of lexical stress on spoken-word recognition. In R. Hoffmann & H. Mixdorff (Eds.), *Speech Prosody 2006* (Vol. 40). TUD press.

Cutler, A., Wales, R., Cooper, N., & Janssen, J. H. (2007). Dutch listeners' use of suprasegmental cues to English stress. In J. Trouvain & W. J. Barry (Eds.), *Proceedings of the 16th International Congress of the Phonetics Sciences* (pp. 1913–1916).

Davis, M. H., Marslen-Wilson, W. D., & Gaskell, M. G. (2002). Leading up the lexical garden path: Segmentation and ambiguity in spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 218–244. DOI: https://doi.org/10.1037/0096-1523.28.1.218

Dogil, G. (1999). The phonetic manifestation of word stress in Lithuanian, Polish, German, and Spanish. In H. van der Hulst (Ed.), *Word Prosodic Systems in the Languages of Europe* (pp. 273–311). De Gruyter. DOI: https://doi.org/10.1515/9783110197082.1.273

Domahs, U., Knaus, J., Orzechowska, P., & Wiese, R. (2012). Stress "deafness" in a language with fixed word stress: An ERP Study on Polish. *Frontiers in Psychology, 3*. DOI: https://doi.org/10.3389/fpsyg.2012.00439

Dupoux, E., Pallier, C., Sebastian, N., & Mehler, J. (1997). A Destressing "Deafness" in French? *Journal of Memory and Language, 36*(3), 406–421. DOI: https://doi.org/10.1006/jmla.1996.2500

Goedemans, R., & van Zanten, E. (2007). Stress and accent in Indonesian. In V. J. van Heuven & E. van Zanten (Eds.), *Prosody in Indonesian Languages (LOT Occasional Series)* (Vol. 9, pp. 35–62). LOT, Netherlands Graduate School of Linguistics. https://dspace.library.uu.nl/handle/1874/296769

Goedemans, R., & van Zanten, E. (2014). No Stress Typology. In J. Caspers, Y. Chen, W. Heeren, J. Pacilly, N. O. Schiller, & E. van Zanten (Eds.), *Above and Beyond the Segments* (pp. 83–95). John Benjamins. DOI: https://doi.org/10.1075/z.189.07goe

Goedemans, R. W. N., Heinz, J., & van der Hulst, H. (2020). *StressTyp2: Home Page.* http://st2.ullet.net/

Gordon, M. (2014). Disentangling stress and pitch-accent: A typology of prominence at different prosodic levels. In H. van der Hulst (Ed.), *Word Stress* (pp. 83–118). Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139600408.005

Gordon, M., & Roettger, T. (2017). Acoustic correlates of word stress: A cross-linguistic survey. *Linguistics Vanguard, 3*(1). DOI: https://doi.org/10.1515/lingvan-2017-0007

Hammarström, H., Forkel, R., & Haspelmath, M. (Eds.). (2021). Glottolog 4.4.7. Max Planck Institute for the Science of Human History. https://glottolog.org/

Hanson, H. M. (1997). Glottal characteristics of female speakers: Acoustic correlates. *The Journal of the Acoustical Society of America, 101*(1), 466–481. DOI: https://doi.org/10.1121/1.417991

Hanulíková, A., McQueen, J. M., & Mitterer, H. (2010). Possible words and fixed stress in the segmentation of Slovak speech. *Quarterly Journal of Experimental Psychology, 63*(3), 555–579. DOI: https://doi.org/10.1080/17470210903038958

Hockett, C. F. (1955). *A manual of phonology*. University of Chicago Press.

Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous Inference in General Parametric Models. *Biometrical Journal, 50*(3), 346–363. DOI: https://doi.org/10.1002/bimj.200810425

Hyman, L. M. (2006). Word-prosodic typology. *Phonology, 23*(02), 225–257. DOI: https://doi.org/10.1017/S0952675706000893

Ikranagara, K. (1975). *Melayu Betawi grammar* (Doctoral dissertation, University of Hawai'i at Mānoa). DOI: https://doi.org/10.1515/ijsl.1975.5.93

Iseli, M., Shue, Y.-L., & Alwan, A. (2007). Age, sex, and vowel dependencies of acoustic measures related to the voice source. *The Journal of the Acoustical Society of America, 121*(4), 2283–2295. DOI: https://doi.org/10.1121/1.2697522

Kaland, C. (2019). Acoustic correlates of word stress in Papuan Malay. *Journal of Phonetics, 74*, 55–74. DOI: https://doi.org/10.1016/j.wocn.2019.02.003

Kaland, C. (2020). Offline and online processing of acoustic cues to word stress in Papuan Malay. *The Journal of the Acoustical Society of America, 147*(2), 731–747. DOI: https://doi.org/10.1121/10.0000578

Kaland, C., & Baumann, S. (2020). Demarcating and highlighting in Papuan Malay phrase prosody. *The Journal of the Acoustical Society of America, 147*(4), 2974–2988. DOI: https://doi.org/10.1121/10.0001008

Kaland, C., Himmelmann, N. P., & Kluge, A. (2019). Stress predictors in a Papuan Malay random forest. In Sasha Calhoun, Paola Escudero, & Marija Tabain (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 2871–2875). Australasian Speech Science and Technology Association Inc.

Kaland, C., Kluge, A., & van Heuven, V. J. (2021). Lexical analyses of the function and phonology of Papuan Malay word stress. *Phonetica, 78*(2), 141–168. DOI: https://doi.org/10.1515/phon-2021-2003

Kaland, C., & van Heuven, V. J. (2020). Papuan Malay word stress reduces lexical alternatives. *Proceedings of the 10th International Conference on Speech Prosody 2020*, 454–458. DOI: https://doi.org/10.21437/SpeechProsody.2020-93

Kaufman, D., & Himmelmann, N. P. (n.d.). Suprasegmental Phonology. In A. Adelaar & A. Schapper (Eds.), *The Oxford Guide to the Western Austronesian Languages*. Oxford University Press.

Kluge, A. (2017). *A grammar of Papuan Malay*. Language Science Press. DOI: https://doi.org/10.17169/langsci.b78.35

Kluge, A., Rumaropen, B., & Aweta, L. (2014). *Papuan Malay data—Word list*. Dallas, TX: SIL International. https://www.sil.org/resources/archives/59649

Laksman, M. (1994). Location of stress in Indonesian words and sentences. In C. Odé, V. van Heuven, & E. van Zanten (Eds.), *Experimental studies of Indonesian prosody* (pp. 108–139). Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit Leiden.

Lindblom, B. (1963). Spectrographic Study of Vowel Reduction. *The Journal of the Acoustical Society of America, 35*(11), 1773–1781. DOI: https://doi.org/10.1121/1.1918816

Lindström, E., & Remijsen, B. (2005). Aspects of the prosody of Kuot, a language where intonation ignores stress. *Linguistics, 43*(4). DOI: https://doi.org/10.1515/ling.2005.43.4.839

Maskikit-Essed, R., & Gussenhoven, C. (2016). No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology, 33*(2), 353–389. DOI: https://doi.org/10.1017/S0952675716000154

McDonnell, B. (2016). Acoustic correlates of stress in Besemah. *NUSA: Linguistic Studies of Languages in and around Indonesia, 60*, 1–28. https://doi.org/10.15026/87442

Mengel, A. (2000). *Deutscher Wortakzent: Symbole, Signale*. Phorm-Verl. [u.a.].

Mohd Don, Z., Knowles, G., & Yong, J. (2008). How Words can be Misleading: A Study of Syllable Timing and "Stress" in Malay. *The Linguistic Journal, 3*.

Nababan, P. W. J. (1981). *A grammar of Toba-Batak*. Dept. of Linguistics, Research School of Pacific Studies, Australian Natl. Univ. https://pure.mpg.de/pubman/faces/ViewItemOverviewPage.jsp?itemId=item_401741

Odé, C. (1994). On the perception of prominence in Indonesian. In C. Odé, V. van Heuven, & E. van Zanten (Eds.), *Experimental studies of Indonesian prosody* (pp. 27–107). Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, Rijksuniversiteit Leiden.

Paauw, S. H. (2009). *The Malay contact varieties of eastern Indonesia: A typological comparison*. http://ubir.buffalo.edu/xmlui/handle/10477/45490

Peperkamp, S., Vendelin, I., & Dupoux, E. (2010). Perception of predictable stress: A cross-linguistic investigation. *Journal of Phonetics, 38*(3), 422–430. DOI: https://doi.org/10.1016/j.wocn.2010.04.001

Potisuk, S., Gandour, J., & Harper, M. P. (1996). Acoustic Correlates of Stress in Thai. *Phonetica, 53*(4), 200–220. DOI: https://doi.org/10.1159/000262201

Prentice, J. (1994). Manado Malay: Product and agent of language change. In T. Dutton & D. T. Tryon (Eds.), *Language Contact and Change in the Austronesian World*. De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110883091.411

R Core Team. (2019). R: The R project for statistical computing (3.5.3) [Computer software]. https://www.r-project.org/

R Studio Team. (2019). RStudio: Integrated Development for R (1.0.143) [Computer software]. RStudio, Inc. https://www.rstudio.com/

Remijsen, B. (2002). Lexically contrastive stress accent and lexical tone in Ma'ya. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7*. De Gruyter. DOI: https://doi.org/10.1515/9783110197105.585

Riesberg, S., Kalbertodt, J., Baumann, S., & Himmelmann, N. P. (2018). *On the Perception of Prosodic Prominences and Boundaries in Papuan Malay*. Zenodo. DOI: https://doi.org/10.5281/ZENODO.1402559

Riesberg, S., Kalbertodt, J., Baumann, S., & Himmelmann, N. P. (2020). Using Rapid Prosody Transcription to probe little-known prosodic systems: The case of Papuan Malay. *Laboratory Phonology: Journal of the Association for Laboratory Phonology, 11*(1), 8. DOI: https://doi.org/10.5334/labphon.192

Roosman, L. (2006). *Phonetic experiments on the word and sentence prosody of Betawi Malay and Toba Batak*. LOT.

Salasoo, A., & Pisoni, D. B. (1985). Interaction of knowledge sources in spoken word identification. *Journal of Memory and Language, 24*(2), 210–231. DOI: https://doi.org/10.1016/0749-596X(85)90025-7

Sluijter, A. M. C., van Heuven, V. J., & Pacilly, J. J. A. (1997). Spectral balance as a cue in the perception of linguistic stress. *The Journal of the Acoustical Society of America, 101*(1), 503–513. DOI: https://doi.org/10.1121/1.417994

Soderberg, C. D., & Olson, K. S. (2008). Indonesian. *Journal of the International Phonetic Association, 38*(2), 209–213. DOI: https://doi.org/10.1017/S0025100308003320

Stoel, R. B. (2005). *Focus in Manado Malay: Grammar, Particles, and Intonation*. Leiden University Press.

Stoel, R. B. (2007). The intonation of Manado Malay. In van Heuven, Vincent J. & van Zanten, E (Eds.), *Prosody in Indonesian Languages* (Vol. 9, pp. 117–150). LOT, Netherlands Graduate School of Linguistics.

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104. DOI: https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology, 44*(1), 24–31. DOI: https://doi.org/10.1177/0098628316677643

Sulpizio, S., & McQueen, J. M. (2012). Italians use abstract knowledge about lexical stress during spoken-word recognition. *Journal of Memory and Language, 66*(1), 177–193. DOI: https://doi.org/10.1016/j.jml.2011.08.001

van der Hulst, H., Goedemans, R., & Zanten, E. van (Eds.). (2010). *A Survey of Word Accentual Patterns in the Languages of the World*. De Gruyter Mouton. DOI: https://doi.org/10.1515/9783110198966

van Heuven, V. (1988). Effects of stress and accent on the human recognition of word fragments in spoken context: Gating and shadowing. W.A. Ainsworth & J.N. Holmes (Eds.), Proceedings of the 7th FASE/Speech-88 Symposium, 811–818 (1988).

van Heuven, V. J. (2018). Acoustic Correlates and Perceptual Cues of Word and Sentence Stress: Towards a Cross-Linguistic Perspective. In R. Goedemans, J. Heinz, & H. Van der Hulst (Eds.), *The Study of Word Stress and Accent* (1st ed., pp. 15–59). Cambridge University Press. DOI: https://doi.org/10.1017/9781316683101.002

van Heuven, V. J., Roosman, L., & van Zanten, E. (2008). Betawi Malay word prosody. *Lingua, 118*(9), 1271–1287. DOI: https://doi.org/10.1016/j.lingua.2007.09.005

van Heuven, V. J., & van Zanten, E. (1997). Effects of substrate language on the localization and perceptual evaluation of pitch movements in Indonesian. *Proceedings of the 7th International Conference on Austronesian Linguistics*, 63–80. https://scholarlypublications.universiteitleiden.nl/handle/1887/63067

van Minde, D. (1997). *Malayu Ambong: Phonology, morphology, syntax*. Research School CNWS.

van Zanten, E., Stoel, R., & Remijsen, B. (2010). Stress types in Austronesian languages. In H. Van der Hulst (Ed.), *A Survey of Word Accentual Patterns in the Languages of the World* (pp. 87–112). De Gruyter.

van Zanten, E., & van Heuven, V. J. (1984). The Indonesian vowels as pronounced and perceived by Toba Batak, Sundanese and Javanese speakers. *Bijdragen Tot de Taal-, Land- En Volkenkunde / Journal of the Humanities and Social Sciences of Southeast Asia, 140*(4), 497–521. DOI: https://doi.org/10.1163/22134379-90003411

van Zanten, E., & van Heuven, V. (1998). Word stress in Indonesian; Its communicative relevance. *Bijdragen Tot de Taal-, Land- En Volkenkunde/Journal of the Humanities and Social Sciences of Southeast Asia and Oceania, 154*. DOI: https://doi.org/10.1163/22134379-90003908

van Zanten, E., & van Heuven, V. J. (2004). Word stress in Indonesian: Fixed or free? *NUSA Linguistic Studies of Indonesian and Other Languages in Indonesia, 53*, 1–20.

Vogel, I., Athanasopoulou, A., & Pincus, N. (2016). Prominence, Contrast, and the Functional Load Hypothesis: An Acoustic Investigation. In J. Heinz, R. Goedemans, & H. van der Hulst (Eds.), *Dimensions of Phonological Stress* (pp. 123–167). Cambridge University Press. DOI: https://doi.org/10.1017/9781316212745.006

Vroomen, J., van Zon, M., & de Gelder, B. (1996). Cues to speech segmentation: Evidence from juncture misperceptions and word spotting. *Memory & Cognition, 24*(6), 744–755. DOI: https://doi.org/10.3758/BF03201099

Wan, A. (2012). *Instrumental phonetic study of the rhythm of Malay* (Doctoral dissertation, Newcastle University). http://theses.ncl.ac.uk/jspui/handle/10443/1682

Wingfield, A., Goodglass, H., & Lindfield, K. C. (1997). Word recognition from acoustic onsets and acoustic offsets: Effects of cohort size and syllabic stress. *Applied Psycholinguistics, 18*(1), 85–100. DOI: https://doi.org/10.1017/S0142716400009887