Phonology **33** (2016) 353–389.
doi:10.1017/S0952675716000154

*No stress, no pitch accent, no prosodic focus: the case of Ambonese Malay**

Raechel Maskikit-Essed
Carlos Gussenhoven
Radboud University Nijmegen

Varieties of Malay, including Indonesian, have been variously described as having word stress on the penultimate syllable, as having variable word stress and as having a phrase-final pitch accent without word stress. In Ambonese Malay, the alignment of sentence-final pitch peaks fails to support the existence of either word stress or phrase-final pitch accents. Also, the shape of its pitch peaks fails to vary systematically with the information status of the phrase-final word. The two intonation melodies of the language include phrase-final boundary-tone complexes which do not associate with any syllables. The declarative rise-fall would appear to be timed so as to occur within the last word of the sentence. Minimal stress pairs presented in earlier descriptions show a contrast between /a/ and a segmentally distinct weak /ă/, a contrast that also appears in positions that have not been claimed to have stress. A preliminary phonological analysis concludes the account.

* E-mail: RACHELMASKIKIT@YAHOO.COM, C.GUSSENHOVEN@LET.RU.NL.
Many people have assisted us with their expertise at various stages of the work. Aone van Engelenhoven, Joop Kerkhoff, Rosina Lekawael, Thommy Leninduan, Alexander Maskikit, Donald Maskikit and Jonathan Maskikit helped us in various ways to get the project going. We are particularly indebted to Vincent van Heuven, Roeland van Hout, Francesca Moro, Jos Pacilly, Toni Rietveld, Hein Steinhauer and Marco van de Ven for indispensable information and advice. We thank our speakers, Corine Bary, Ata Maskikit, Ois Maskikit, Oly Maskikit, Simon Maskikit, Kees de Schepper, Peter de Swart and Sammy Tarenskeen, for their help and patience. Results of this investigation were presented on thirteen occasions between 2013 and 2015, and we are grateful for the helpful input we received from the audiences as well as for the critical comments of three anonymous reviewers and the editors. The first author wishes to thank the Stichting Nijmeegs Universiteitsfonds (SNUF) for a travel grant to go to Ambon as part of the work for her M.A. thesis.
A report on Principal Component Analyses of the segmental landmarks in the Dutch and Ambonese Malay data is available as online supplementary materials at http://www.journals.cambridge.org/issue_Phonology/Vol33No02. Links to sound files for Figs. 7, 10, 11 and 13 can be found in the online version of the journal.

1 Introduction

The word prosody of varieties of Malay, including Indonesian, has posed a persistent conundrum, with analyses ranging from the assumption of regular stress to proposals involving the absence of any word-prosodic structure. Broadly, three views can be identified. Traditionally, word stress has been assumed to be on the penultimate syllable (e.g. Halim 1974, for four speakers in Sumatra; Cohn 1989, for four speakers from unspecified but different parts of Indonesia; Laksman 1994, for one speaker from Jakarta; Stack 2005, for one speaker of Manado Malay). The term 'Indonesian' is generally applied to varieties of the standard language, which is spoken in many parts alongside local varieties of Malay, as well as indigenous languages. Because substrates may be variably present in what is described as Indonesian, we specify the locations of the speakers in the above references. Non-default stress is generally attributed to the occurrence of schwa in the penultimate syllable, in which case most accounts describe stress on the final syllable in disyllables (e.g. Halim 1974: 76, Soderberg 2014). The assumption that Indonesian has word stress is fairly generally accepted in the phonological literature (e.g. Cohn 1989, Halle & Idsardi 1995).

The second view abandons the link between word stress and a specific syllable, with a variety of interpretations of whatever structure remains. For instance, word stress has been presented as a feature that may appear in different syllables of the Indonesian word spoken on different occasions (van Zanten & van Heuven 1998). Recognising the mobility of (putative) word stress in as a function of sentence position, Halim (1974: 63) observes that words at the end of final phrases have penultimate stress, while those occurring at the end of pre-final phrases have final stress. Going one step further, Odé (1994: 63) claims that prominence in Indonesian 'cannot be described in terms of stressed or accented syllables … the syllable does not seem to be the level on which prominence must be studied'. Mohd Don *et al*. (2008: 67) echo this suspicion, concluding that the syllable has no relevance to the prosody of Malaysian Malay, based on a lack of durational evidence for a prosodically privileged status of the penultimate syllable. In a similar spirit, van Heuven *et al*. (2008: 1285) adopt the notion of a 'pitch accent', while arguing that in Betawi Malay there is no dedicated syllable to which it is assigned.

The third class of responses have in common that they unambiguously see stress as a property of phrases, not of words. Stack (2005: 171) observes for Manado Malay that 'stress typically falls at the end of each phrase within the sentence' (cf. also Stoel 2005). Similarly, Goedemans & van Zanten (2007) suggest that Indonesian uses pitch accents as boundary markers, a view also expressed for Brunei Malay by Clynes & Deterding (2011). Based on his work on the Austronesian language Waima'a (East Timor), Himmelmann (2010) suggests that penultimate phrasal pitch accents may be characteristic of the western region of the Malay archipelago.

Ambonese Malay (ISO 639-3, also known as Malayu Ambong; van Minde 1997) is spoken by approximately 200,000 speakers in Ambon and other parts of the Central Moluccas in Indonesia (Lewis *et al.* 2014). As elsewhere in what is now Indonesia and Malaysia, Malay was introduced in precolonial times in the Moluccas, commercially important because of the trade in cloves and nutmeg, in which Ambon has held a central position since around 1600. Moluccan or Ambonese Malay has been variously counted as 'Indonesian' and as an indigenous language in censuses (Steinhauer 1994). Work on the language has focused on the lexicon and the syntax (e.g. Collins 1983); there are no studies of its prosody, other than the section on word stress in van Minde (1997). (On Ambonese Malay as a heritage language in the Netherlands, see van Engelenhoven 2002 and Moro 2016.)

§2 gives four possible interpretations of a phrase-final F0 peak, and poses the question which of these applies to Ambonese Malay, while outlining a production experiment designed to address this issue. §3 reports on the alignment of the pitch peak that putatively marks the word stress, and presents duration and spectral tilt data, together with a discussion of the extent to which these measurements reveal the existence of word stress. This discussion also addresses the claims made by van Minde (1997). To put the alignment data in perspective, we consider comparable data in Dutch, a language with undisputed word stress, and French, a language with a phrase-final pitch accent. Next, §4 considers the question whether the location or the shape of the F0 peak is sensitive to the information status of the target word, and §5 provides a phonological analysis of the sentence prosody, showing that the F0 peak represents a declarative contour that contrasts with a non-declarative rise. We summarise the findings in §6, where we also offer our conclusions.

2 A production experiment

2.1 The research issue

In the autosegmental-metrical model (Pierrehumbert 1980; cf. also Ladd 2008: ch. 3, Gussenhoven 2004: ch. 7), there are four ways in which an F0 peak can be analysed, shown in (1).

(1) a. … ma ma **ma**)$_{\omega}$)$_{\varphi}$)$_{\iota}$
H*+L

b. … ma ma ma)$_{\omega}$)$_{\varphi}$)$_{\iota}$
H* L%

c. … ma ma ma)$_{\omega}$)$_{\varphi}$)$_{\iota}$
H%

d. … ma ma ma)$_{\omega}$)$_{\varphi}$)$_{\iota}$
HL%

(1a) has a pitch accent which is associated with a lexically stressed syllable (in bold); it creates the sentence melody in combination with any boundary tones (cf. Bruce 1977, Pierrehumbert 1980 and subsequent work on numerous languages).

(1b) is exemplified by French (Post 2000, Jun & Fougeron 2002, D'Imperio *et al*. 2007), which has no word-prosodic structure (Peperkamp & Dupoux 2002, Dupoux *et al*. 2008), and where the distribution of pitch accents is a function of the prosodic phrasing. The distinction between (a) and (b) has been described as one between 'stress accents' (Beckman 1986), which are distinguished from other syllables in duration and voice quality in addition to possible F0 differences, and 'non-stress accents', which differ from non-accented stressed syllables in terms of F0 (e.g. Levi 2005, Remijsen & van Heuven 2005, Abolhasanizadeh *et al*. 2012, Hualde 2012). That is, in (b), the H* pitch accent will end up on the phrase-final syllable, which does not otherwise have the properties of a stressed syllable.

In (1c), the HL melody consists of boundary tones, at least one of which associates to a syllable (cf. Grice *et al*. 2000; we have marked both tones with '%', so as to be explicit about their status as boundary tones). Jun (1993, 2005) (see also Ko 2013) describes this case for the Seoul Korean accentual phrase, which has LH (or HH, depending on the nature of the onset consonant in the first syllable) associating with the first and second syllables, while there is a larger number of final melodies with a similar association behaviour at the phrase end. The difference between (b) and (c) may seem vacuous, since in both cases an H tone ends up on an IP-final syllable which has no stress. However, (1b), where the final syllable is accented, is required, because accents, interpreted as abstract markers for the insertion of tone (Goldsmith 1976, Hyman 1978, 2006, 2012, Gussenhoven 2004: 47), may be subject to overriding morphosyntactic generalisations, and can therefore in principle escape the final position. A case in point for French is *Que sais-je?* 'What do I know?', which has the pitch accent on *sais*, due to the unaccentable status of the 1st person singular pronoun *je* (cf. *Prends-le!* 'Take it!', which has the pitch accent on *le*).

(1d) presents boundary tones that remain floating. It differs empirically from (1c) in lacking a near-constant timing of the peak relative to some syllable. In (d), there is no need to assume either word stress or pitch accents, whether associated lexically or postlexically. Whereas in English the F0 peak in a case like (1a) marks word stress, indicates the end of the focus constituent and signals some discourse meaning, in (d) it might only indicate that the sentence has some intonationally signalled function, like 'declarative'. And if hyperarticulation were to be systematically used for signalling narrow focus, it would not be concentrated on the last syllable, but be more diffusely distributed over some prosodic phrase. The questions addressed by our experiment are which representation is appropriate for Ambonese Malay, and whether there is any systematic marking of information status in the prosody in the language.

2.2 Data collection

We designed a reading task aimed at eliciting scripted speech consisting of 80 mini-dialogues, in which either the question or the answer served as the carrier sentence of a polysyllabic monomorphemic target noun. Monosyllabic content words are rare in Malay, and, while trisyllabic words do occur, they are usually morphologically complex, often involving affixation (Adelaar 1992, van Minde 1997, Clynes & Deterding 2011). We selected nine target words whose phonological structure reflects that of the vocabulary in general, including the six disyllabic nouns in (2a), whose segmental composition varies within the limits of the permitted syllable structure, (C)V(C), where word-internal CC sequences consist of a homorganic nasal and plosive. The disyllabic word in (2b) has exceptional final stress according to van Minde (1997), while (2c) is trisyllabic. By selecting target words whose structure is representative of the language, we ensured that there was enough segmental variation to facilitate our research on the correlation between F0 turning points and segmental measures. Syllable boundaries, which are uncontroversial, are indicated by a dot. (The transcriptions of the first syllables of (2b, c) will be revised in §3.4.)

(2)	a.	anjing	[aɲ.ɟiŋ]	‘dog’
		mangga	[maŋ.ga]	‘mango’
		loteng	[lo.teŋ]	‘attic’
		mobil	[mo.bil]	‘car’
		rumah	[ru.ma]	‘house’
		ular	[u.lar]	‘snake’
	b.	tamang	[ta.maŋ]	‘friend’
	c.	balalang	[ba.la.laŋ]	‘grasshopper’

The target words were embedded in carrier sentences occurring in mini-dialogues in which both discourse mode (utterance-final declarative, utterance-final polar interrogative and non-utterance-final, henceforth ‘continuative’) and position in the IP (final and medial) were varied. This yielded 48 sentences (3 discourse conditions × (8 IP-final + 8 IP-medial)). The variation in the position of the target word (underlined) in the sentence is illustrated in (3) and (4). IP-medial target words, like *balalang* in (3b), were between one and four words from the end of the IP.

(3) a. A: Biking apa? ‘What is he doing?’
B: Dorang pung kaka mau makang <u>balalang</u>.
‘Their brother wants to eat a grasshopper.’
b. A: Tina mau apa? ‘What does Tina want?’
B: Tina suka making <u>balalang</u> deng garam.
‘Tina wants to eat a grasshopper with salt.’

The three discourse conditions are illustrated in (4a) for the declarative condition, which employs a B-sentence as an answer to a wh-question, in (4b) for the polar interrogative, where the A-sentence is the carrier sentence, and in (4c) for the continuative, where the B-sentence is the carrier sentence. Declarative sentences like the B-sentence in (4a) differed only from polar interrogative sentences like the A-sentence in (4b) in the presence of a question mark in the script.

(4) a. A: Dirk buwat apa? 'What's up with Dirk?'
B: Sekarang, Dirk mau pulang ka rumah. 'Dirk wants to go home now.'
b. A: Sekarang, Dirk mau pulang ka rumah? 'Does Dirk want to go home now?'
B: Iyo. 'Yes.'
c. A: Biking apa itu Gina? 'What is Gina doing?'
B: Gina mau masa dolo di rumah, abis itu undeng Tina. 'Gina first wants to cook at home, and then she will invite Tina.'

In addition, the information status of the eight target words was varied by including each of them in final on-focus and final post-focus positions in two declarative sentences. This yielded a further 32 sentences (2 focus conditions × 2 sentences × 8). The focus and non-focus conditions are illustrated in (5a) and (5b) respectively. Five of the sentences with *anjing* as the target word also contained a second focused constituent, as in *Tidak, ini Timo pung anjing* '(Is this Lucas's cat?) No, this is Timo's dog'.

(5) a. A: Dia mau piara monyet? 'Does she want to keep a monkey?'
B: Tidak, dia mau piara anjing. 'No, she wants to keep a dog.'
b. A: Dia mau piara anjing? 'Does she want to keep a dog?'
B: Tidak, dorang mau piara anjing. 'No, they want to keep a dog.'

Two male (M1, M2) and two female (F1, F2) native speakers were recruited in Ambon, aged 26, 55, 26 and 19 respectively. M1, F1 and F2 were monolingual, and had been born and raised in Ambon, while M2, who was also fluent in a local village language, had lived in Ambon for 28 years and used Ambonese Malay in his daily communication. They were provided with a written version of the corpus on the day of the recording, together with oral instructions, and were asked to read through the corpus at their leisure. Subsequently, they were asked to read out both the questions and the answers of all dialogues twice.[1]

[1] The elicitation of this scripted speech was first attempted with pairs of speakers, one producing the A sentence and the other the B sentence. We found this procedure less

Nineteen non-experimental dialogues were included in the corpus as fillers, ten of which occurred at the beginning and nine at the end. Recordings were made in the homes of the speakers with the help of a Sony minidisk recorder MZ-NH700 and a Sony microphone ECM-MS907, at a sampling rate of 44 kHz. Ambient noise from animals and gusts of wind was avoided during speech production by asking subjects to repeat utterances whenever this occurred.[2] One version of each of the eighty dialogues as produced by each speaker was selected, and analysed using the digital speech analysis software in Praat (Boersma & Weenink 2011). In the absence of a theory of the prosodic structure, we annotated these dialogues quite liberally, drawing on categories that are familiar from European and other languages – these labels served a heuristic purpose only. Specifically, we segmented the words, the IPs, the putatively stressed syllables of the target words and the rhymes of these syllables, adding preliminary annotations in terms of pitch accents and boundary tones.

3 Searching for phonetic evidence for word stress and pitch accents

In order to establish if there is phonetic evidence for the presence of word stress or the existence of pitch accents, we collected duration, F0 and spectral measurements. §3.1 evaluates the evidence provided by duration measurements of final and penultimate syllables of words with putative stress. §3.2 reports a timestamp analysis aimed at establishing the best segmental anchor in the word by correlating the H-timestamp with the timestamps of a number of potential anchoring landmarks, on the assumption that, if there is word stress, the location of the F0 peak should to a large extent be determined by an edge of the stressed syllable or its rhyme. Second, since the location of the target of H may be a segmentally defined point ('segmental anchoring' in the terminology of Welby & Lœvenbruck 2006) or a zone covering a syllabic constituent ('segmental anchorage'), we also carried out a latency analysis to evaluate the role of syllabic constituents in the temporal definition of the target of H. An argument for the absence of prosodic structure which is based on a series of null results may be open to the objection that more detailed measurements might have revealed the requisite evidence. To mitigate the force of any such criticism, we validated our methodology by setting up a baseline with

helpful, as it took time for participants to get attuned to each other in the way they acted out the conversations and thus to produce a consistently natural effect. The speakers preferred reading out both sentences themselves and had no problem taking on both roles in the same quasi-natural style which the mini-dialogues were intended to elicit.

2 In one instance, all recordings of a target sentence, the one representing the IP-medial occurrence of *balalang* in the continuative intonation, were spoilt by ambient noise.

comparable alignment data from a language with uncontroversial stress. To this end, we recruited four speakers of Dutch, two females (F3, F4) and two males (M3, M4), between 30 and 35 years old, who had at least one parent who, like themselves, was born and raised in Nijmegen or its immediate surroundings. We composed a list of seven target words that were segmentally comparable to the Ambonese Malay target words, all with penultimate stress. These were *Anjum* [ˈɑn.jəm] (place-name) (cf. *anjing*), *bemaling* [bə.ˈmaː.lɪŋ] 'drainage' (cf. *balalang*), *loting* [ˈloː.tɪŋ] 'lottery' (cf. *loteng*), *mango* [ˈmɑŋ.go] 'mango' (cf. *mangga*), *meubel* [ˈmøː.bəl] 'piece of furniture' (cf. *mobil*), *Roma* [ˈroː.ma] 'Rome' (cf. *rumah*) and *Oeral* [ˈuː.rɑl] 'Urals' (cf. *ular*). To match the three utterance types in the Ambonese Malay subcorpus shown in (4a) and (5), we constructed mini-dialogues in which the target word appeared in final position in the B sentence in three conditions. Two of these were a declarative and focus condition, corresponding to (4a) and (5a). Because a post-focus condition in Dutch would have yielded deaccented target words, we instead used a third condition, which yielded a question giving two alternatives, as illustrated in (6).

(6) A: Willen ze die fabriek nou in de Oeral of in de Gobiwoestijn?
'Do they want that factory in the Urals or in the Gobi desert?'
B: Die zijn ze aan 't bouwen in de Oeral.
'They're building it in the Urals.'

3.1 Duration

Mohd Don *et al*. (2008) argue that the penultimate syllable in Malaysian Malay has no stress, because its duration falls short of what they claim it should be if it did have stress. The bars in Fig. 1 gives mean durations and standard deviations of the penultimate rhymes over seven words with putative penultimate stress per speaker for Ambonese Malay (a) and Dutch (b), aligned at their mid-points. A one-way ANOVA shows that the Dutch speakers have longer penultimate rhymes than the Ambonese Malay speakers (154 *vs*. 126 ms; $F(1,166) = 16{,}829$, $p < 0.0001$) and display less variation in duration (SD 23.8 *vs*. 59.0 ms), even though word durations are not significantly different. Dutch stressed rhymes take up 33.8% of the word durations, against 26.2% in Ambonese Malay. A reviewer observes that since five of the seven Dutch target words have a long vowel in the stressed rhyme, the longer durations in the Dutch data might be due to a bias in vowel quantity. We recalculated rhyme and word durations with hybrid vowel durations on the basis of average durations of 105 ms for lax (short) vowels and 155 ms for tense monophthongs (Rietveld *et al*. 2004).[3] This reduced the proportion taken up by the

[3] In choosing the phonological structures of the Dutch words, we were guided by the segmental interpretations by Dutch native speakers of Indonesian words. Lax vowels are used in closed syllables and tense vowels in open syllables. Lax vowels

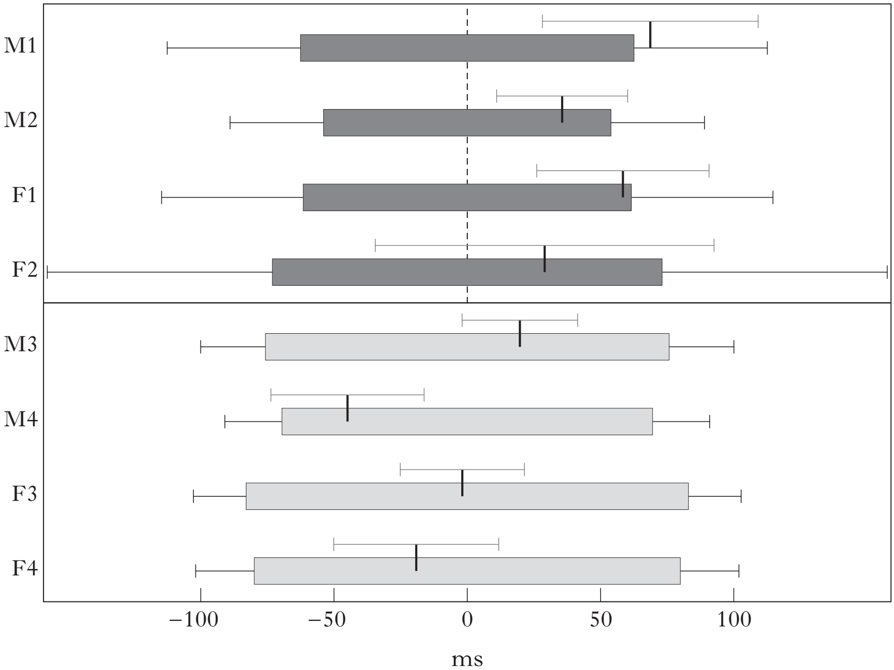

Figure 1

Mean durations of the penultimate rhyme (standard deviations indicated by hairlines), together with mean peak alignments relative to the rhyme midpoint (0 ms; standard deviations indicated by hairlines), pooled over three sentence-types, for four speakers of Ambonese Malay (M1, M2, F1, F2) and of Dutch (M3, M4, F3, F4) ($n = 21$).

stressed rhyme in the Dutch word to 30.4%. These hybrid rhymes are not significantly different from the Ambonese Malay stressed rhymes, but Ambonese word durations are now significantly longer than the Dutch hybrid vowel durations (434 *vs*. 480 ms; $F(1,166) = 9{,}181$, $p < 0.01$). Thus both methods indicate that the duration of penultimate syllables in Ambonese Malay does not stand out in any way, as opposed to the situation in Dutch.

are short, while non-high tense vowels are short in unstressed open syllables and long otherwise. High tense vowels are short in stressed and unstressed syllables, except when followed by /r/ in the same foot. This means that only the rhymes of *mango* and *Anjum* have short vowels in the stressed syllable. The hybridisations are based on original rhyme durations (ORD) multiplied by 105/130 in the case of the long vowels and $0.5 \times \mathrm{ORD} + 130/105(0.5 \times \mathrm{ORD})$ in the case of short vowels, on the realistic assumption that short vowels take up half the rhyme in syllables closed by a sonorant consonant.

Another stress-related durational effect might be leakage of final lengthening to an IP-penultimate stressed syllable, as has been found for Dutch in words ending in a syllable with schwa (Cambier-Langeveld 2000: 50). Such effects have also been found for English (Shattuck-Hufnagel & Turk 1998, Cho *et al*. 2013). A comparison of the penultimate syllable in Ambonese Malay IP-final words with the same syllable in IP-medial words failed to reveal any leakage of final lengthening to the penultimate syllable, as can be seen in Fig. 2. Separate repeated measures ANOVAs were run for the durations of the penultimate and final syllables, with Position (two levels), Word (seven levels) and Discourse (three levels) as factors. The duration of the penultimate syllable of words in IP-final position was virtually identical to that in IP-medial position, while for the final syllable in IP-final position it was significantly greater ($F(1,3) = 8.375$, $p = 0.031$, one-sided).[4] While there is a phrase-finality effect on the word-final syllable, the penultimate syllable retains its duration regardless of its position in the IP.

While these data do not suffice to demonstrate the absence of stress in Ambonese Malay, we have seen how two measures which may be seen as indicating penultimate stress, penultimate rhyme/word duration and final lengthening of stressed penults, do show up in Dutch, but not in Ambonese Malay. Fig. 1 also reveals a greater variability of the F0 peak in Ambonese Malay than in Dutch, as shown by the greater mean standard deviations of the distances between H and the beginning and end of the stressed rhyme. The next section therefore considers the alignment of the H target.

3.2 F0 peak alignment

Approximate synchronisation of locations in the F0 contour with locations in the segment string has been confirmed in various ways for various languages. Prieto *et al*. (1995) found a fairly stable segmental anchor for the beginning of the prenuclear rise in Spanish, reproducing a result for Dutch nuclear pitch accents found by Caspers & van Heuven (1993). Arvaniti *et al*. (1998) reported a stable alignment of the beginning and end of a prenuclear pitch rise in Greek with the stretch of speech defined by the vowel in the accented syllable and the strings of consonants around it. The location of the accentual peak in Dutch was found to be stably located towards the end of the accented syllable by Ladd *et al*. (2000) and, across different focus conditions, by Peters *et al*. (2014). Schepman *et al*. (2006) concluded that the beginning of the rhyme ('Silverman peak delay') and the beginning of the accented syllable ('Prieto peak delay') were better predictors of peak alignment than the

[4] We ignore the expected main effects for Word. For final syllable duration, there was a main effect of Discourse ($F(2,6) = 7.699$, $p < 0.05$), with Declarative being significantly different from Continuative in a Sidak post hoc comparison ($p < 0.05$), as well as a significant Word × Discourse interaction ($F(12,36) = 2.784$, $p < 0.01$), due to an unexpectedly short duration of *loteng* and unexpectedly long durations of *anjing* and *ular* in the Continuative condition.

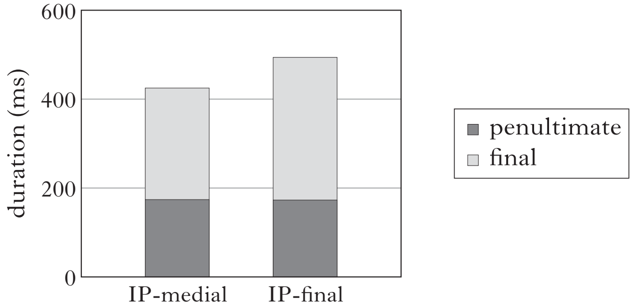

Figure 2

Duration of penultimate and final syllables of the seven Ambonese Malay target words with putative penultimate stress in IP-medial and IP-final positions in (2a, b), pooled over the four speakers ($n = 84$).

end of the accented syllable, provided the mean peak location was closer to the syllable or rhyme beginning. Moreover, variation in the distances between F0 turning points and segmental anchoring landmarks are subject to a variety of factors (e.g. Silverman & Pierrehumbert 1990). These findings suggest that, in addition to timestamp analyses, in which the correlation between the timing of an F0 turning point and the timing of a segmental landmark is investigated, we should carry out a latency analysis, in which the distance between an F0 point and some landmark is correlated with the distance between segmental landmarks that define a syllabic constituent.

3.2.1 *Timestamp analysis*. The variation in the alignment of the F0 peak in Fig. 1 for each of the speakers suggests that peaks are more stably aligned with the stressed rhyme in Dutch than in Ambonese, as indicated by the lower standard deviations of the timestamp of the peak relative to the midpoint of the penultimate rhyme. Since varieties of the same language may vary in the alignment of tonal targets (Atterer & Ladd 2004), it is the variability in target alignment rather than the average target locations which indicates how tightly a target is tied to a landmark. Timestamps were collected for six segmental landmarks in the seven target words, as well as the peak of the utterance-final rise-fall in three declarative sentences per speaker, taken from the discourse subcorpus and the focused and non-focused pronunciations from the focus subcorpus, giving 84 utterances in all.[5] The segmental landmarks were the beginning of the word (BegWd), the beginning of the penultimate syllable (BegσPen), the beginning of the penultimate rhyme (BegRhPen), the end of the

[5] The beginning and end of the fall were hard to define, because of the difficulty of deciding where the rise to the peak should be taken to begin and the considerable temporal variation in the lowest point in the final, low-intensity part of the contour.

penultimate syllable (EndσPen), the beginning of the final rhyme (BegRhPen) and the end of the word (EndWd). No major problems were encountered in the determination of the landmarks, except in the case of the NC clusters [ɲɟ] and [ŋg] in *anjing* and *mangga*, where it was hard to decide where the oral stop began. We opted to place it at the beginning of the fricated release in the case of *anjing* and at a point near the release where nasality appeared to be auditorily absent in *mangga*. Onsetless syllables were taken to begin at the start of the voicing for the vowel rather than at the start of any glottal stop or creaky phonation. The timestamp for the peak was determined automatically within the domain of the word, except in two situations. If the highest pitch occurred at the beginning of the pitch contour and was followed by a dip leading to a new high that was lower than the first, as often happened after a glottal stop at the beginning of words with onsetless syllables (*anjing, ular*), we adopted the second high as the location of the maximum. Second, if the automatic measurement was located in a voiceless stretch, as in a number of occurrences of the medial consonant in *loteng*, we adopted the higher edge of the voiceless gap. In order to calculate correlations between the H-timestamp and the six landmarks, the word beginning was set at a randomly chosen point between 1 and 99 ms before the beginning of each target word. This procedure ensured that our correlation coefficients did not spuriously reflect the large variation between the distances from the beginning of the speech file to the target word in it.

The phonological status of the melody is likely to be reflected in the temporal relation between the timing of H and segmental landmarks. For instance, if we were to find that the peak was highly correlated with the end of the word, that would provide evidence for the status of the melody as a boundary tone complex. By contrast, if it were to correlate best with the beginning of the penultimate rhyme, that could indicate an alignment with a stressed penultimate syllable. Pearson's correlation coefficients (r) were calculated between the H-timestamps and each of the six landmarks for the Ambonese Malay and Dutch data sets separately (Fig. 3a). These data discriminate between the languages in terms of variability, as shown by the lower correlations found for Ambonese Malay for Dutch. Second, the Dutch bars reach a peak at the beginning of the penultimate rhyme, while the Ambonese Malay bars are weighted towards the end of the word. Comparisons between the r's within each language show a significant difference in Dutch between those for the beginning of the penultimate rhyme and the beginning of the penultimate syllable (Fisher $z = 2.63$, $p < 0.01$), suggesting that Dutch uses 'Silverman alignment', as assumed earlier by 't Hart *et al.* (1990) and suggested by findings in Schepman *et al.* (2006). The r between the H-timestamp and the beginning of the penultimate rhyme is also significantly different from the r between the H-timestamp and the word end (Fisher $z = 2.28$, $p < 0.05$). In Ambonese Malay, the correlation of the timestamp of H with the beginning of the final rhyme is significantly different from those with the word

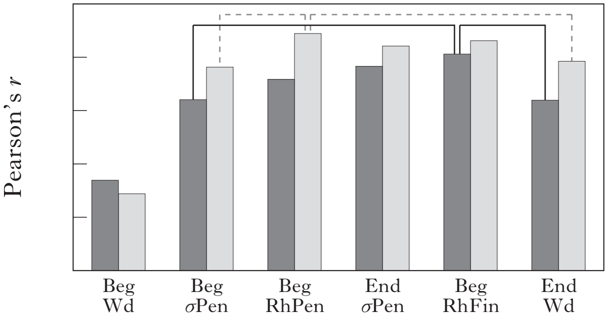

Figure 3

Pearson's correlation coefficients between the H-timestamps and each of the six landmarks for Ambonese Malay (dark) and Dutch (light). Brackets indicate significant differences among the coefficients for the last five landmarks within each language ($p < 0.05$; $n = 84$).

end and word beginning and from that with the beginning of the penultimate syllable.

The fact that the H-timestamp in Ambonese Malay is not mainly determined by any anchor in the penultimate syllable suggests that the putatively stressed syllable does not in fact attract the pitch peak, while the fact that it has a low correlation with the timestamp of the word end suggests that it is also not aligned with the right-edge IP boundary.

3.2.2 *Latency analysis*. If a segmental anchor such as the beginning of a stressed syllable is known or assumed, the duration of the constituent it begins or ends with may co-vary with the time that elapses between that segmental anchor and the H-timestamp. Welby (2006) reports the alignment of the final peak in the French accentual phrase (AP), equivalent to the IP as used here, in terms of its distance from the beginning of the AP-final syllable. Using that measure, H-latency, as the dependent variable in a stepwise linear regression analysis, she shows that 73.6% of its variation is accounted for by the duration of the AP-final syllable and a binary variable splitting the contour types into two groups on the basis of tone density. Applying this method heuristically to our Ambonese Malay data is hazardous, because of the fairly high correlations between word, syllable and rhyme durations. Putting these worries aside, the duration of the final syllable explained only 7% of the variance in the distance of the H-timestamp from the beginning of the final syllable. This improved to 30% when word duration was added, while rhyme duration did not independently contribute anything to the explained variance. This falls far short of what would be expected if, as in French, the final syllable is the docking site of a pitch accent (Post 2000, Jun & Fougeron 2002, D'Imperio *et al.* 2007). Repeating the procedure with the distance between the peak and the beginning of the penultimate rhyme as the dependent variable, we found that 40% of its variance was explained by the duration of the penultimate syllable, which improved to 44% by adding word duration. The equivalent procedures for the Dutch data yielded 50% explained variance in the H-latency from the beginning of the penultimate rhyme by the duration of the penultimate syllable. No other results were obtained, because the inclusion criteria were too low. The finding for Dutch indicates that, as well as alignment with the beginning of the stressed rhyme (Fig. 3), duration of the stressed syllable is an additional factor in determining the location of the F0 peak. No such conclusion can be drawn for either the penultimate or the final syllable in Ambonese Malay.

Instead of assuming that the H-latency to the *beginning* of some constituent is a relevant variable for the location of the F0 peak in Ambonese Malay, we considered whether its latency to the word end might co-vary with the duration of some constituent in the word. Instead of aligning with some landmark, the speaker's strategy might be to locate the peak comfortably within an IP-final domain, like the IP-final word. This conjecture turned out to be correct. The variation in the H-to-word end latency was largely explained by the duration of the final rhyme (77%), the final syllable (71.2%) or the final word (67.2%). Fig. 4a gives the correlation coefficients, with the Dutch values for comparison. Despite the greater variability in the H-timestamp in Ambonese Malay, its co-variation with the duration of final constituents is considerably greater than in Dutch. That is, the location of the peak in Ambonese Malay has an orientation with the word end, while being primarily sensitive to the available time in the final rhyme, syllable or word. To underscore the point that mean distance of the H-timestamp from the word end fails to reveal the difference between the languages, Fig. 4b shows mean H-to-word-end latencies. It is only the standard deviations which indicate that different strategies are used.

Given van Minde's (1997) claim that *tamang* has final stress, we excluded this target word from the above analyses. In order to see whether it behaves differently from the six disyllabic target words with putative penultimate stress in (2a), we carried out exhaustive pairwise *t*-tests for these seven words (i.e. $7 \times 6/2 = 21$ pairs). The word *anjing* appeared to have a significantly longer H-to-word-end latency than all other words. Among the remaining 15 comparisons, six were significant, none of which involved *tamang*. That is, contrary to what van Minde's claim would predict, the peak alignment in a word with putative final stress is no different from that in words with penultimate stress. Fig. 5 shows mean onset and rhyme durations for *tamang* and the pooled data for the six disyllabic words in (2a), as well as H-timestamps with standard deviations. Statistical analyses reported in the supplementary materials allow the further conclusion that Ambonese Malay shows greater variability among the segmental landmarks themselves. The fact that the temporal coherence of the syllabic constituents is lower in Ambonese Malay

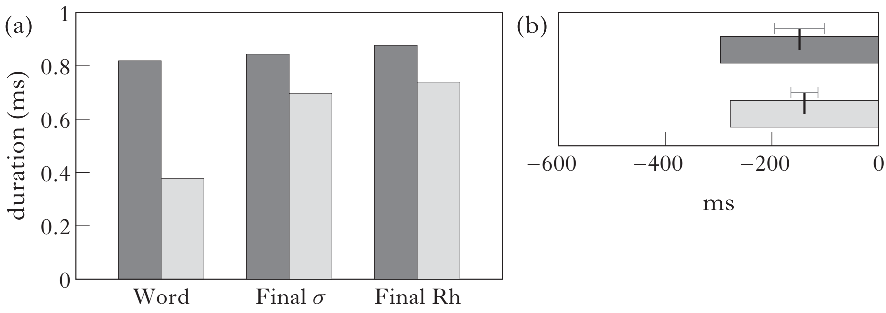

Figure 4

(a) Pearson's correlation coefficients between H-latency from the word end and the duration of the word, the final syllable and the final rhyme, for Ambonese Malay (dark) and Dutch (light) ($n = 84$). (b) Mean H-latencies from the word end, with standard deviations.

than in Dutch may reflect an absence of feet in Ambonese Malay, on the assumption that a stable pattern of syllable duration depends on the presence of foot structure.

3.3 Spectral tilt

In van Minde's (1997) analysis, the target words *tamang* and *balalang* would be characterised as having final and penultimate stress respectively. As observed by a reviewer, spectral balance may be a measure distinguishing the final syllables or rhymes in these words, such that the putatively stressed rhyme has a more even spectral balance than the putatively unstressed one. Spectral balance as a function of prosodic properties has been measured in terms of the intensity roll-off across some frequency band (Sluijter & van Heuven 1996) and the spectral centre of gravity (van Son & van Santen 2005). Various analyses of the latter measure failed to reveal any effects. Using the first measure, which would appear to be more suitable for vowels, we calculated the dB decrease per decade over logarithmic intensity levels across 35 contiguous bands within the 500–4000 Hz frequency range in utterances from the four speakers containing these words in the declarative, interrogative, prefinal and focus contexts ($n = 16$). A repeated measures ANOVA with Word (two levels) and Discourse (four levels) as fixed factors yielded a main effect for Word ($F(1,3) = 10.878$, $p < 0.05$), due to a less steep spectral balance for *tamang* than *balalang* (−49.8 dB/dec *vs*. −54.7 dB/dec). Apart from the small effect size, this result is somewhat fragile. No difference was found when the frequency band was reduced to 500–3000 Hz or increased to 500–6000 Hz. As an indication of a difference in word stress, this result is not convincing.

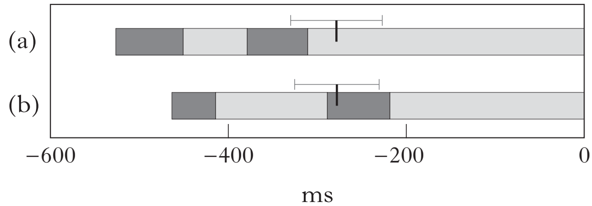

Figure 5

Onset (dark grey) and rhyme (light grey) durations for (a) *tamang* ($n = 20$) and (b) the six disyllabic words in (2a) ($n = 120$), with H-timestamps and their standard deviations, aligned at the word end.

3.4 Evaluating symptoms of stress

In §3.1, we saw that there are no indications in the duration measurements that the putatively stressed penultimate syllables in the seven target words are actually stressed. We found no effect of prefinal lengthening on the penultimate syllable, nor was the penultimate syllable longer than might be expected if it had no stress. In §3.2, we saw that the F0 peak did not align with the beginning or end of the penultimate syllable or with the beginning of its rhyme. By contrast, we found strong correlations between the distance of the F0 peak from the word end with the duration of the final rhyme, the final syllable and the word. These latter findings suggest that the F0 peak is due to a final boundary melody, whose alignment is sensitive to the space available in the word and the final syllable. The rise-fall melody is thus placed over the last word in the IP, with a tendency to be timed earlier as the word is longer. These conclusions can be seen against the perspective of measurements in a segmentally equivalent set of Dutch words, all of which indicate that the penultimate syllable has stress and that the rise-fall melody is not a boundary tone complex.

Since the F0 peak in Ambonese Malay is located near the boundary between the penultimate and final syllables, not unlike the F0 peak for French non-downstepped H* (Welby 2006), an argument for an association of H with the final syllable in Ambonese Malay might have been found in a tendency to locate the F0 peak at some distance from the beginning of the last syllable as a function of the duration of that syllable, as Welby found for French. No such tendency was found, however, and so no plausible case for an association of the peak with the final syllable can be made. The F0 data suggest that of the four representations in (1), the correct one for Ambonese Malay is (d). That is, the tones are boundary complexes which are not associated to any tone-bearing units.

If *tamang* is prosodically no different from any of the other disyllabic words, an explanation is required for the minimal pairs contrasting penultimate and final stress (van Minde 1997: 23). Some of his examples are given in (7).

(7)	ˈbacang	(name of island)	baˈcang	(kind of fruit)
	ˈbale	'to return'	baˈle	'to coil up'
	ˈbali	(name of island)	baˈli	'to buy'
	ˈbarat	'west'	baˈrat	'heavy'
	ˈmasing	'salty'	maˈsing	'engine'
	ˈparang	'machete'	paˈrang	'war'
	ˈsalak	(kind of fruit)	saˈlak	'at once'

It is unclear why the penultimate syllables in these pairs contain /a/. Since van Minde describes the language as having five vowels, /i e a o u/, this implies that if Ambonese Malay has stress, only /a/ will reject it.[6] The question therefore arises why /i e o u/ cannot appear in unstressed penults, i.e. why they unexceptionally attract stress in that position (e.g. **biˈlang*, **uˈlir*). In languages in which stress is sensitive to vowel quality, open vowels tend to attract stress (as opposed to being stress-repelling) (Wordick 1982: 42, Crowhurst & Michael 2005). Equally oddly, the vowel /a/ would appear in three prosodic guises, because short [ă], while being very frequent in syllables before the penult, is not obligatory there. Words like *kalapa* /kălapa/ 'coconut', *carita* /cărita/ 'story, to tell' and *calana* /călana/ 'pants' contrast with a less common set of words that have regular /a/ in antepenultimate position, like *kakarlak* /kakărlak/ 'cockroach', *angkatang* /aŋkataŋ/ 'generation', *lalamung* /lalamuŋ/ (kind of algae) (Francesca Moro, personal communication). The contrast between the vowels in the antepenultimate syllables would have to be analysed as one between 'unstressed' /a/ and secondary stressed /a/, since the putative primary stress is either on the final syllable (/kakărlak/) or on the penult. Moreover, secondary stress would not be governed by foot structure, at least not to the extent that a trisyllabic word with final stress would have a footed initial syllable, as witnessed by words with 'unstressed' /a/ in both penultimate and antepenultimate position, like *tabala* (ta + bala) /tăbăla/ 'to be split' or *bacarita* (ba + carita) /băcărita/ 'to tell' (van Minde 1997: 96, 307).

The assumption of a three-way stress contrast for /a/, together with stress-prone /i e o u/, does not yield a plausible stress system. This suggests that 'unstressed' [a] must be a sixth vowel. Two interpretations can be given. First, it could be categorically stressless, like English schwa. This would entail that the set of unstressed vowels would be larger than the set of stressed vowels (five), given that there are many polysyllabic words that do not contain the categorically unstressed vowel. This analysis

[6] Van Minde also lists *ˈbobo* 'sleep (child)' – *boˈbo* 'to smell'. Since *bobo* 'to smell' is a reduplicated word, it is not minimally different from *ˈbobo* 'sleep' (Francesca Moro, personal communication). Van Minde points out that he considers his analysis provisional and explicitly leaves room for further insights by phonetic measurements.

would moreover imply that 'unstressed' /ă/, the sixth vowel, and 'unstressed' regular /a/ are contrastive in unstressed antepenultimate syllables. The second interpretation of the sixth vowel is that it is a weak vowel which is independent of stress. The vowel is comparable to French schwa, which contrasts with /e/, as in *peler* /pəle/ 'peal' *vs*. *Pelé* /pele/ (proper name), without there being an implication of a difference in stress. By analogy, we will refer to the Ambonese Malay weak vowel as 'a-*caduc*', and assume that it is distinguished from the other five vowels in being moraless.[7] The fact that Indonesian, but not Ambonese Malay, has been analysed as having the six vowels /i e a o u ə/ must be attributed to the fact that the spectral quality of Indonesian /ə/ differs from the spectral profiles of the other five, while that of the Ambonese vowel is very similar to that of /a/, as shown in Fig. 6. It presents plots of the six vowels in a normalised F1−F2 space, as spoken in IP-final target words in two declarative sentences and two interrogative sentences.[8] The equation in the literature of a-*caduc* with regular /a/ is evidently due to their close similarity in vowel quality.

Neither our phonetic measurements nor the data in van Minde (1997) lead to the conclusion that Ambonese Malay possesses word-prosodic structure, whether lexical, in the form of word stress, or postlexical, in the form of a phrase-final pitch accent. The only finding that is consistent with the existence of word stress is a fragile effect on spectral tilt, as reported in §3.3. Arguably, it can be attributed to greater F0 change in the final syllable of *tamang* /tămaŋ/, which is longer (cf. Fig. 5) and has a larger portion of the F0 peak in the final rhyme than the final syllable of *balalang* /bălalaŋ/. Despite the absence of word prosody, the language clearly has syllable structure, as evidenced by /tu.u/ [tuʔu] 'kind of fish' *vs*. /pus/ [pus] 'cat' (1997: 24, 30), the general (C)V(C) formula, with an additional optional C at the beginning of loanwords (e.g. /stip/ [stip] 'eraser', from Dutch *stuf* /stʏf/; 1997: 52), and speakers' ability to count the syllables of words. It also has prosodic words, as evidenced by the fact that reduplicated forms contrast with segmentally identical monomorphemes. The reduplication of hypothetical /goŋ/ (a potential loan from Dutch /ɣɔŋ/ 'gong') would be prosodically distinct from the noun /goŋgoŋ/ 'to bark' (van Minde 1997: 29), while the results of the H-latency analysis indicate the word and the syllable are prosodic domains.

7 The analogy with French e-*caduc* is imperfect to the extent that the French vowel is subject to alternation with [œ] when accented, and appears in monosyllabic words as well as in non-final syllables of polysyllables.

8 The sixteen exemplars of [u] came from *rumah* (F1, M1) and *ular* (F2, M2), of [i] from *anjing* (F1, M1) and *mobil* (F2, M2), of [e] from *loteng* (F1, M1), of [o] from *loteng* (F1, M1) and *mobil* (F2, M2), of [a] from the second syllables of *balalang* (F1, M1) en *mangga* (F2, M2). The values for [a] *caduc* come from the first syllable of *tamang*. In all cases, two measurements were taken, one at 25% and one at 75% of the vowel.

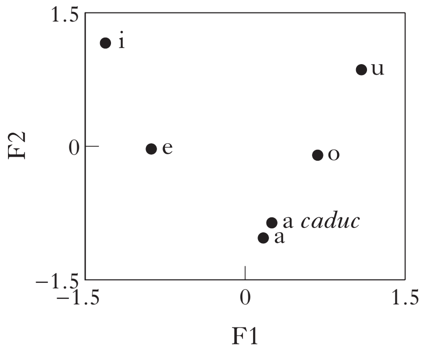

Figure 6

Normalised F1 – F2 plots (Lobanov 1971) of six Ambonese vowels pooled over four speakers, based on the means of measurements taken at points at 25% and 75% of the duration of the vowel ($n = 16$).

4 Focus effects

The prosodic marking of the end of the focus constituent is a well-known feature of languages spoken in Europe and Asia, with words following the focus constituent being spoken with compressed pitch range ('post-focus compression'; Xu *et al*. 2012). Phonologically, post-focus compression is diverse. In English, it is due to the absence of pitch accents after the focus, causing the post-focal pitch to be low after an L-ending pitch accent and mid or high after an H-ending pitch accent (e.g. Pierrehumbert 1980; a different view is taken by Xu & Xu 2005). In Portuguese, it is due to the occurrence of an H+L* pitch accent after the focus, as opposed to H*+L in the focus constituent (Frota 2000). In Mandarin Chinese, a phonetic reduction of the pitch range in the post-focal words occurs, whereby the tonal representation is the same as in the focused case (cf. Xu 1999). In Japanese, it has been argued that the phrasing structure is affected in such a way that the focus constituent is never downstepped, and thus has a relatively wide pitch range (Pierrehumbert & Beckman 1988, Kubozono 1993), while the post-focal words may lose their status as an accentual phrase, causing a reduction of the pitch range relative to pronunciation of those words as a separate accentual phrase (Sugahara 2003; see Ishihara 2007 for an alternative view). In this section, we report on an investigation of the prosodic effects of focus in our Ambonese Malay data.

Thirty-two sentences with eight targets words in final position were included in our recordings, to test for effects of information focus on prosody. The B response in (5a) served as the on-focus condition and that in (5b) as the post-focus condition. The 16 sentences in the on-focus condition included five with double focus. In (8), the first contrastive focus is a proper name and the second, *rumah*, the target word.

(8) A: Johan tidur di Henk pung kamar?
'Is Johan sleeping in Henk's room?'
B: Tidak, Johan tidur di Stefan pung rumah.
'No, Johan is sleeping in Stefan's house.'

4.1 Duration and focus

Duration measurements were taken of the penultimate syllable, the final syllable and the entire word of all target words. Repeated measures ANOVAs were performed on each set, with Word (eight levels) and Focus (two levels) as factors. In all three cases, there were significant main effects of Word, as expected in view of the variation in their segmental structures, where penultimate syllables varied from [u] to [maŋ] and final syllables from [ga] to [maŋ] (for word duration ($F(7) = 21.844$, $p < 0.001$), for the penultimate syllable ($F(7) = 16.156$, $p < 0.001$) and for the final syllable ($F(7) = 24.680$, $p < 0.001$)). In no case, however, was there a significant effect of Focus or any interaction between Word and Focus.

4.2 Pitch and focus

A preliminary inspection of sentences with narrow corrective focus on IP-final target words and those with post-focal target words did not reveal any systematic pitch range reduction in the latter condition. Fig. 7 shows illustrative pronunciations of the same sentence with the target word in the narrow corrective focus (a) and the post-focus (b) context. The pitch contours are similar, not only in the target words, but overall.

However, the effects of focus on the prosody may be more subtle. In order to establish if there might be small phonetic differences in pitch height between corrected and repeated IP-final words, we averaged the F0 measurements of the contours in the two focus conditions for each target word separately. In 5.1% of cases, empty cells due to measurement gaps in voiced sections of the contours of up to three measuring points were supplied by interpolated values. In 3% of cases, measurement errors due to octave jumps were corrected. This left us with a further 1.5% of missing values, which were ignored in the calculation of the means. The results, reported in semitones in an attempt to reduce effects of gender, are shown in Fig. 8, where time scales have been normalised. Pooled over speakers, the contours appear to be very similar indeed. A hypothesis that the on-focus contour is pronounced with an expanded pitch range can be tested on the basis of the number of times that the difference between the values for the averaged on-focus and post-focus contours at equivalent normalised time points is positive. The mean difference

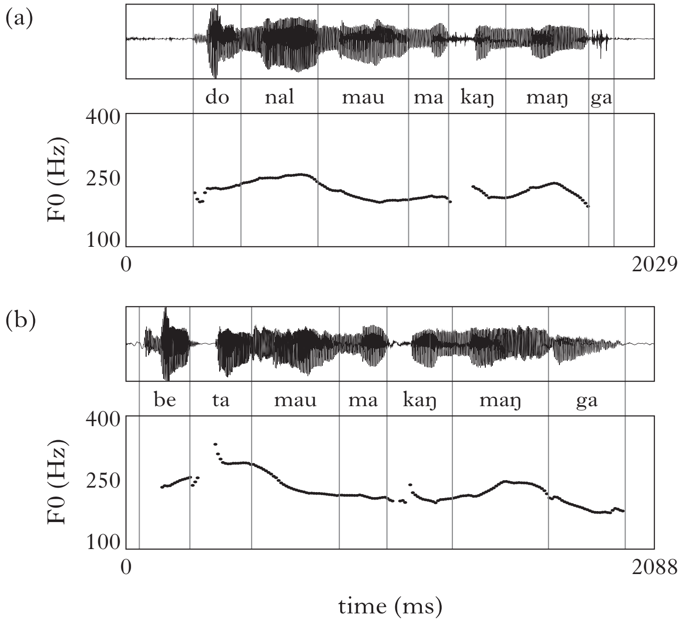

Figure 7

F0 tracks and waveforms of (a) *Donal mau makang mangga* 'Donald wants to eat a mango', with an IP-final target word in focus, and (b) *Beta mau makang mangga* 'I want to eat a mango', with the same target word after the focus. (Speaker F2.)

between measuring points is 0.23 semitones, meaning that the on-focus condition has a negligibly higher F0 overall than the post-focus condition. Since we cannot assume that our speakers behaved uniformly, we tested for the independence between the four speakers and numbers of positive and negative ΔSTs, whereby nil differences were equally divided over the two categories. We obtained a contingency coefficient of 0.32 ($p < 0.001$) and a significant result ($\chi^2[3] = 107{,}959$, $p < 0.001$), meaning that there is a significant effect of the focus condition, but that speakers did not behave uniformly. Tests of independence for the four speakers separately showed significant effects for speakers M1 ($\chi^2[1] = 18.612$, $p < 0.001$, contingency coefficient = 0.19) and M2 ($\chi^2[1] = 66.353$, $p < 0.001$, contingency coefficient = 0.35), with mean ΔSTs of 0.37 and 0.33 semitones respectively. Thus two of the four speakers had higher pitch in the on-focus condition than in the post-focus condition. Four of the eight target words showed a significant difference in the data pooled over the four speakers: *anjing* ($\chi^2[1] = 15.6$, $p < 0.001$), *balalang* ($\chi^2[1] = 6.32$, $p < 0.05$), *tamang* ($\chi^2[1] = 4.0$, $p < 0.05$) and *rumah* ($\chi^2[1] = 21.1$, $p < 0.001$). We conclude that Ambonese Malay does not express information focus in its prosody, but that a phonetically detailed investigation

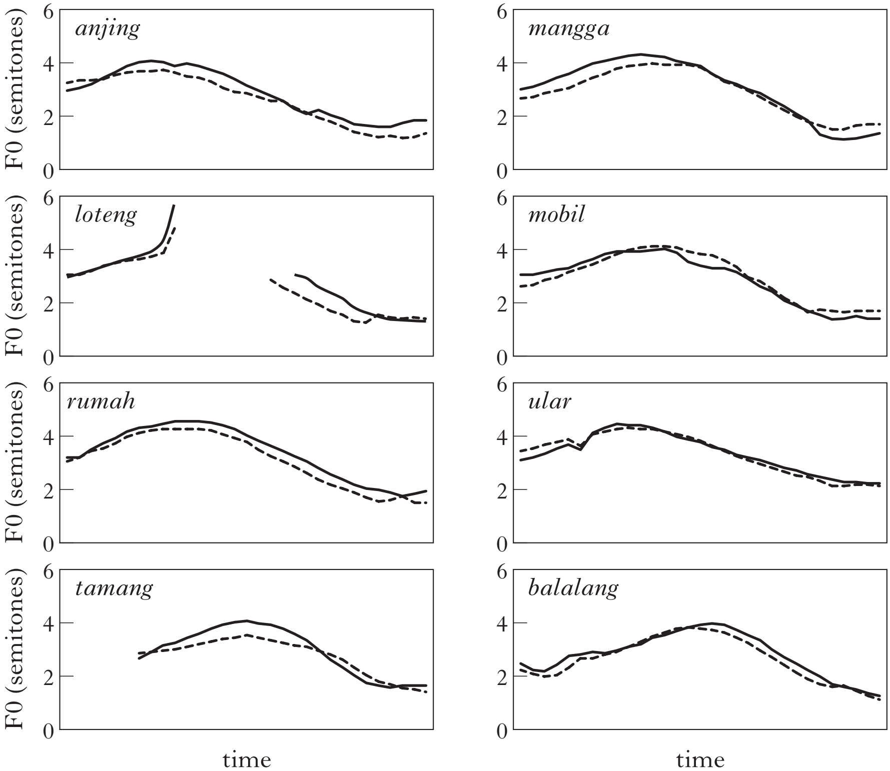

Figure 8

F0 tracks (semitones, relative to 100 Hz) of the target words in (2) in sentence-final position with corrective declarative focus (solid lines) and neutral declarative focus (dashed lines) ($n = 8$). Normalised time scale.

may reveal unsystematic and small effects on the pitch range of words with narrow corrective focus compared to repeated words, all else being equal. This result is relevant to the role of meaning in models of phonetic implementation, even if they do not support a structural difference between corrective and repeated NPs. Our conclusion is in general agreement with those drawn by Zerbian (2007) for Sotho, by Kügler & Skopeteas (2007) and Gussenhoven & Teeuw (2008) for Yucatec Maya, Gut *et al*. (2013) for Malaysian English, Wang *et al*. (2011) for the Mon Khmer languages Wa and Daeng and the Sino-Tibetan language Yi, and Xu *et al*. (2012) for Taiwanese Mandarin. Downing (2008) shows that Chichewa, Chitumbuka and Zulu lack prominence marking of the focus, but may express focus in the prosodic phrasing.

5 Sentence prosody

If the rising-falling F0 movement which we investigated in §4 and §5 signals neither stress nor focus, the remaining option is that it signals declarative intonation, and as such functions in a system of intonational pitch contours. This section investigates the duration and F0 of the target words as a function of the three discourse conditions, declarative, interrogative and continuative, illustrated in (4) above. §5.1 reports duration measurements, while §5.2 reports and analyses the F0 properties of the clause-wide contours observed in these three discourse conditions.

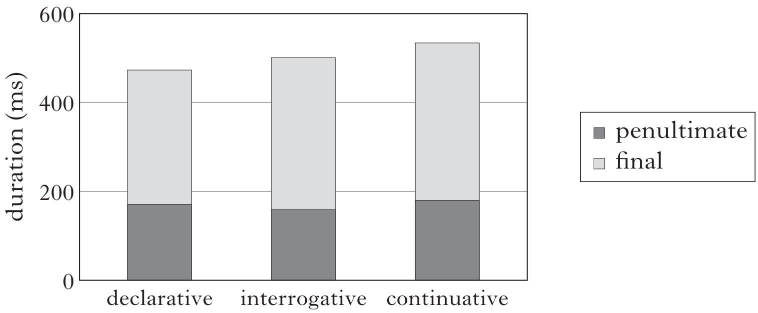

Figure 9

Duration of penultimate and final syllables of the eight Ambonese Malay IP-final target words in (2) in three discourse conditions, pooled over four speakers ($n = 32$).

5.1 Duration of target words in three discourse conditions

While syllable durations are not affected by the putative stress described in the literature, durational effects may occur as a function of the intonational melody or the discourse condition. Figure 9 gives the durations of the final and penultimate syllables of eight IP-final target words across the three discourse conditions. Repeated measures ANOVAs with Word (eight levels) and Discourse (three levels) as factors yielded a significant effect of Word for final syllables ($F(7,21) = 10{,}672$, $p < 0.001$), while no effect of Discourse was observed. The duration of the penultimate syllable was significantly affected by both Word ($F(7,21) = 27{,}804$, $p < 0.001$) and Discourse ($F(2,6) = 6{,}101$, $p < 0.05$), while there was a significant interaction between Word and Discourse ($F(14,42) = 2{,}513$, $p < 0.05$; all Huynh-Feldt corrected). For total word durations the same effects were found as for the penultimate syllable: Word ($F(7,21) = 27{,}791$, $p < 0.001$) and Discourse ($F(2,6) = 7{,}454$, $p < 0.05$), with a significant interaction between Word and Discourse ($F(14,42) = 3{,}024$, $p < 0.01$).[9] Post hoc comparisons showed that the continuative condition is significantly longer than

[9] The penultimate syllable of *ular* was disproportionately lengthened in the interrogative condition, while those of *anjing* and *balalang* were disproportionately

the interrogative condition. Therefore, despite considerable variation, final lengthening as measured over the final two syllables of the IP is stronger in the non-utterance-final continuative condition than in the utterance-final conditions.

Two findings require an explanation. First, because the lower-ranked IP shows more final lengthening than the higher-ranked utterance, the regularity that final lengthening increases with the rank of the boundary (Gordon & Munro 2007 and references therein) is absent in our data. We believe that our finding shows the manipulative use of final lengthening for signalling non-finality utterance-medially. However, more research is required to see if this strategy is used in interactive speech styles. Second, penultimate shortening affected interrogative contours in comparison with continuative ones. It is well-known that IP-final syllables tend to be longer in interrogatives. Interestingly, this was found to occur at the expense of IP-medial syllables by Stoel (2007) for Fataluku and Abolhasanizadeh *et al*. (2012) for Persian.[10] Van Heuven & van Zanten (2005) report tempo increases in various sections of interrogative utterances in Manado Malay and Orkney English, as well as to a lesser extent in Dutch, which they tentatively explain as being due to the use of fast speech as a substitute for high pitch (cf. Rietveld & Gussenhoven 1987) or to the effect of a greater tension in questions relative to statements (cf. Bolinger 1989). Rejecting these explanations, we suggest that penultimate shortening may apply in interrogatives as a result of an utterance-internal higher tempo implemented by way of compensation for longer final syllables. The longer final syllable itself can be explained by the fact that rising pitch takes longer to produce than falling pitch (Ohala & Ewan 1973), a durational difference that may be maintained in the absence of any actual pitch differences (Smith 2002).

5.2 Pitch contours

An inspection of the pitch melodies in our corpus revealed two types, 'declarative' and 'non-declarative' contours. The declarative contour was used in the final IPs of declarative sentences as well as in the 32 declarative sentences intended to reveal effects of focus. The non-declarative melody was used in polar questions and in continuatives (i.e. non-final IPs). While we had no wh-questions among our target sentences, inspection of the many non-target wh-questions shows that they too have the non-declarative contour.

---

lengthened in the continuative condition. The word duration of *mangga* was shorter and that of *tamang* longer than expected in the non-final condition.

[10] The compensatory shortening of non-final syllables in questions is confined to the penultimate syllable in our data. There was no effect of Discourse on the duration of either the final syllable or the penultimate syllable. In this corpus, the missing data for the one instance of *balalang* were supplied by taking the mean of the values for the three other speakers and correcting it for speaker by multiplying it by the ratio between the overall average duration of this speaker and the three other speakers.

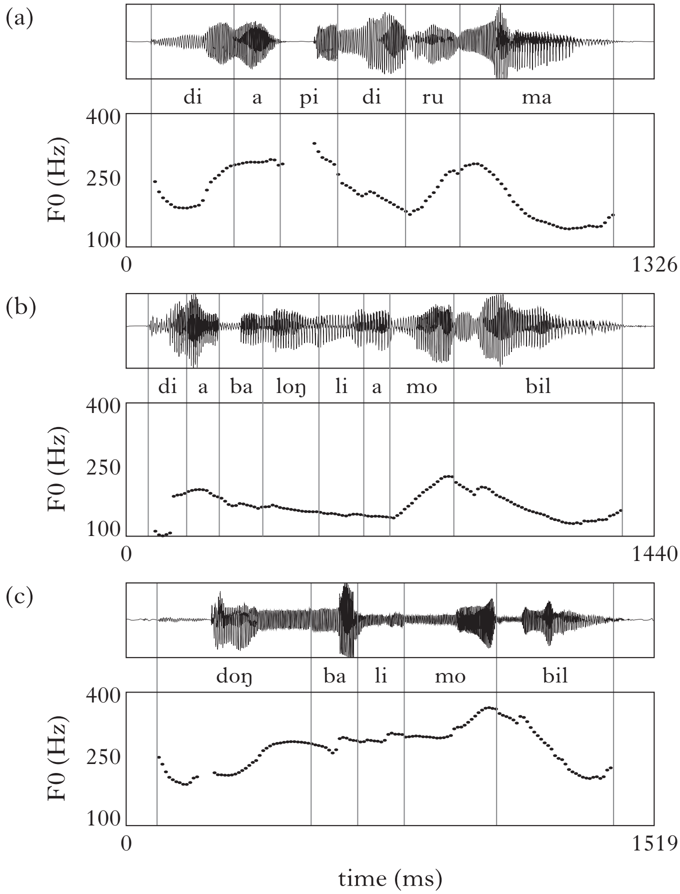

Figure 10

F0 tracks and waveforms of three declarative IPs: (a) (*Kim pi di gereja? Tidak.*) *Dia pi di rumah* '(Is Kim going to church? No.) She is going home' (speaker M1); (b) (*Anak itu su lia mobil? Tidak.*) *Dia balong lia mobil* ('Did the child see the car? No.) He didn't see the car yet' (speaker M1); (c) (*Dorang su bali sepeda? Tidak.*) *Dong bali mobil* '(Did they buy a bike? No.) They bought a car' (speaker F1).

The IP-wide declarative contour has an initial rise and a final fall, with the part in between showing varying lowered pitch shapes, broadly confirming a similar observation by Mohd Don *et al*. (2008) for Malaysian Malay. The initial rise occurs over the first two syllables, and may be completed within the first syllable in the case of monosyllabic words. An example of this two-peak contour is shown in Fig. 10a. The initial peak is often lower, and the gradual fall accordingly shallower, as in (b), spoken by the same speaker. In shorter phrases, the initial peak may lack a fall, such that the pitch between the peaks has a rising trend, as in (c), with an acceleration of the rise in the final word (cf. Jun 1993, Jun & Fougeron 2002 and Welby 2006 for the effect of the number of syllables on the realisation of tones in Korean and French). The shape of the final falls is considerably more stable than that of the preceding stretch.

In interrogative IPs, we observe a similar pattern at the beginning of the IP, while a pitch rise occurs at the right edge, often occurring in the final syllable, as illustrated in Fig. 11a. As in the declaratives, the initial peak may be low or perhaps absent, while the final rise frequently begins before the final syllable, as in (b). Two rising sections in the final rise can sometimes be distinguished, the second faster than the first, whereby a weak dip during the onset consonant of the final syllable may act as a divider, as in (c), where the rising trajectory begins in the penultimate [ru], dips down during [m] and continues upwards from there.

For the continuative contour, we first observe that the IP-final pitch rise has the same shape as in interrogative IPs, as shown in Fig. 12. The downturn at the end of the interrogative contour is due to the utterance-final pitch reversal in the fading final portion of the speech signal (cf. Gussenhoven 2004: 9). The graphs are based on the pitch contours over the final two syllables of eight target words in the two conditions by four speakers, and are time-normalised. Due to unvoiced sections, 10% of the data points are missing, largely at the beginning of the first syllable and the end of the second.

IP-internally, further intonationally marked phrasing occurs. Many clausewide contours show clause-internal prosodic breaks marked by a steep final rise, but without initial rise, and typically followed by a steep fall. The following IP-internal phrase has slowly rising pitch towards the beginning of the pitch movement at the next boundary. The post-modifying demonstratives *ini* 'this' and *itu* 'that' are almost always followed by such clause-internal boundaries. Figure 13a shows a contour consisting of two IPs, with the IP boundary after *anjing* in the continuative IP. The second IP is divided over the two lower-ranking phrases *abis itu* 'after that' and *baru belajar* 'next (he will) study'. In addition to demonstratives, proper names also often attract these lower-ranking phrase breaks after them. The contour in Fig. 13b shows clause-internal final rises after *Jonathan* and *kaka* 'brother', followed by an IP-final declarative contour. This phrase-final pronunciation of *Jonathan* should be compared with the pronunciation of the same name without a following boundary in Fig. 11b, where it has the IP-initial rise, occurring in the penultimate syllable. The contour in Fig. 13c consists of two IPs, the first ending after *tamang* '(my) friend'. Each has an IP-internal boundary, after *dolo* 'first' in the first IP, and after *baru beta pi* 'next (I will) go' in the second.

A provisional tonal analysis of the IP should include two final boundary tone complexes, since the non-declarative final melody has the same shape when marking continuatives and interrogatives. Each of these has an initial

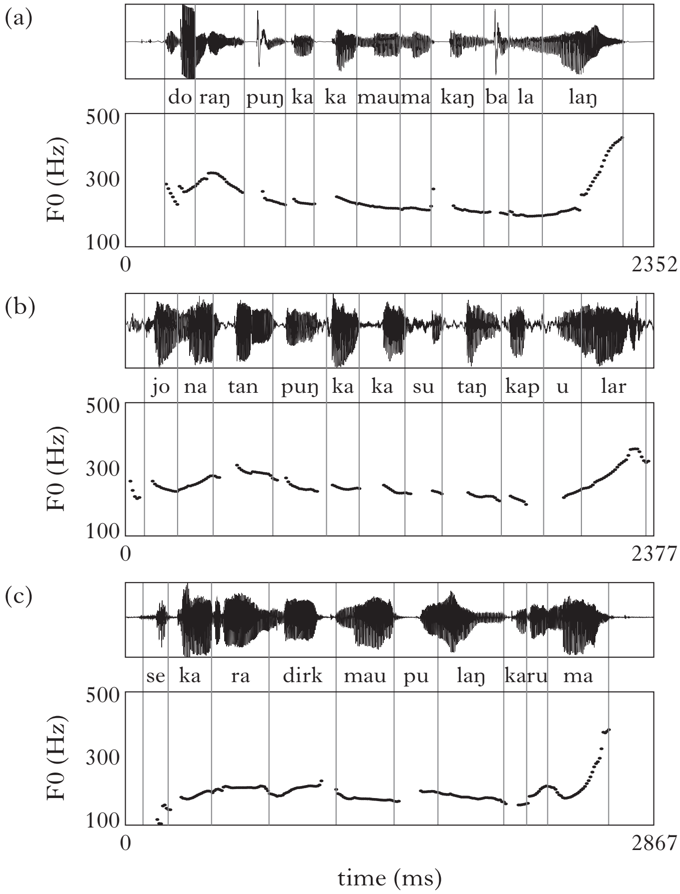

Figure 11

F0 tracks and waveforms of three interrogative IPs: (a) *Dorang pung kaka mau makang balalang?* 'Does their brother want to eat a grasshopper?' (speaker F1); (b) *Jonathan pung kaka su tangkap ular?* 'Is Jonathan's brother catching a snake?' (speaker F2); (c) *Sekarang, Dirk mau pulang ka rumah?* 'Does Dirk want to go home now?' (speaker M1).

rise, analysed as $L_\iota H_\iota$, occurring over the first one or two syllables of the IP, as illustrated in the contours in (9a) and (c). Since IPs always end with a rise, with or without a following fall, and IP-internal, lower-ranking prosodic breaks are marked by pre-boundary rises, we will assume a tonally sensitive prosodic constituent below the IP, like the intermediate phrase or the phonological phrase, labelled φ in (9), which has $L_\varphi H_\varphi$ at the right edge. The declarative IP can then be analysed as having L_ι following the last $L_\varphi H_\varphi$, as in (9b, c). For non-declaratives we

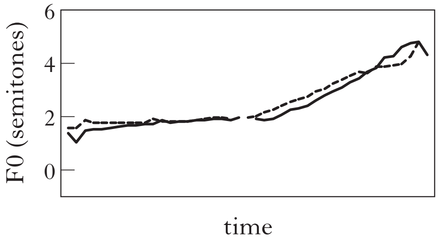

Figure 12

Time-normalised F0 tracks (semitones relative to 100 Hz) over the penultimate and final syllables of eight IP-final target words in final interrogative IPs (solid line) and in non-final IPs (dashed line) ($n = 32$).

can assume an IP-final H_ι, as shown in (9a, c). The double occurrence of an H tone at the end of the IP may be motivated by the occasional occurrence of two clearly discernable rising sections, as in Fig. 11c, but more research into the durational and F0 properties of IP-internal φ-breaks, as in (9b) and the first and third breaks in (9c), and utterance-internal IP-breaks, like the second break in (9c), is clearly required.

(9) a.

$(((\text{Jonathan})_\omega\ (\text{pung})_\omega\ (\text{kaka})_\omega\ (\text{su})_\omega\ (\text{tangkap})_\omega\ \text{ular})_\omega)_\varphi)_\iota$

$L_\iota H_\iota$ … $L_\varphi H_\varphi H_\iota$

'Is Jonathan's brother catching a snake?'

b.

$(((\text{Jonathan})_\omega)_\varphi\ ((\text{pung})_\omega\ (\text{kaka})_\omega)_\varphi\ ((\text{ada})_\omega\ (\text{tangkap})_\omega\ \text{ular})_\omega)_\varphi)_\iota$

$L_\varphi H_\varphi$ … $L_\varphi H_\varphi$ … $L_\varphi H_\varphi L_\iota$

'Jonathan's brother is catching a snake.'

c.

$(((\text{Beta})_\omega\ (\text{pangel})_\omega\ (\text{dolo})_\omega)_\varphi\ (\text{tamang})_\omega)_\varphi)_\iota$

$L_\iota H_\iota$ … $L_\varphi H_\varphi$ … $L_\varphi H_\varphi H_\iota$

$(((\text{baru})_\omega\ (\text{beta})_\omega\ (\text{pi})_\omega)_\varphi\ ((\text{di}))_\omega\ (\text{pasar})_\omega)_\varphi)_\iota$

$L_\varphi H_\varphi$ … $L_\varphi H_\varphi L_\iota$

'I will first call my friend and then go to the market.'

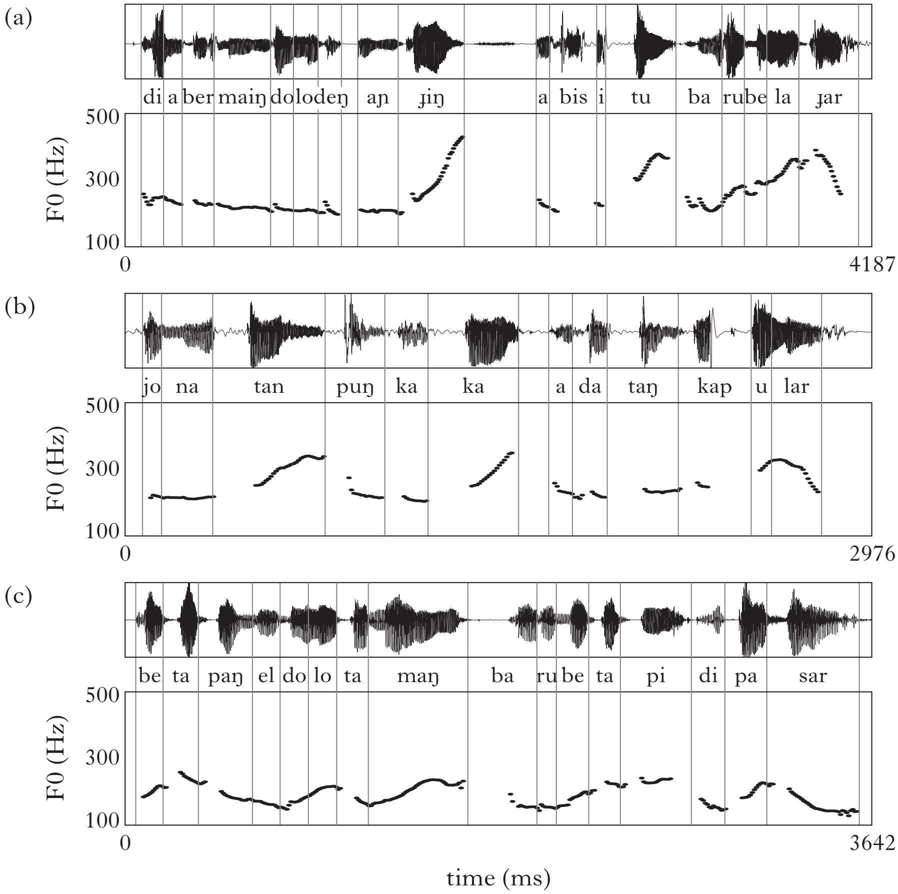

Figure 13

F0 tracks and waveforms of three continuative IPs: (a) *Dia bermaing dolo deng anjing, abis itu baru belajar* 'He is first going to play with the dog, and then he will study' (speaker F1); (b) *Jonathan pung kaka ada tangkap ular* 'Jonathan's brother is catching a snake' (speaker F1); (c) *Beta pangel dolo tamang, baru beta pi di pasar* 'I will first call my friend and then I will go to the market' (speaker M1).

6 Conclusion

A detailed phonetic investigation of temporal alignments of the IP-final declarative pitch peaks and of syllable durations revealed that in no case did those measurements suggest that Ambonese Malay has word stress. An exhaustive set of segmental landmarks failed to correlate with the F0 peak in a way that was interpretable as being due to word stress, and no word-internal constituent could be identified that serves as a synchronisation point for the peak. Second, the duration of the penultimate syllable, which in earlier studies has been designated as a stressed syllable, is neither affected by preboundary lengthening nor in any way attributable to stress. These findings support similar conclusions reached by Odé (1994) and Mohd *et al*. (2008) for other varieties of Malay. Our results contrast with findings obtained for Dutch, reported here and by Cambier-Langeveld (2000), which show that the beginning of the rhyme of the stressed penultimate syllable shows a high correlation with the timing of the peak, confirming data presented by Schepman *et al*. (2006) and a practice-based assumption by 't Hart *et al*. (1990). For Dutch, the beginning of the penultimate (stressed) rhyme accounted for 81% of the variance in the location of the pitch peak, against 57% in the Ambonese Malay data, less than for segmental landmarks that have not been claimed to be relevant to stress in the language. Phrase-final lengthening of penultimate stressed syllables was established for Dutch by Cambier-Langeveld (2000).

We also found that the distance between the beginning of the final syllable and the pitch peak failed to correlate with the duration of the final syllable, quite unlike the strong correlation for French presented by Welby (2006). This rules out an interpretation of the pitch peak as a pitch accent associating with the phrase-final syllable, as in French. The one measure for which Ambonese Malay scores higher than Dutch concerns the correlation between the distance from the F0 peak to the word end and the duration of the final rhyme, syllable or word. This finding prompted the conclusion that while Ambonese words have no word-prosodic structure, words are referred to as domains within which the rising-falling pitch movement is placed.

Correlations among the Ambonese Malay landmarks are low, reflecting a looser temporal word structure than is found in Dutch. To underscore this relative lack of word-internal temporal integration, we performed Principal Component Analyses on the segmental landmarks in the Dutch as well as the Ambonese Malay data, reported in the online supplementary materials. While the *patterns* of correlations among the time-stamps of the various segmental landmarks are similar in the two languages, the degree to which they correlate is substantially smaller in Ambonese Malay. This may reflect the absence of any word-prosodic (metrical) structure, if it is assumed that temporal integration of segmental landmarks requires a foot structure, and that without it, cross-syllabic articulations are less easily anchored to each other.

Van Minde's (1997) claim that Ambonese Malay has contrastive word stress was refuted by pointing out that under that assumption the only vowel in a five-vowel system /i e a o u/ that can occur without stress in penultimate syllables would be /a/. Incongruously, only non-low vowels would obligatorily attract stress. His data can be explained by assuming that 'unstressed a' is an additional, featureless vowel, which is durationally reduced and realised as an open vocoid, restricted to non-final positions. Its status is thus comparable to that of schwa in Indonesian. The slightly steeper intensity roll-off in the final rhyme of *balalang*, which putatively has prefinal stress, than in that of *tamang*, with putative final stress, can be explained as due to less F0 movement in the final rhyme of *balalang*.

Very small F0 increases in the rising-falling contour in corrective focus as compared to post-focus pronunciations were found for two out of four speakers. There were no durational differences between the two focus conditions for any of the relevant syllables. In effect, this means that Ambonese Malay does not express focus in its prosody. In view of its geographical location, Ambonese Malay conforms to Xu *et al.*'s (2012) claim that post-focus compression is an areal feature of the European-Asian landmass, with the exception of South-east Asia.

An investigation of the intonation system yielded two IP-final melodies, a rise-fall, analysed as $L_\varphi H_\varphi L_\iota$, and a rise, which is analysable as either $L_\varphi H_\varphi H_\iota$ or $L_\varphi H_\varphi$. At the beginning of the IP, a $L_\iota H_\iota$ melody is frequently observed. Its peak may be reduced, while in one-word IPs it is categorically absent. The shape of the final $L_\varphi H_\varphi$ varies from a straight rise at the end of the IP to a sequence of a slow rise and a faster rise. In some cases, the two rising sections are separated by a brief fall at the beginning of the final syllable, but we failed to observe different rise categories in this varied collection. The IP-final melodies have a highly consistent distribution, with $L_\varphi H_\varphi L_\iota$ being used in final declaratives and $L_\varphi H_\varphi(H_\iota)$ in non-final IPs and in final IPs of polar and wh-interrogatives (the non-declarative contour). IP-internally, there were obligatory prosodic breaks after demonstrative words marked by final $L_\varphi H_\varphi$.

We found small duration differences among the three discourse conditions. Lengthening of the final syllable was more extreme in non-final IPs than in final (declarative and interrogative) IPs, showing that this phonetic feature is used to mark non-finality *within* the utterance. This communicatively motivated lengthening in utterance-internal position goes against the well-known pattern of progressive lengthening before higher-ranking boundaries. In addition, penultimate syllables in interrogative IPs are shorter than in other IPs, while the final syllable is longer than in declaratives. We interpret this pattern as a strategy of lengthening final syllables in questions, with a compensatory shortening of IP-medial syllables, as earlier suggested in Abolhasanizadeh *et al.* (2012).

Our investigation only concerns the variety of Malay spoken in Ambon. However, the conclusion that there is no word stress may well apply to other varieties, including Indonesian. The reason why these languages have been described as having penultimate word stress is arguably due to the L2 perception of the intonational melody by listeners with a language background that includes word stress. As Gordon (2014: 111–112) observes, 'a large portion, if not the majority, of the prosodic typology that is currently understood to refer to stress may actually reflect phrasal prominence associated with tonal events occurring at or near phrase edges'. It may be instructive here to consider the other side of this coin. Indonesian listeners will accept that words in European languages have a stressed syllable, one which is phonetically more salient than other syllables, but, without a deeper understanding of the structural connections between word stress, pitch accents and focus, they may be confused about the way in which phonetic salience is distributed over those functions. In the question session after a talk on Dutch word prosody, an Indonesian student, who had earlier learnt that stress in Dutch was cued mainly by pitch and duration features, asked how Dutch listeners kept the duration and pitch cues for emphasis apart from the cues to word stress. Might not a word like /ˈsɛr.vis/ 'Serbian' be confused with /sɛrˈvis/ 'dinner service' when emphasised, or *vice versa*? Beneath the silence with which the question was met by the presenter and her audience, there lay a division which was created by the difference in word prosody. The Dutch students in the audience must have wondered how the question could ever have arisen, while it appeared to be an entirely straightforward one to the Indonesian members of the audience. The answer is that to emphasise a Dutch word, its stressed syllable is hyperarticulated, and emphasising a word therefore creates an even clearer set of cues to stress location than will a non-emphatic pronunciation.

A further indication of the absence of stress in Indonesian is that there is considerable variation in the formulation of stress rules, as documented by Odé (1994) for 13 authors. On the basis of a finding that the F0 peak is always located in the penultimate syllable, Laksman (1994) concludes that Indonesian stress is penultimate without exception, i.e. also when the vowel in the penult is schwa. For Halim (1974: 76), a penultimate syllable with schwa (e.g. *menang* 'to win') will cause stress to move to the final syllable if the word is disyllabic, but to the antepenult if it is trisyllabic or longer (*puteri* 'girl'), unless the antepenultimate syllable also has schwa (*telepon* 'telephone'). Prentice (1987) claims that penultimate schwa will cause the stress to be final in the variety spoken in Java, while elsewhere it is final unless two consonants separate schwa from the vowel in the final syllable. To a large extent, this variation must be due to the way variation in peak location as a function of segmental make-up and language variety gives rise to perceptions of stress on different syllables by speakers of languages that have word stress. A second type of observation in publications by adherents of the stress hypothesis concerns the mobility of stress in derivations. Halim (1974) notes that prefixing a stem like [ˈcat] 'paint' with the agentive marker [məŋ] will leave the stress on the final syllable, as in [məŋˈcat] 'to paint', while other affixations lead to penultimate stress, as in [ˈcatan] 'painting' and [pəngəcaˈtaɲa] 'the painting of it'. Again, these observations may be based on the perception of L2 stress by speakers of languages with stress, or else on a preconceived notion, created by the ubiquitous presence of stress in European languages, that Indonesian must have stress. To quote Hyman (2006: 246), 'if word-stress is so hard to find, maybe it is not there at all … On the other hand, some linguists seem to apply the flip side of this motto: if it sounds like stress, it must be stress!'.

Earlier arguments against the existence of stress were based on behavioural criteria. Odé (1994) finds that listeners had great difficulty identifying the prominent words in spontaneous speech. Not only did the subjects report that they found the task hard, they also indicated different words as prominent and varied considerably in the number of words they considered prominent. (Overall, only 71% of words indicated as 'prominent' had some F0 feature making it stand out from level pitch.) Significantly, Indonesian listeners appear to be largely insensitive to differences in the peak location in Indonesian (van Zanten & van Heuven 1998: 141) or Dutch (van Heuven & Faust 2009) words. Although the task using Dutch words addressed the perception of focus, which may be a skill that is independent of the perception of stress location (§3), these behavioural data do not support the notion of word-based stress that is marked with a pitch accent.

Returning to §2.1 and the four representations in (1), we conclude that Ambonese Malay has no tonal associations, and thus conforms to type (1d). Its final boundary tones are timed to occur in such a way that H is located in the last syllable of IP-final words in non-declarative intonation ($L_\varphi H_\varphi$) and near the beginning of the final syllable in IP-final polysyllabic words in declarative intonation ($L_\varphi H_\varphi L_\iota$). Although we have not investigated IP-final monosyllabic words, we have informally observed placement of the H well within their syllable.

REFERENCES

Abolhasanizadeh, Vahideh, Mahmood Bijankhan & Carlos Gussenhoven (2012). The Persian pitch accent and its retention after the focus. *Lingua* **122**. 1380–1394.

Adelaar, K. Alexander (1992). *Proto Malayic: the reconstruction of its phonology and parts of its lexicon and morphology*. Canberra: Australian National University.

Arvaniti, Amalia, D. Robert Ladd & Ineke Mennen (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *JPh* **26**. 3–25.

Atterer, Michaela & D. Robert Ladd (2004). On the phonetics and phonology of 'segmental anchoring' of *F*0: evidence from German. *JPh* **32**. 177–197.

Beckman, Mary E. (1986). *Stress and non-stress accent*. Dordrecht: Foris.

Boersma, Paul & David Weenink (2011). *Praat: doing phonetics by computer* (version 4.6.09). http://www.praat.org/.

Bolinger, Dwight (1989). *Intonation and its uses: melody in grammar and discourse*. London: Arnold.

Bruce, Gösta (1977). *Swedish word accents in sentence perspective*. Lund: Gleerup.

Cambier-Langeveld, Tina (2000). *Temporal making of accents and boundaries*. PhD dissertation, University of Leiden.

Caspers, J. & Vincent J. van Heuven (1993). Effects of time pressure on the phonetic realization of the Dutch accent-lending pitch rise and fall. *Phonetica* **50**. 161–171.

Cho, Taehong, Jiseung Kim & Sahyang Kim (2013). Preboundary lengthening and preaccentual shortening across syllables in a trisyllabic word in English. *JASA* **133**. 384–390.

Clynes, Adrian & David Deterding (2011). Standard Malay (Brunei). *Journal of the International Phonetic Association* **41**. 259–268.

Cohn, Abigail C. (1989). Stress in Indonesian and bracketing paradoxes. *NLLT* **7**. 167–216.

Collins, James T. (1983). Syntactic change in Ambonese Malay: the possessive construction. In James T. Collins (ed.) *Studies in Malay dialects*. Part 2. Jakarta: Badan Penyelenggara Seri NUSA. 28–41.

Crowhurst, Megan J. & Lev D. Michael (2005). Iterative footing and prominence-driven stress in Nanti (Kampa). *Lg* **81**. 47–95.

D'Imperio, M., R. Bertrand, A. Di Cristo & C. Portes (2007). Investigating phrasing levels in French: is there a difference between nuclear and prenuclear accents? In José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez & María José Cabrera (eds.) *Romance linguistics 2006: selected papers from the 36th Linguistic Symposium on Romance Languages* (*LSRL*). Amsterdam & Philadelphia: Benjamins. 97–110.

Downing, Laura J. (2008). Focus and prominence in Chichewa, Chitumbuka and Durban Zulu. *ZAS Papers in Linguistics* **49**. 47–65.

Dupoux, Emmanuel, Núria Sebastián-Gallés, Eduardo Navarrete & Sharon Peperkamp (2008). Persistent stress 'deafness': the case of French learners of Spanish. *Cognition* **106**. 682–706.

Engelenhoven, Aone van (2002). Concealment, maintenance and renaissance: language and ethnicity in the Moluccan community in the Netherlands. In David Bradley & Maya Bradley (eds.) *Language endangerment and language maintenance*. London: Routledge Curzon. 272–309.

Frota, Sónia (2000). *Prosody and focus in European Portuguese: phonological phrasing and intonation*. New York: Garland.

Goedemans, Rob & Ellen van Zanten (2007). Stress and accent in Indonesian. In Vincent J. van Heuven & Ellen van Zanten (eds.) *Prosody in Indonesian languages*. Utrecht: LOT. 35–62.

Goldsmith, John A. (1976). *Autosegmental phonology*. PhD dissertation, MIT.

Gordon, Matthew (2014). Disentangling stress and pitch-accent: a typology of prominence at different prosodic levels. In Harry van der Hulst (ed.) *Word stress: theoretical and typological issues*. Cambridge: Cambridge University Press. 83–118.

Gordon, Matthew & Pamela Munro (2007). A phonetic study of final vowel lengthening in Chickasaw. *IJAL* **73**. 293–330.

Grice, Martine, D. Robert Ladd & Amalia Arvaniti (2000). On the place of phrase accents in intonational phonology. *Phonology* **17**. 143–185.

Gussenhoven, Carlos (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

Gussenhoven, Carlos & Renske Teeuw (2008). A moraic and a syllabic-H tone in Yucatec Maya. In Esther Herrera Z. & Pedro Martín Butragueño (eds.) *Fonología instrumental: patrones fónicos y variación*. México, D.F.: El Collegio de México. 49–71.

Gut, Ulrike, Stefanie Pillai & Zuraidah Mohd Don (2013). The prosodic marking of information status in Malaysian English. *World Englishes* **32**. 185–197.

Halim, Amran (1974). *Intonation in relation to syntax in bahasa Indonesia*. Jakarta: Djambatan.

Halle, Morris & William Idsardi (1995). General properties of stress and metrical structure. In John A. Goldsmith (ed.) *The handbook of phonological theory*. Oxford: Blackwell. 403–443.

Hart, Johan 't, René Collier & Antonie Cohen (1990). *A perceptual study of intonation: an experimental-phonetic approach to speech melody*. Cambridge: Cambridge University Press.

Heuven, Vincent J. van & Vera Faust (2009). Are Indonesians sensitive to contrastive accentuation below the word level? *Wacana* (*Jurnal Ilmu Pengetahuan Budaya* [*Journal of the Humanities of Indonesia*]) **11**. 226–240.

Heuven, Vincent J. van, Lilie Roosman & Ellen van Zanten (2008). Betawi Malay word prosody. *Lingua* **118**. 1271–1287.

Heuven, Vincent J. van & Ellen van Zanten (2005). Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication* **47**. 87–99.

Himmelmann, Nikolaus (2010). Notes on Waima'a intonational structure. In Michael C. Ewing & Marian Klamer (eds.) *East Nusantara: typological and areal analyses*. Canberra: Pacific Linguistics. 47–69.

Hualde, José Ignacio (2012). Two Basque accentual systems and the notion of pitch-accent language. *Lingua* **122**. 1335–1351.

Hyman, Larry M. (1978). Tone and/or accent. In Donna Jo Napoli (ed.) *Elements of tone, stress and intonation*. Washington: Georgetown University Press. 1–20.

Hyman, Larry M. (2006). Word-prosodic typology. *Phonology* **23**. 225–257.

Hyman, Larry M. (2012). In defense of prosodic typology: a response to Beckman & Venditti. *Linguistic Typology* **16**. 341–385.

Ishihara, Shinichiro (2007). Major Phrase, Focus Intonation and Multiple Spell-Out (MaP, FI, MSO). *The Linguistic Review* **24**. 137–167.

Jun, Sun-Ah (1993). *The phonetics and phonology of Korean prosody*. PhD dissertation, Ohio State University. Published 1996, New York: Garland.

Jun, Sun-Ah (2005). Korean intonational phonology and prosodic transcription. In Sun-Ah Jun (ed.) *Prosodic typology: the phonology of intonation and phrasing*. Oxford: Oxford University Press. 201–229.

Jun, Sun-Ah & Cécile Fougeron (2002). Realizations of accentual phrase in French intonation. *Probus* **14**. 147–172.

Ko, Eon-Suk (2013). A metrical theory of Korean word prosody. *The Linguistic Review* **30**. 79–115.

Kubozono, Haruo (1993). *The organization of Japanese prosody*. Tokyo: Kurosio.

Kügler, Frank & Stavros Skopeteas (2007). On the universality of prosodic reflexes of contrast: the case of Yucatec Maya. In Jürgen Trouvain & William J. Barry (eds.) *Proceedings of the 16th International Congress of Phonetic Sciences*. Saarbrücken: Saarland University. 1025–1028.

Ladd, D. Robert (2008). *Intonational phonology*. 2nd edn. Cambridge: Cambridge University Press.

Ladd, D. Robert, Ineke Mennen & Astrid Schepman (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *JASA* **107**. 2685–2696.

Laksman, Myrna (1994). Location of word stress in Indonesian words and sentences. In Odé *et al*. (1994). 108–139.

Levi, Susannah V. (2005). Acoustic correlates of lexical accent in Turkish. *Journal of the International Phonetic Association* **35**. 73–97.

Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.) (2014). *Ethnologue: languages of the world*. 17th edn. Dallas: SIL International. Available at http://www.ethnologue.com.

Lobanov, Boris M. (1971). Classification of Russian vowels spoken by different speakers. *JASA* **49**. 606–608.

Minde, Don van (1997). *Malayu Ambong: phonology, morphology, syntax*. PhD dissertation, University of Leiden.

Mohd Don, Zuraidah, Gerry Knowles & Janet Yong (2008). How words can be misleading: a study of syllable timing and 'stress' in Malay. *The Linguistics Journal* **3:2**. 66–81.

Moro, Francesca R. (2016). *Dynamics of Ambon Malay: comparing Ambon and the Netherlands*. PhD dissertation, Radboud University Nijmegen.

Odé, Cecilia (1994). On the perception of prominence in Indonesian. In Odé *et al*. (1994). 27–107.

Odé, Cecilia, Vincent J. van Heuven & Ellen van Zanten (eds.) (1994). *Experimental studies of Indonesian prosody*. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië, University of Leiden.

Ohala, John J. & William G. Ewan (1973). Speed of pitch change. *JASA* **53**. 345.

Peperkamp, Sharon & Emmanuel Dupoux (2002). A typological study of stress 'deafness'. In Carlos Gussenhoven & Natasha Warner (eds.) *Laboratory phonology 7*. Berlin & New York: Mouton de Gruyter. 203–240.

Peters, Jörg, Judith Hanssen & Carlos Gussenhoven (2014). The phonetic realization of focus in West Frisian, Low Saxon, High German, and three varieties of Dutch. *JPh* **46**. 185–209.

Pierrehumbert, Janet B. (1980). *The phonology and phonetics of English intonation*. PhD dissertation, MIT. Distributed 1987, Indiana University Linguistics Club.

Pierrehumbert, Janet B. & Mary E. Beckman (1988). *Japanese tone structure*. Cambridge, Mass.: MIT Press.

Post, Brechtje (2000). *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.

Prentice, D. J. (1987). Malay (Indonesian and Malaysian). In Bernard Comrie (ed.) *The world's major languages*. London & Sydney: Croom Helm. 913–935.

Prieto, Pilar, Jan van Santen & Julia Hirschberg (1995). Tonal alignment patterns in Spanish. *JPh* **23**. 429–451.

Remijsen, Bert & Vincent J. van Heuven (2005). Stress, tone and discourse prominence in the Curaçao dialect of Papiamentu. *Phonology* **22**. 205–235.

Rietveld, Toni & Carlos Gussenhoven (1987). Perceived speech rate and intonation. *JPh* **15**. 273–285.

Rietveld, Toni, Joop Kerkhoff & Carlos Gussenhoven (2004). Word prosodic structure and vowel duration in Dutch. *JPh* **32**. 349–371.

Schepman, Astrid, Robin Lickley & D. Robert Ladd (2006). Effects of vowel length and 'right context' on the alignment of Dutch nuclear accents. *JPh* **34**. 1–28.

Shattuck-Hufnagel, Stefanie & Alice Turk (1998). The domain of phrase-final lengthening in English. (Abstract.) *JASA* **103**. 2889.

Silverman, Kim E. A. & Janet B. Pierrehumbert (1990). The timing of prenuclear high accents in English. In John Kingston & Mary E. Beckman (eds.) *Papers in laboratory phonology I: between the grammar and physics of speech*. Cambridge: Cambridge University Press. 72–106.

Sluijter, Agaath M. C. & Vincent J. van Heuven (1996). Spectral balance as an acoustic correlate of linguistic stress. *JASA* **100**. 2471–2485.

Smith, Caroline L. (2002). Prosodic finality and sentence type in French. *Language and Speech* **45**. 141–178.

Soderberg, Craig D. (2014). Cocos Malay. *Journal of the International Phonetic Association* **44**. 103–107.

Son, Robert J. J. H. van & Jan P. H. van Santen (2005). Duration and spectral balance of intervocalic consonants: a case for efficient communication. *Speech Communication* **47**. 100–123.

Stack, Maggie (2005). Word order and intonation in Indonesian. *LSO Working Papers in Linguistics* **5**. 168–182.

Steinhauer, Hein (1994). The Indonesian language situation and linguistics: prospects and possibilities. *Bijdragen tot de Taal-, Land- en Volkenkunde* **150**. 755–784.

Stoel, Ruben (2005). *Focus in Manado Malay: grammar, particles, and intonation*. PhD dissertation, University of Leiden.

Stoel, Ruben (2007). Question intonation in Fataluku. Paper presented at the 5th East Nusantara Conference, Kupang, Indonesia.

Sugahara, Mariko (2003). Downtrends and post-focus intonation in Tokyo Japanese. PhD dissertation, University of Massachusetts, Amherst. Available (April 2016) at http://scholarworks.umass.edu/dissertations/AAI3079597.

Wang, Bei, Ling Wang & Tursun Qadir (2011). Prosodic realization of focus in six languages/dialects in China. In Wai-Sum Lee & Eric Zee (eds.) *Proceedings of the 17th International Congress of Phonetic Sciences, Hong Kong 2011*. Hong Kong: University of Hong Kong. 144–147.

Welby, Pauline (2006). French intonational structure: evidence from tonal alignment. *JPh* **34**. 343–371.

Welby, Pauline & Hélène Lœvenbruck (2006). Anchored down in Anchorage: syllable structure, rate, and segmental anchoring in French. *Italian Journal of Linguistics* **18**. 73–124.

Wordick, F. J. F. (1982). *The Yindjibarndi language*. Canberra: Australian National University.

Xu, Yi (1999). Effects of tone and focus on the formation and alignment of f_0 contours. *JPh* **27**. 55–105.

Xu, Yi & Ching X. Xu (2005). Phonetic realization of focus in English declarative intonation. *JPh* **33**. 159–197.

Xu, Yi, Szu-Wei Chen & Bei Wang (2012). Prosodic focus with and without post-focus compression: a typological divide within the same language family? *The Linguistic Review* **29**. 131–147.

Zanten, Ellen van & Vincent J. van Heuven (1998). Word stress in Indonesian: its communicative relevance. *Bijdragen tot de Taal-, Land- en Volkenkunde* **154**. 129–149.

Zerbian, Sabine (2007). Investigating prosodic focus marking in Northern Sotho. In Enoch Oladé Aboh, Katharina Hartmann & Malthe Zimmermann (eds.) *Focus strategies in African languages: the interaction of focus and grammar in Niger-Congo and Afro-Asiatic*. Berlin & New York: Mouton de Gruyter. 55–79.