

# Some preliminary observations on prosody and information structure in Austronesian languages of Indonesia and East Timor

Nikolaus P. Himmelmann  
Universität zu Köln

## *Abstract*

*This chapter provides a brief overview of what is known about prosody and information structure in the Austronesian languages of Indonesia and East Timor. It emphasizes the fact that the prosodic systems found in these languages appear to differ substantially from the better known systems found in languages such as English and German and finds that to date there is little evidence that prosody plays a major role in conveying information-structural distinctions. Of major import in this regard appears to be the fact that many Austronesian language in the area appear to lack lexical stress as well as lexical tone. Consequently, intonational phrases lack (postlexical) pitch accents, the tonal inventory being restricted to a smallish number of edge tone combinations on the intonational phrase level plus a single boundary tone on the level of intermediate phrases. The chapter concludes with a brief discussion of a major exception to these generalisations, i.e. the group of (Austronesian) West New Guinea languages that show a bewildering variety of tonal and stress-related distinctions.*

*Keywords: prosodic typology, phrase accent, prosodic phrasing, lexical stress, Malayic varieties*

## 1. Introduction

Apart from some varieties of Malay, the prosodic systems found in Indonesian and East Timorese languages have not been investigated in detail to date. Still, from what is known from Malayic varieties<sup>1</sup> and the few studies on other languages that have been published (in particular, Stoel 2006 on Javanese and Himmelmann 2010 on Waima'a) it seems likely that prosodic *prominence* does not have a major role to play in marking information-structural categories. If at all, prosodic *phrasing* may be of relevance in this regard inasmuch as it is not determined by syntactic or processing constraints.

Current ideas on the prosodic marking of information-structural categories, in particular focus and activation status (i.e. the distinction between given, accessible and new discourse referents),

---

<sup>1</sup> See Riesberg et al. (in this volume) for a brief summary of the relevant literature on Malayic varieties.

are based on, and heavily biased towards, what is found in some western European languages, in particular the West Germanic languages English, German and Dutch. From a cross-linguistic point of view, the prosodic marking of information-structural categories in these languages is quite unusual and does not provide a good starting point for investigating the relationship between prosody and information structure in Austronesian languages of Indonesia. Rather, as it is argued here, it will be more productive to start with much simpler assumptions and only take on board more complex prosodic features, if the data require them.

Paradoxically, it will be useful to look at the basic ingredients of West Germanic systems in order to make clear what is meant by “simpler assumptions”. Consequently, Section 2 briefly lists the essential features of a West Germanic system. Section 3 presents a general proposal for a stepwise build-up of prosodic systems, not necessarily confined to the languages under investigation. It starts from the most minimal assumptions about prosodic phrasing and stops at the level of complexity that appears to be widespread in the languages of Indonesia. Sections 4 and 5 introduce complications to the relatively simple prosodic system sketched in Section 3. Section 4 is concerned with the further subdivision of intonational phrases (IPs) into smaller (lower-level) prosodic phrases, while Section 5 briefly looks at languages in eastern Indonesia where highly unusual word-prosodic systems are attested. Section 6 concludes.

The exposition is couched in the terminology and formalisms used in the autosegmental-metrical framework for prosodic analysis (Ladd 2008), and more specifically the Tone and Break Indices (ToBI) framework (Beckman et al. 2005). This framework is chosen because it is the currently most widely used and understood approach to prosodic analysis, and there are a number of cross-linguistic studies which make use of it (cp. for example the two volumes edited by Jun 2005; 2014). But the current argument does not depend on the autosegmental framework and can, in principle, also be expressed in other frameworks for prosodic analysis.

The limitation to Austronesian languages of Indonesia and East Timor is arbitrary in the sense that there are Austronesian languages outside this area that may show similar characteristics, in particular the ones in Brunei and Malaysia. However, the author is not sufficiently familiar with these other languages to be able to make useful observations with regard to western Austronesian languages more generally, not to mention Oceanic languages. This, in fact, also holds for the Indonesian part of Borneo, known as Kalimantan, where the generalizations put forward here possibly do not apply. The prosodic systems found in the languages of the Philippines very likely differ in important regards.<sup>2</sup>

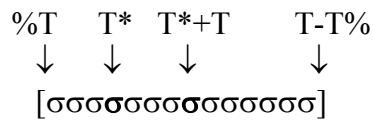
## 2. The West Germanic ‘prototype’

Abstracting away from many details, the phonological structure of an intonation phrase (IP) in West Germanic can be represented as in Figure 1. Using the conventions of the ToBI framework, the T here represents tonal targets which can be either H(igh) or L(ow). These

---

<sup>2</sup> Thus, for example, Blust (2013: 175) notes: “The most distinctive typological feature in the sound systems of Philippine languages is the widespread occurrence of phonemic stress.” The Sangiric languages in northern Sulawesi may show related contrasts. See Himmelmann & Kaufman for a more detailed assessment of this claim.

targets can be anchored either to the edge of an IP (hence edge tones) or to a metrically strong syllable (represented by a bolded  $\sigma$  in Figure 1).<sup>3</sup> The latter are widely referred to as (postlexical) pitch accents. The difference between the two edge tone types, i.e. boundary tones and phrase accents, is further explained in Section 3.



*Metrically-anchored tones*

T\* = pitch accent (monotonal)

T\*+T = pitch accent (bitonal)

*Edge tones*

T- = phrase accent

T% = IP boundary tone (final)

%T = IP boundary tone (initial)

*Figure 1: Phonological Structure of a West Germanic Intonation Phrase*

The occurrence of (postlexical) pitch accents presupposes lexical stress, i.e. the phonologically organized highlighting of a syllable relative to adjacent ones by way of modulating phonetic parameters such as pitch, intensity and duration. It is doubtful that all languages have lexical stress in this sense,<sup>4</sup> and as further detailed in Section 3.2, this appears to be the case for most of the Austronesian languages of Indonesia and East Timor. Note also that even if all languages had lexical stress, it could be the case that the intonational system is organized independently of it. Thus, for example, Lindström & Remijsen (2005) claim for the Papuan language Kuot, spoken on New Ireland, that it is “a language where intonation ignores stress” (2005: 839). In fact, it may be the case that the occurrence of postlexical pitch accents of the West Germanic type is rather rare cross-linguistically.

Most research on intonation, especially in the last three decades, has focussed on (postlexical) pitch accents and their function in marking information-structural categories, in particular focus and activation status (cp. Ladd 2008; Wagner & Watson 2010; Baumann & Kügler 2015; Zimmermann 2016 for recent reviews). While this makes sense with regard to West Germanic and possibly other European languages, it may be counterproductive simply to transfer this model to other languages, as further argued in the following section.

The major emphasis on postlexical pitch accents goes hand in hand with prioritizing information-structural categories among the three main functions that intonational marking may serve. The other two main functions are marking sentence mood (declarative vs. interrogative, etc.) and delimiting phrases on various levels (phrasing or chunking function). While the

<sup>3</sup> Actually, the division is not as straightforward as it is made out to be here. The placement of edge tones, in particular phrase accents, may also make reference to metrically strong syllables, as discussed in the Grice et al. (2000).

<sup>4</sup> The term *stress* is used throughout this chapter in exactly this sense. Note that *stress* as widely used in the literature often includes other notions such as regular rhythmic alternations between strong and weak syllables (foot structure). Claiming that Austronesian languages in the area under discussion usually do not make use of lexical stress hence does not necessarily imply that they do not make use of foot structure or other word-level prosodic distinctions.

sentence mood marking function also has received considerable attention in the investigation of European languages, the phrasing function has not played a very prominent role.<sup>5</sup>

### 3. Start simple!

When looking at prosody in Austronesian languages of Indonesia (and perhaps also in many other parts of the world), it is useful to start with the simplest possible assumptions regarding a prosodic system rather than with the complex model provided by the West Germanic languages.

#### 3.1. Lexical level

With regard to the lexical level, the simplest assumption would be that there is neither lexical stress nor lexical tone (including so-called lexical pitch accents). In the literature on Austronesian languages, as conveniently summarized in van Zanten et al. (2010), it has been widely assumed that these languages, including the ones spoken in Indonesia and East Timor, have lexical stress systems of various kinds, with a strong preference for stress to occur on the penultimate syllable. However, as van Zanten et al. remark in Section 4 of their survey (2010: 99–102; see also van Heuven & van Zanten 2007: 194), there are good reasons to doubt that the prominence phenomena discussed in the literature actually belong to the lexical and not rather to the phrasal level. As we will see further below, what has often been described as regular penultimate stress is in fact the regular occurrence of a rising-falling edge tone combination at the end of intonational phrases. The classic example is Standard Indonesian as spoken in Java, to which a wide variety of stress systems have been attributed, but where there is solid evidence that it actually lacks lexical stress (see Goedemans & van Zanten 2007 for a summary of the relevant research). The work by van Heuven and colleagues shows that this holds true for production (acoustics) as well as perception. With regard to the latter, van Zanten & van Heuven (1998) report a gating experiment which shows that Indonesian listeners were unable to make use of prosodic information in predicting word endings, unlike Dutch listeners who performed much better on the same (Indonesian!) stimuli.

Goedemans & van Zanten (2014) go a step further. Reflecting on the experiences accrued in over two decades of compiling and maintaining a database on stress systems attested in the world's languages (StressTyp) and in particular the fact that in more recent years a number of languages initially classified as having lexical stress had to be reclassified as having no stress, they propose the following list of criteria (or indicators) for descriptions of prominence phenomena where the proposed analysis as lexical stress is doubtful (Goedemans & van Zanten 2014: 88):

1. Stress is reported to vary in different utterances of the same word
2. Stress is reported to be a phenomenon related to phrases

---

<sup>5</sup> Féry (2013) argues that prosodic phrasing is actually more relevant for focus marking than prosodic prominence, hence questioning the emphasis on prosodic prominences in intonation research. However, the marking of information structure (in particular focus) is still considered the primary function of prosody in this line of argument. In line with much of the literature, it is assumed here that the phrasing function of prosody is not necessarily related to information-structural distinction. This does not preclude the possibility that phrasing sometimes may be indicative of focus domains. This issue is taken up again at the end of Section 3.2.

3. Stress is reported to be very weak and unstable
4. Fundamental differences in stress use of various speakers reported
5. The reported stress rule makes no sense in any current metrical theory

The following three indicators are considered not to raise suspicion on their own but to strengthen doubts in case at least one of the above indicators holds true (Goedemans & van Zanten 2014: 88):

6. Position in the intonation contour influences stress location in an unexpected way
7. Numerous exceptions to the rule are reported
8. Only schwa in penultimate position rejects stress

A quick glance at the usually terse sections on stress in many a description of an Austronesian language makes it clear that more often than not several of these indicators apply and that therefore it is not advisable to assume the existence of lexical stress in a particular Austronesian language without further validation.

As for the Austronesian languages of Indonesia, the existence of lexical stress has been properly demonstrated only for very few of them. All of these languages are spoken in the easternmost part of the archipelago, in the Indonesian part of New Guinea and surrounding islands (henceforth simply called Papua in this chapter), as further discussed in Section 5 below.

Note also that it may be the case that there is sufficient evidence for stress-like distinctions in some languages of Sulawesi, in particular Central Sulawesi languages, as briefly discussed in Himmelmann & Kaufman (forthc.). Utsumi (2011) claims that in Bantik, a Sangiric language of northern Sulawesi, lexical pitch accents regularly occur on either the penultimate or the ultimate syllable (hence having a distinctive function).

### **3.2. Postlexical level**

With regard to the postlexical level, the simplest assumption – apart from no use of prosody<sup>6</sup> – would be something like the structure depicted in Figure 2. It shows a string of syllables which is separated from adjacent strings of the same type by melodic and rhythmic cues. Typical rhythmic cues are lengthening the final syllable of the string and pausing. The basic melodic cue pertains to the fact that syllable strings in natural languages are produced with a coherent melody, one string being delimited from the preceding and following ones by the on- and offsets of a coherent pitch contour. Typically, there is a noticeable jump in pitch (up or down) between the offset and the next onset. Additional optional cues include non-modal voice quality such as devoicing at the end of the string or the occurrence of (non-phonemic) glottal stops at the beginning. See Himmelmann et al. (forthc.) for further discussion and experimentation.

---

<sup>6</sup> ‘No prosody’ would mean only purely physiologically conditioned variation in pitch and chunking of speech production. Speech chunks would then be completely determined by breathing requirements, with no regard for content or structure, each chunk probably starting on a relatively high pitch and gradually declining till the end of the unit. Lieberman’s (1967) model of intonation is considerably more refined, but is based on a model of speech physiology which would roughly produce this kind of output, if the speaker were not allowed to control and thereby modulate the basic physiological necessities.

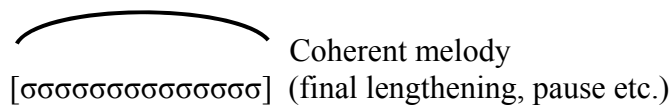


Figure 2: A minimal prosodic structure

The minimal structure in Figure 2 can serve a basic phrasing function inasmuch as the units thus delimited are useful processing units (in terms of planning and/or comprehension). As such, their size would not be primarily determined by the exigencies of airflow management (breathing), but rather by other factors such as semantic and pragmatic considerations of information packaging. Thus, for example, Chafe (1994: 108–119) proposes that intonational phrases (intonation units in his terminology) are designed to introduce one new idea (one piece of new information) at a time into the ongoing discourse.<sup>7</sup> Such a function can easily be served by the units in Figure 2. To serve this function, there is no need to highlight the new information in some way, or to distinguish different types of boundaries. The only requirement is that such a unit never contain more than one piece of new information.<sup>8</sup> An alternative, though partially overlapping, hypothesis for the functional basis of the kind of unit depicted in Figure 2 is the idea that it corresponds to speech acts, i.e. each speech act is packaged as one prosodic unit.<sup>9</sup> A first maximally simple enrichment of the minimal structure in Figure 2 would be the addition of one type of tonal target, i.e. boundary tones, as shown in Figure 3.



%T, T% = boundary tone(s) (initial, final)

Figure 3: Minimal tonal prosodic structure (boundary tones only)

In terms of function, units with the structure in Figure 3 would be capable of marking sentence mood-like distinctions in addition to the very basic information packaging function served by units of the type in Figure 2.

Further enrichment of the basic structure in Figure 3 would consist in adding a so-called phrase accent, i.e. a pitch target anchored to the edge of the phrase, but not necessarily to the very final segment.

<sup>7</sup> A very similar proposal is Pawley & Syder's (2000) one-clause-at-a-time hypothesis.

<sup>8</sup> Obviously, the validity of Chafe's one-new-idea constraint depends on being able to provide an independent and operationalizable definition of how to identify one piece of new information. As this hypothesis only serves as an illustration of what kind of function the minimal structure in Figure 2 has, there is no need here to get into this quite complicated issue.

<sup>9</sup> This idea is suggested by the widely recognized sentence mood marking functions of intonation mentioned above. See Cresti (1996) for further elaboration.

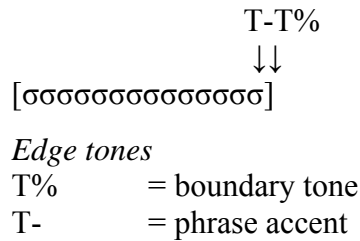


Figure 4: Enriched minimal tonal prosodic structure (edge tones only)

The term *phrase accent* has been used for at least three, prima facie different kinds of phenomena:

1. for a postnuclear prominence occurring at the right edge of an intonational phrase which is part of a complex phrase-final edge tone and typically anchored to a metrically strong syllable (ToBI labels T-T%; cp. Grice, Ladd & Arvaniti 2000; Ladd 2008 142–147);
2. for the boundary tone of a so-called intermediate phrase, i.e. a phrase that is smaller than an IP but larger than a prosodic word (cp. Beckman & Pierrehumbert 1986 and much subsequent work);
3. for a major pitch excursion occurring at the right or – much more rarely – the left edge of an intonational phrase in some languages which otherwise do not appear to make phonological use of pitch changes, i.e. lacking lexical tone distinctions as well as postlexical pitch accents (e.g. French, Indonesian, Waima'a). This pitch excursion may, or may not, be accompanied by extra duration.

It is a matter for further research to determine whether these three phenomena have enough in common to warrant subsumption under a common notion. Alternatively, we are dealing with three phenomena which share the positional feature of edge placement but otherwise have different properties. Part of resolving this issue will be the question of which functions phrase accents serve. Type 2 phrase accents appear to have a boundary-marking function but is this also true for the other two types, both of which are followed by boundary tones proper? And, inasmuch as phrase accents have a boundary-marking function, what is the function of the units thus delimited?

Many Austronesian languages of Indonesia seem to have prosodic units of the basic type shown in Figure 4. The available descriptions report differences with regard to the exact placement of the phrase accents. Common options include:

- ‘free’ variation within a 2-syllable window (e.g. Javanese);
- mostly penultimate syllable, but sometimes also on ultima (e.g. Manado and Papuan Malay, Waima'a).

It is unclear whether these reported differences are actually factual differences. Alternatively, they arise from different analytical procedures and theoretical frameworks. With regard to Ambon Malay, Maskikit & Gussenhoven (2016) provide production evidence for an analysis which considers the configuration depicted in Figure 4 as free floating boundary tones not

anchored to segmental landmarks in the same way as phrase accents in European languages (including in particular French). Although the pre-boundary pitch movement distinguishing this configuration from simple boundary tones as in Figure 3 is often perceived by Western researchers as being anchored to either the penultimate or the ultimate syllable, various measurements indicate that – at least in the case of Ambon Malay – the position of the peak of this pitch movement is highly variable and correlates much less strongly with potential segmental landmarks than typical European postlexical accents.

It is not clear whether this analysis for Ambon Malay also applies to other languages, which to date have not been investigated to the same degree of detail as Ambon Malay. Only such more detailed analyses will show whether the different descriptions reported above correspond to factual differences. In the remainder of this chapter, we will continue to use the term *phrase accent* to refer to the configuration in Figure 4, with the understanding that the details of the analysis, and in particular the specifics of tune-text association, are yet to be worked out.

In the current context, the question of what functions phrase accents may have in marking information structure is a major concern. In languages with relatively fixed word order, it is unlikely that the phrase accent directly marks information-structural categories such as focus, as its position is constrained to a relatively small window (usually two syllables) at the IP edge. The text occurring in this window is often just a single word (or part thereof) or the final syllable of a content word plus a phrase-final particle. Hence the phrase accent occurs on the word/word + particle that happens to be in edge position, regardless of its information-structural status. Inasmuch as speakers do not have a choice with regard to determining which word occurs in edge position, they do not have a choice to determine which word (or syllable) is ‘highlighted’ by the phrase accent. Compare the two examples from Papuan Malay in (1).

- (1) Papuan Malay [elicited]
- |         |              |
|---------|--------------|
| a. baju | b. baju mera |
| shirt   | shirt red    |
|         | ‘red shirt’  |



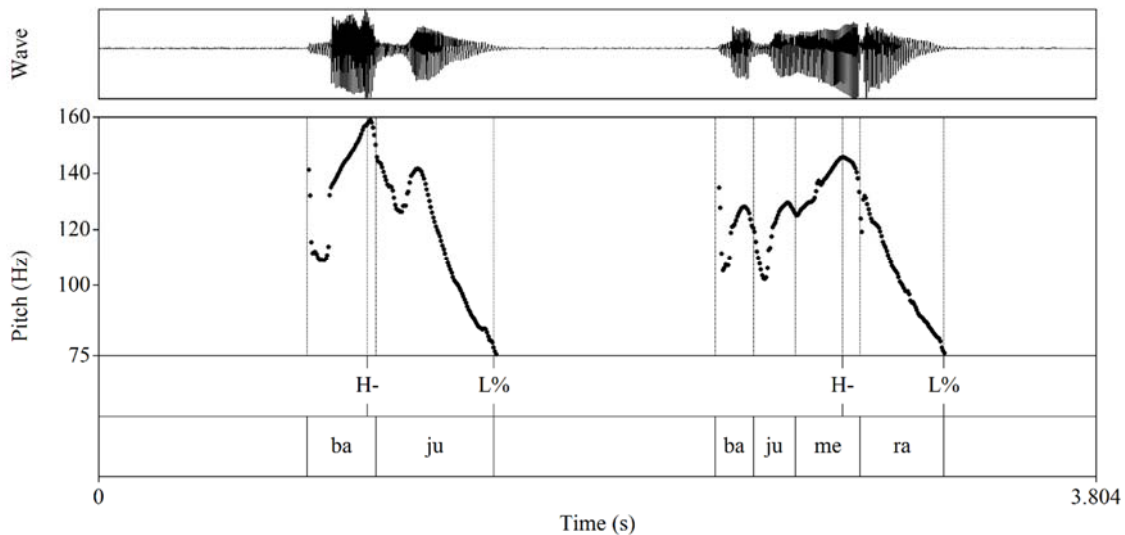


Figure 5: F0 and waveform for ex. (1)a)<sup>10</sup>

As Figure 5 shows, the phrase accent remains at the right edge when another word is added to the phrase. That is, in a phrase such as *baju mera* the phrase accent cannot occur on *baju*, but necessarily occurs on *mera*, because the order of these two constituents cannot be changed. Consequently, in languages which make use of a phrase accent and have fixed word order in at least some phrase types, it is prosodically impossible to mark a difference in (contrastive) focus of the kind seen in English *BLUE car* vs. *blue CAR*. This is nicely illustrated by the following example provided in Stoel (2007), which comes from a corpus of Manado Malay spontaneous narrative speech.

(2) Manado Malay (Stoel 2007: 121)

dorang mo cari tiga: -  
3p ASP look.for three

tiga oto (=)  
three car

cuma: (0.4)  
only

kita cuma dapa satu oto  
1s only get one car

‘(We planned to go to the wedding party by car.) They were looking for three ... three cars. But ... I only got one car.’

<sup>10</sup> All acoustic analyses presented in this chapter were carried out and plotted with PRAAT (Boersma & Weenink 2015).

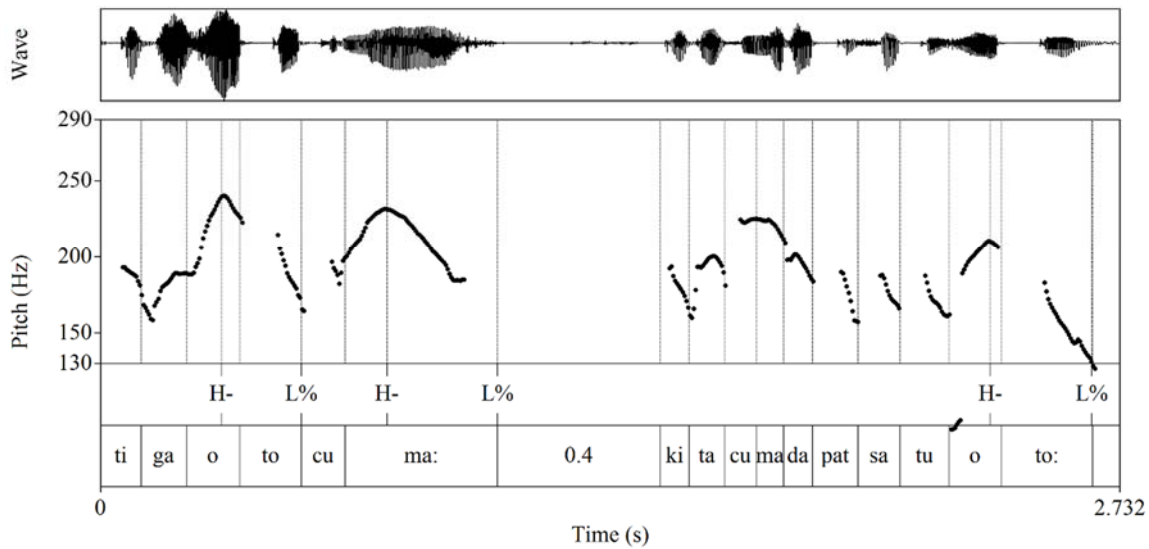


Figure 6: F0 and waveform for example (2)<sup>11</sup>

Despite the fact that in example (2) *tiga* ‘three’ is contrasted with *satu* ‘one’, the phrase accent occurs on *oto* in both instances, because this word takes up the final two syllables in both IPs.

The preceding examples should also make it clear why so many descriptions of Austronesian languages of Indonesia contain the assertion that lexical stress mostly/always occurs on the penultimate syllable. In elicitation, lexical items tend to be produced with the most unmarked declarative intonation pattern, which in many languages includes a phrase accent heard by the researcher to be located in the penultimate syllable, as illustrated by (1a). But when occurring in non-final position in larger structures as in (1b), more often than not no trace of this presumed lexical stress can be detected.

This also brings us back to the prosodic structure shown in Figure 1, which differs from the one in Figure 4 in that it includes metrically-anchored tonal targets in addition to edge tones. As already noted in Section 2, metrical anchoring presupposes lexically-based prominence differences (i.e. lexical stress) which specify the syllables that may serve as anchors for (postlexical) pitch accents. (Postlexical) pitch accents are a major way to prosodically mark the information status of individual words and phrases. Hence, the fact that prosody only plays a limited role in the expression of information structure in many Austronesian languages of Indonesia is related to the lack of lexical stress (lexically encoded prominence differences) in many of these languages.

However, as shown by many languages around the globe including many African languages and Korean,<sup>12</sup> (postlexical) pitch accents may not be the only prosodic means to mark information-structural categories. Prosodic (re- or de-)phrasing may also serve this purpose. A classic and much discussed example comes from the Bantu language Chichewa as analysed by Kanerva (1990). While broad focus utterances tend to be presented in single prosodic phrases,

<sup>11</sup> Only IP edge tones are tonally annotated here. The analysis of the rise across *kita cuma* is discussed in the following section.

<sup>12</sup> See Jun (1998) for a detailed study of Korean dephrasing.

narrow focus on one of the constituents requires the insertion of a prosodic boundary after the focussed word, as can be seen when comparing (3a) with (3b):

(3) Prosodic rephrasing in Chichewa (Kanerva 1990: 98)

- a. (a-na-ményá                      nyu<sup>m</sup>bá    <sup>n</sup>dí    mwáála)<sup>13</sup>  
 1.SBJ-RECENT.PAST-hit      9.house    with    3.rock  
 ‘He hit the house with a rock.’ (answering: What did he do?)
- b. (anaméenya) (nyuú<sup>m</sup>ba) (<sup>n</sup>dí mwáála)  
 ‘He *hit* the house with a rock.’ (answering: What did he do to the house with the rock?)

In fact, Féry (2013: 683) proposes “that the most common prosodic realization of focus can be subsumed typologically under the notion of alignment: a focussed constituent is preferably aligned prosodically with the right or left edge of a prosodic domain the size of either a prosodic phrase or an intonation phrase.”

This view would appear to contradict our assessment above that phrase accents, which occur at the right edge of IPs, do not mark information-structural categories for the simple reason that most Austronesian languages of Indonesia do not freely allow to move focussed words into this position. While speakers thus cannot simply move words around so that they occur in IP-final position, speakers have great freedom in determining the size of an IP. So, in principle, there is the possibility to bring a focussed word into IP-final position by inserting a prosodic boundary behind it (similar to the Chichewa example (3)). However, this possibility does not appear to be systematically used in Austronesian languages of Indonesia. As we will see in the following section, there is some variability with regard to the placement of the phrase accent (at least in Manado Malay), but it is not the case that the word occurring in IP-final position always belongs to the focus domain. Furthermore, most instances of narrow focus are not signalled by inserting an IP boundary. Thus, looking again at example (2), there is no example known to this author where an IP boundary is inserted after a numeral in narrow contrastive focus (here *satu*), separating it from the following non-focussed nominal head of the phrase (here *oto*).

However, IP boundaries are not the only kind of prosodic boundary. In the literature, a fairly heterogeneous group of phrase types is distinguished in between the phonological word and the intonational phrase, these two levels being uncontroversially recognized in all frameworks and widely believed to be found in all languages (but see Schiering et al. 2010). Thus, for example, Kanerva (1990) speaks of *focus phrases* in reference to the smaller kind of phrases seen in example (3b) and argues that these are one level below the intonational phrase in Chichewa. Consequently, to provide a reasonably comprehensive survey of the relation between prosodic phrasing and information structure in Austronesian languages of Indonesia, we have to see whether there is evidence for phrasing units smaller than IPs but larger than phonological words. This will be the topic of the next section.

<sup>13</sup> Morpheme breaks and glossing from Downing & Pompino-Marshall (2013: 651). Numbers refer to noun classes.

## 4. Where things get more complicated 1: Intonation (postlexical prosody)

The structure given in Figure 4 above is not the whole story for the Austronesian languages of Indonesia disposing of this basic type of prosodic structure. In addition to the pitch modulation occurring in a two syllable window at the right edge of an IP, which here is interpreted as a phrase accent, there tend to be further pitch rises earlier on in an IP which are not accounted for by Figure 4. Example (4) from Totoli illustrates.

(4) Totoli [elicited]

[[*i* *Ali*] *anu* *nangaan=ko*]

PN REL AV.RLS:eat=AND

‘Ali was the one who ate it.’ (Answering the question ‘who ate the banana’.)

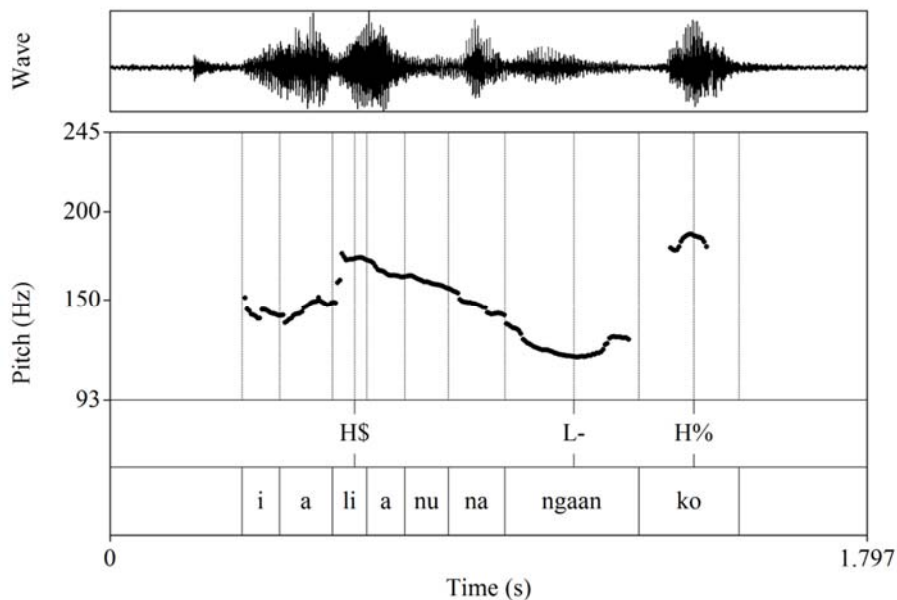


Figure 7: F0 and waveform for example (4)

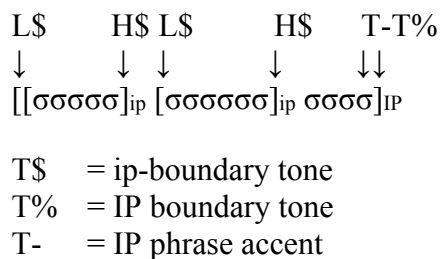
That is, IPs can be prosodically chunked into smaller units which do not interrupt the melodic and rhythmic coherence of the larger IP. These smaller chunks here are called intermediate phrases (ip) and the boundary tone that delimits their right edge is represented by H\$ in the tonal tier. They exhibit the following features throughout the area of investigation:

- The major boundary marker for ips is a H(igh) tone on the unit-final syllable, the peak usually being located at the very end of it. This syllable is not markedly lengthened or otherwise prosodically highlighted in addition to bearing the boundary tone. In theory, the boundary marker for ips could also be a L(ow) tone, but in all the Austronesian languages of Indonesia and East Timor the author is familiar with, it is always H.
- In case there are two or more consecutive ips in an IP, the unit-final Hs tend to be downstepped. However, downstepping does not regularly include IP-final phrase accents and boundary tones, i.e. an IP-final high phrase accent or boundary tone is

often (but not necessarily) higher than any of the preceding H\$ targets. Example (4) illustrates this for a final H-boundary tone (H%).

- No pauses or other rhythmic boundary markers may occur at an ip boundary.
- Similarly, there is no interruption of the overall pitch contour (i.e. no offset-onset phenomena). However, there is always a tonal reset in that the beginning of the following unit always involves a (consecutive) fall to a lower pitch level. This ip-initial low(er) target is often reached within the first syllable of the following ip, but it may also occur somewhat later (2<sup>nd</sup> or even 3<sup>rd</sup> syllable). Cp. the discussion of example (5) below.
- While the rise towards the final H may begin earlier on in the unit, the penultimate syllable is not prosodically highlighted in any special way.

To date, the details of the more fine-grained prosodic structure of the ip in Austronesian languages of Indonesia are not yet well understood. There appears to be considerable variability both within and between languages. Stoel (2005; 2007) observes that in Manado Malay pitch tends to continuously rise across the ip, therefore postulating an initial L\$-boundary tone. Figure 8 provides an abstract representation for the resulting structure.



*Figure 8: The intermediate phrase (ip)*

The analysis shown in Figure 8 is of course only one of a number of different possible analyses for the observed state of affairs. A more standard ToBI analysis, for example, would not make use of a special symbol (\$) for ip boundaries, but analyse these as phrase accents (T-) as well, i.e. conflating the 2<sup>nd</sup> and 3<sup>rd</sup> meanings of ‘pitch accent’ distinguished in reference to Figure 4 above. A major reason for this alternative analysis is Selkirk’s (1984) *Strict Layer Hypothesis* which predicts that each layer on the prosodic hierarchy is exhaustively parsed into constituents of the same type on the next lower level. Thus IPs should exclusively consist of ips. But in in Figure 8, the IP is parsed into two ips plus a third unit of an apparently different status rather than into three ips. The major reason for not following the more mainstream analysis here is that it is not clear that the boundary tones of ips and the phrase accent, which is a part of the edge tone combination marking IPs, really are similar enough to be considered tonal targets of the same type. We come back to this issue at the end of this section after providing more detail on the form and function of ips.

Intermediate phrases with the structure in Figure 8 are also found in East Timorese Waima'a. However, in Waima'a, ips are often essentially flat, the peak of the H\$ rise being followed by a short fall back to the base line, as seen in (5).<sup>14</sup>

<sup>14</sup> This pattern is also found in Manado Malay, e.g. example 30 in Stoel (2007: 130).

(5) Waima'a [pesawat\_41]<sup>15</sup>

ne kara data naha barse ne whaka ige la rihu ne'i wake nin(i)  
 3s want alight if seem 3s fly PTL LOC fog PRX below POSS  
 'if it were about to land, then it should fly below the clouds'

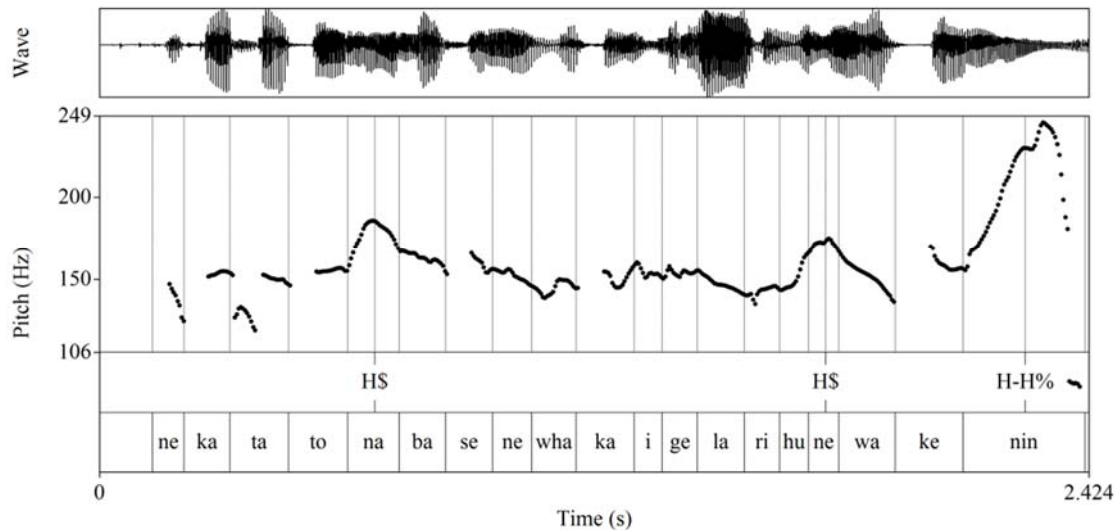


Figure 9: F0 and waveform for example (5)

There are various possibilities for analysing the pitch trajectory in the two ips seen in Figure 9, including also an initial L(ow) boundary tone (the difference between the structure in Figure 8 and the one seen in Figure 9 would then have to be captured by different specifications for phonetic implementation). Alternatively, one could analyse this configuration as involving a final HL\$ boundary tone, with the low target usually being reached on the first or second syllable of the following unit. This is not the place to argue one or the other solution. The important point to keep in mind is that despite considerable variability regarding the details of the pitch contour, all ips have in common that there is an H target in the final syllable. While IPs may also end on a final H target, this final target is immediately preceded by another pitch target – the phrase accent – which is not found in ips.

As seen in example (5), an ip may be quite long and span a number of words and even complete (subordinate) clauses. It is thus clearly larger than the units analysed as *phonological words* and *accentual phrases* in the literature. Instead of *intermediary phrase*, the units under discussion could also be called *prosodic phrases* or *phonological phrases*.<sup>16</sup> All of these three terms are

<sup>15</sup> Elision of syllables is common in natural Waima'a discourse. In (5), for example, the initial conditional clause *ne kara data naha* is shortened to *ne katatona*. The regularities of syllable elision and concomitant sound changes are, however, not yet understood.

<sup>16</sup> The latter term is used by Stoel (2007) who uses a slightly different analytical framework but his *phonological phrase* clearly matches what is labelled *intermediate phrase* here. To wit: "There are two prosodic constituents that are particularly relevant for the description of Manado Malay intonation: the Phonological Phrase (PhP) and the Intonation Phrase (IP). The PhP is defined here as a prosodic constituent that begins and/or ends with an edge tone. The IP is defined as a prosodic constituent that contains one or more PhPs, but no more than one pitch accent [i.e. phrase accent in the terminology used here, NPH]. IPs do not have any associated edge tones. A PhP corresponds roughly to an XP at the syntactic level, and an IP to a clause. An IP may be followed by a short pause, while a PhP may not. It is characteristic for Manado Malay that the accent-bearing unit is a relatively high-level unit, whereas in many European languages, not only the IP, but also the PhP, may have more than one accent" (Stoel 2007: 121).

used in very different ways in the literature and it is not clear that the units thus labelled are actually instances of a common general type. Hence, *intermediate phrase* here specifically applies to the kinds of units defined at the beginning of this section. It is a matter for further research to determine whether these units have essential features in common with units referred to by the same label in other languages (English and Japanese, for example, as analysed in Beckman & Pierrehumbert 1986).

As a general rule, the size of ips is determined by syntax. That is, their boundaries usually match syntactic constituents such as NPs, VPs or clauses, with the possibility that heavy constituents such as NPs which include a relative clause are chunked into two ips. A fully worked out analysis of ip chunking is not yet available for any Austronesian language of Indonesia. The following patterns have been observed in the author's data for Totoli and Waima'a.

It is quite common that the initial word in an IP is chunked as an ip if it is a question word, a conjunction, or an imperative marker. In (6), the initial word is the question word *isei* 'who', in (7) it is the conjunction *tamba* 'because', and in (8) it is the negative imperative marker *deme'e* 'don't'. For such relatively short ips it appears to be the rule that pitch rises continuously throughout the ip, as seen in the corresponding figures.

(6) Totoli [elicited]

isei      nangaanko              saginna  
 who?    AV.RLS:eat:AND          banana:3s.POSS  
 'Who ate his/her banana?'

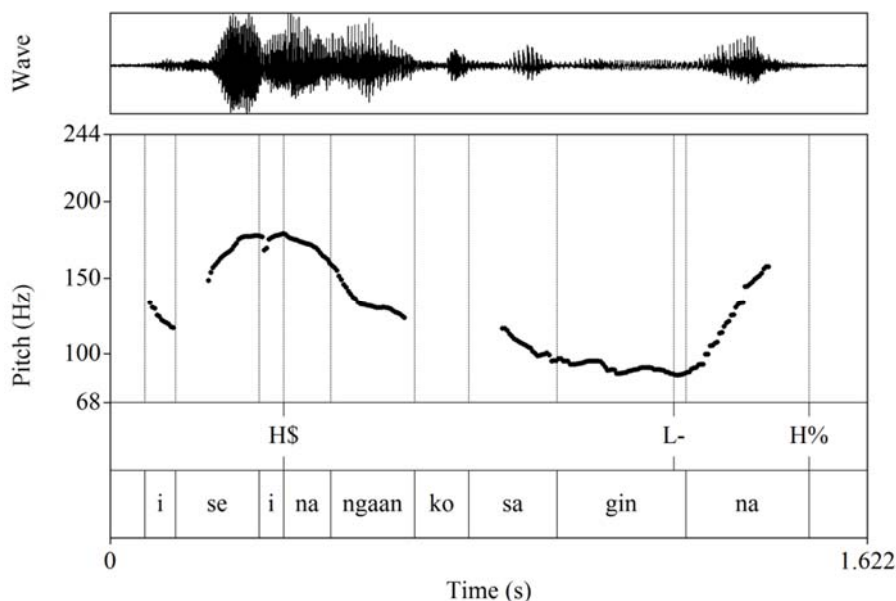


Figure 10: F0 and waveform for example (6)

(7) Waima'a [elicited]

tamba    ai-sa'i    aku    bira  
 because yesterday 1s    sick  
 'because I was sick yesterday'

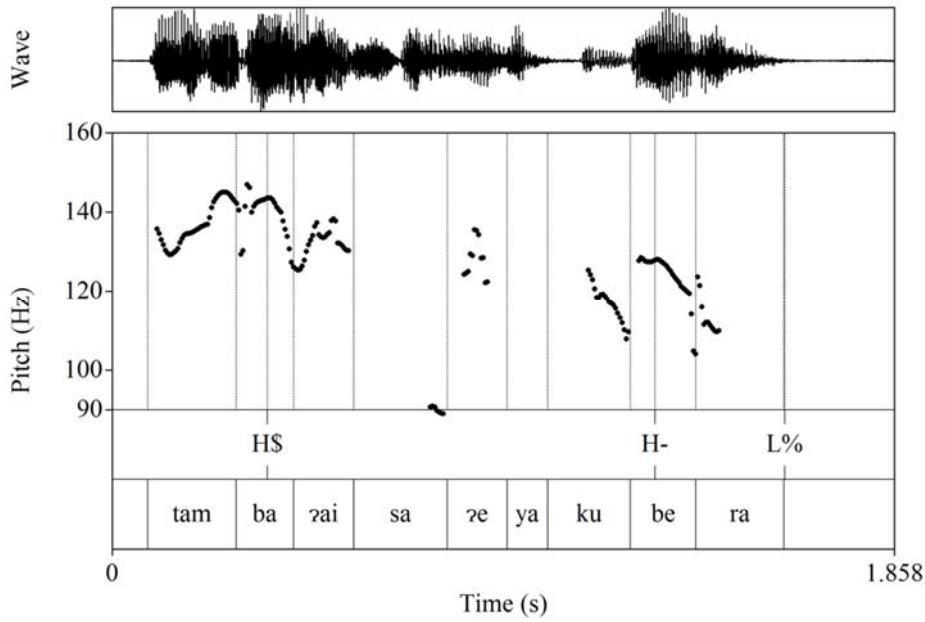


Figure 11: F0 and waveform for example (7)

- (8) Waima'a [elicited]  
deme'e sike mala ne'i  
NEG.IMP touch box PRX  
'Don't touch this box.'

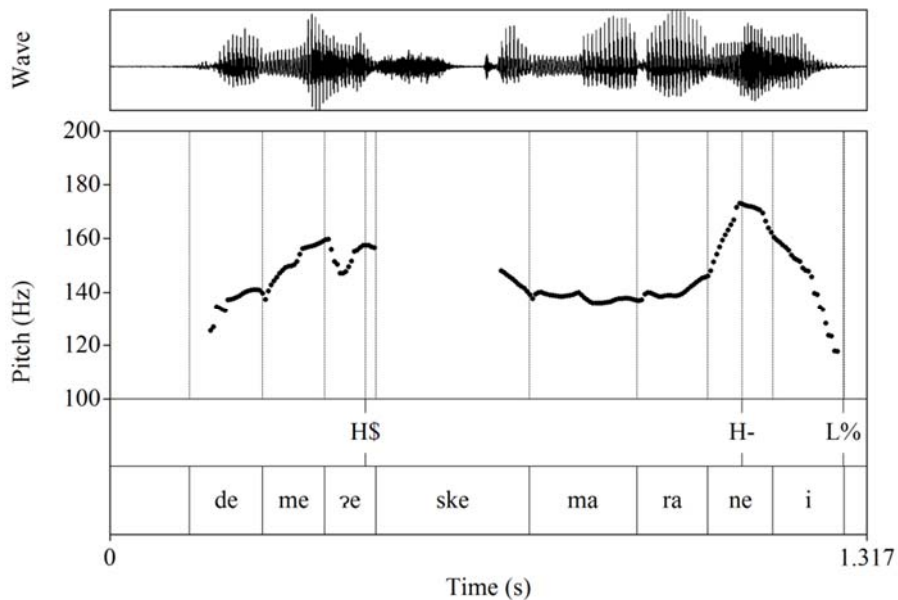


Figure 12: F0 and waveform for example (8)

Similarly, initial adverbial phrases such as *pas la n'iki* 'right here' in (9) form their own ip. In this example, the pronominal subject and the verb (*ne soke* 'he crashes') also form an ip of their own.



- (9) Waima'a [pearcarlito\_101]  
 pas la n'iki ne soke la watu see  
 right LOC here 3s crash LOC stone one  
 'and right then he crashes into a stone.'

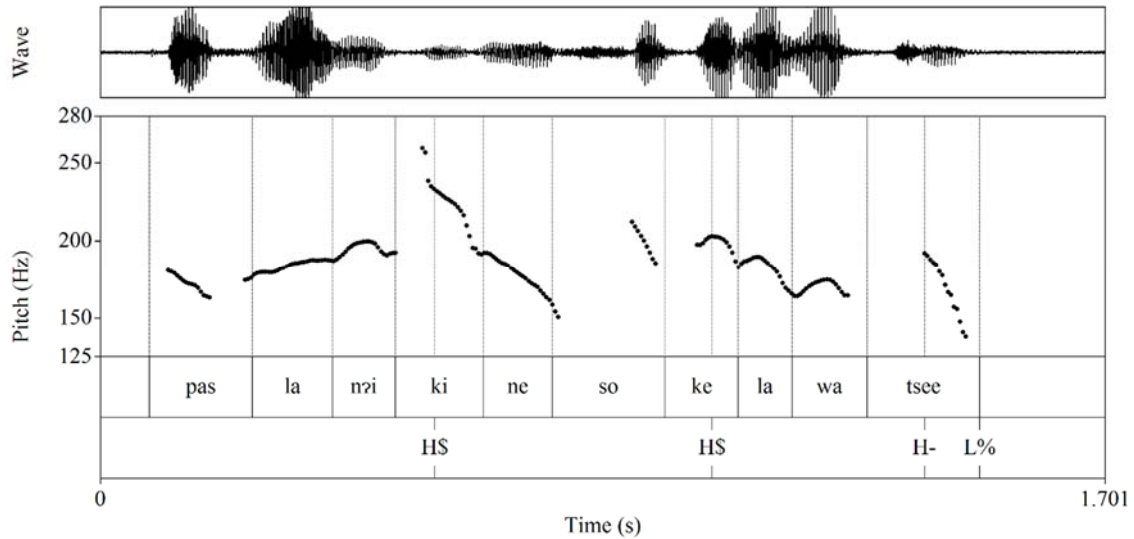


Figure 13: F0 and waveform for example (9)

Initial subject or topic NPs also tend to be phrased as ips. In (10), *kii ba'an ke* 'the old man' exemplifies this preference. As opposed to the preceding example, here the following VP is not phrased independently but forms one longish final phrase together with the local adjunct *la kai-oo kai-oo ta* 'in the tree tops'. To date, the phrasing regularities for VPs and constituents following VPs are not yet well understood.

- (10) Waima'a [pearcarlito\_79]  
 kii ba'an ke uhu naga kai-wuo la kai-oo kai-oo ta  
 person HON-N DEM pick CONT fruit LOC tree-top tree-top DIST  
 'the old man just keeps on picking fruits in the tree tops'

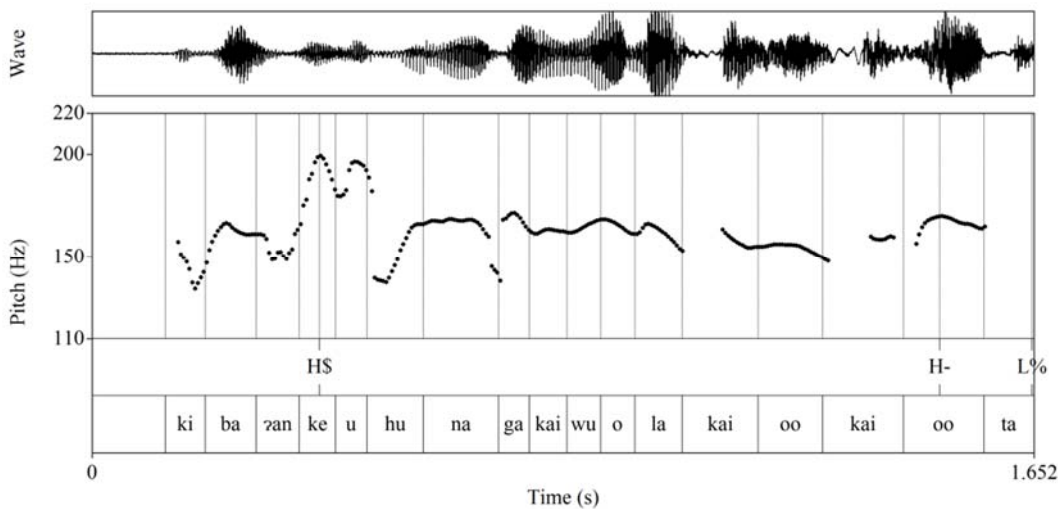


Figure 14: F0 and waveform for example (10)

Finally, there is a strong tendency to phrase clauses separately in case a single IP contains more than one clause. This holds for subordinate clauses, as already exemplified with example (5). Example (11) shows that the length of subordinate and main clauses does not appear to play a major role in this regard, i.e. ip boundaries are inserted even when the overall IP is relatively short.

- (11) Waima'a [pearcarlito\_103]  
 soke watu see ne lo'i  
 crash stone one 3s fall  
 'crashing into a stone, he falls'

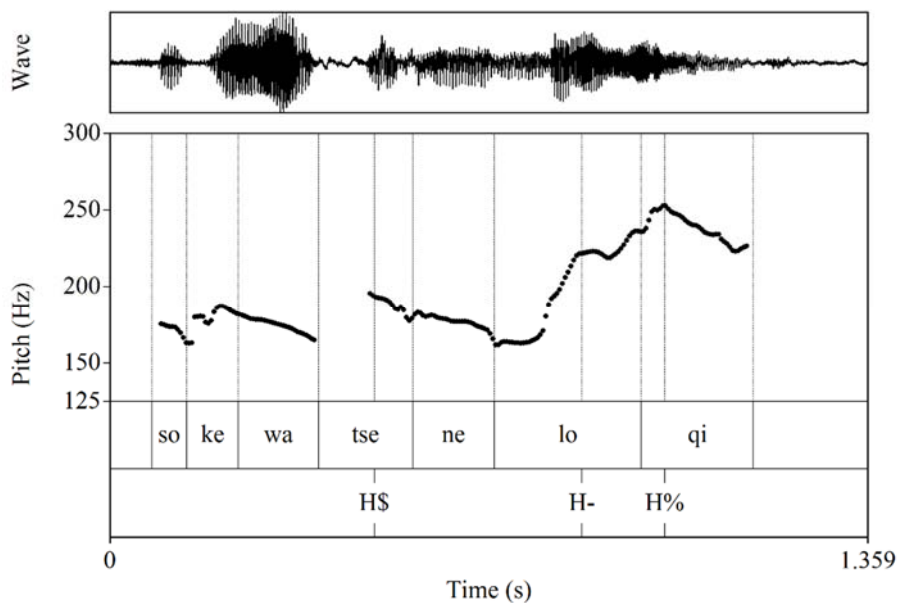


Figure 15: F0 and waveform for example (11)

Two clauses may also be combined in a single IP if they are parallel in structure. In example (12), the parallelism is emphasized by the preverbal particle *oo* 'too, as well' in both clauses.

- (12) Waima'a [pearcarlito\_143]  
 wuo-telu ana oo laka ne oo laka  
 CLF-three DIM too go 3s too go  
 'the three of them walk off, (and) he also walks off'

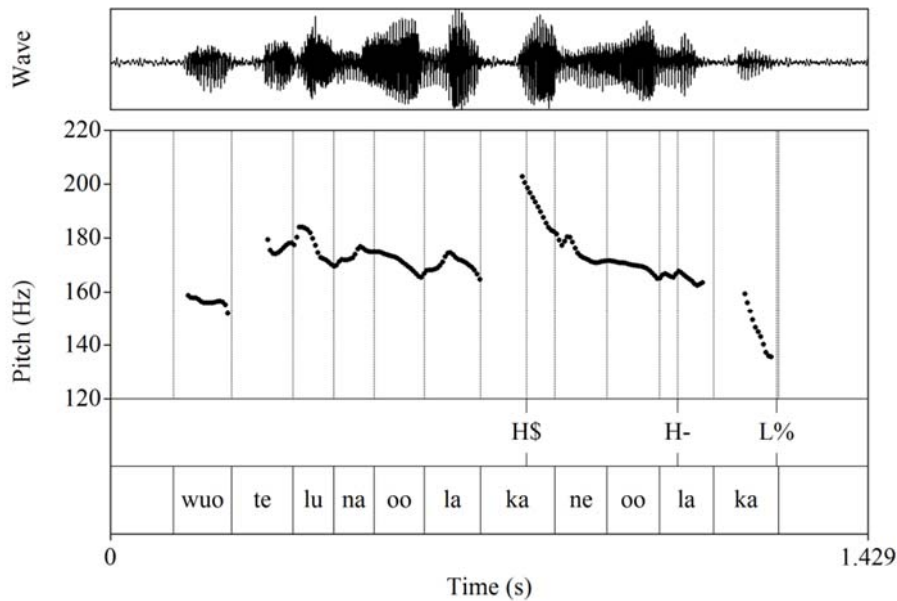


Figure 16: F0 and waveform for example (12)

The preceding examples illustrate strong tendencies where it seems reasonable to account for ip chunking in terms of syntactic structure. For these examples, it is not immediately obvious that information structure has a role to play. Nevertheless, it is also clear that a purely syntactic account will not suffice in all instances as none of these tendencies is actually obligatory (i.e. subordinate clauses do not have to be phrased as separate ips, for example). It may thus very well turn out that some aspects of ip chunking are sensitive to information-structural factors.

A case in point are VPs and constituents following them. For this syntactic configuration, no clear syntactic tendencies have been detected so far (cp. examples (9) and (10) above). A particularly conspicuous example in this regard is example (5) where the final H\$-boundary splits a complex PP/NP into two parts: *la rihu ne'i* ‘LOC this fog(s)’ and *wake nini* ‘below POSS’. The first part of this PP, which contains the preposed possessor NP ‘this fog’, is chunked with the remainder of the clause (*barse ne whaka ige* ‘it should fly’ (= modal particle + subject pronoun + verb + particle). The second part consists of what is formally the head of the complex NP, the possessum ‘below’ (more literally: ‘its underside’). Here, it may be speculated that ip chunking puts special emphasis on ‘below’, as flying below the fog may have prevented the plane crash reported in the narrative from which this segment is taken.

Another example for the possible influence of information structure on ip chunking comes from Manado Malay. Stoel (2007) observes the following possibility for prosodically highlighting a word which does not occur in IP-final position. In Manado Malay, it is possible to add one, and exactly one, ip after the word carrying the phrase accent. This post-accentual ip is characterized by a compressed pitch range, usually being almost flat on a low tonal level, with the possibility of ending with a smallish final fall. Example (13) can be produced in the two different prosodic shapes presented in Figure 17 and Figure 18, respectively (both elicited).<sup>17</sup>

<sup>17</sup> The following three figures are directly quoted from Stoel (2007), hence the difference in layout and annotation detail. No sound files were available to the present author.

## (13) Manado Malay verb focus (Stoel 2007: 126)

dia da bamara pa Weni  
 3s ASP angry at Weni  
 ‘She is angry at Weni.’

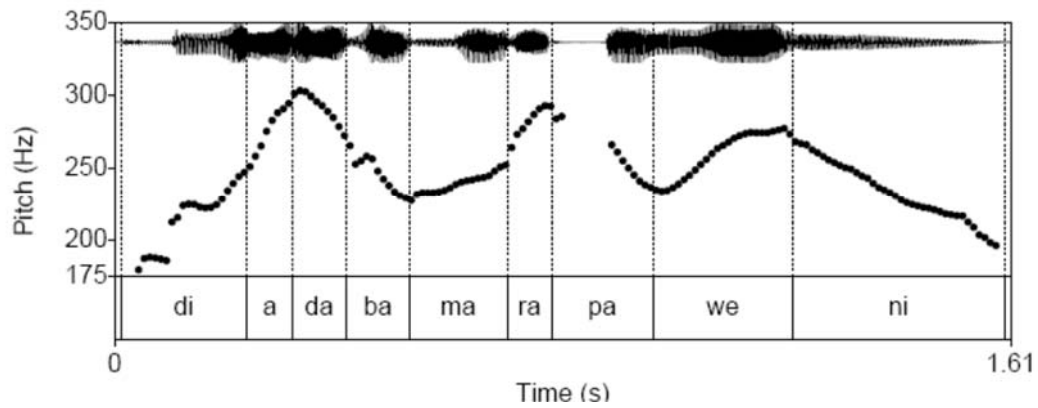


Figure 17: Example (13) with unmarked prosodic phrasing (= Figure 5 in Stoel 2007:126)

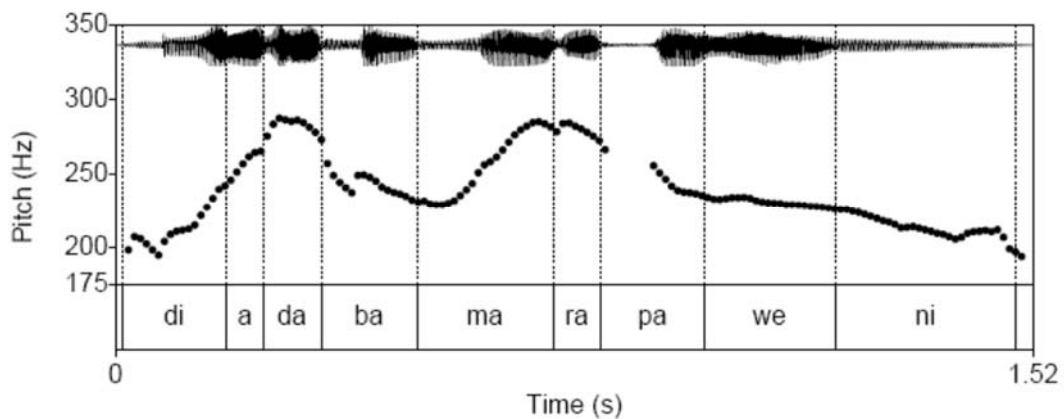


Figure 18: Example (13) with verb focus (= Figure 6 in Stoel 2007: 127)

Note that in both instances, according to the analysis proposed by Stoel, the sentence is chunked into three ips, i.e. [dia] [da bamara] [pa Weni]. In Figure 17, the first two units are characterized by the rise from a low initial target to a H tone on the final syllable of the ip, as is typical for ips. In the third unit, *pa Weni*, there is a phrase accent on the penultimate syllable *We*, followed by a fall in the final syllable, i.e. the typical pattern marking the end of an IP. Information-structurally this is a relatively neutral rendering, compatible with broad focus and object focus contexts.

In Figure 18, on the other hand, the phrase accent occurs on the penultimate syllable of the second unit [*da baMara*], followed by a fall which continues throughout the third unit [*pa Weni*]. The post-accentual ip is analysed by Stoel as involving only a single L boundary tone at the right edge, as opposed to the continuous L to H-rise typical for ips preceding the pitch accent. He calls it “encliticized” (Stoel 2007 121), as it appears to contain (usually already) well-known information that is added to a unit which in itself is already complete. The rendering in Figure 18 is appropriate for contexts involving a narrow focus on the predicate. Example (14) illustrates narrow focus on the predicate from spontaneous speech.

(14) Manado Malay verb focus (Stoel 2007: 126)

da orang cari pa ngana  
 ASP person look.for at 2s  
 ‘Somebody was looking for you.’

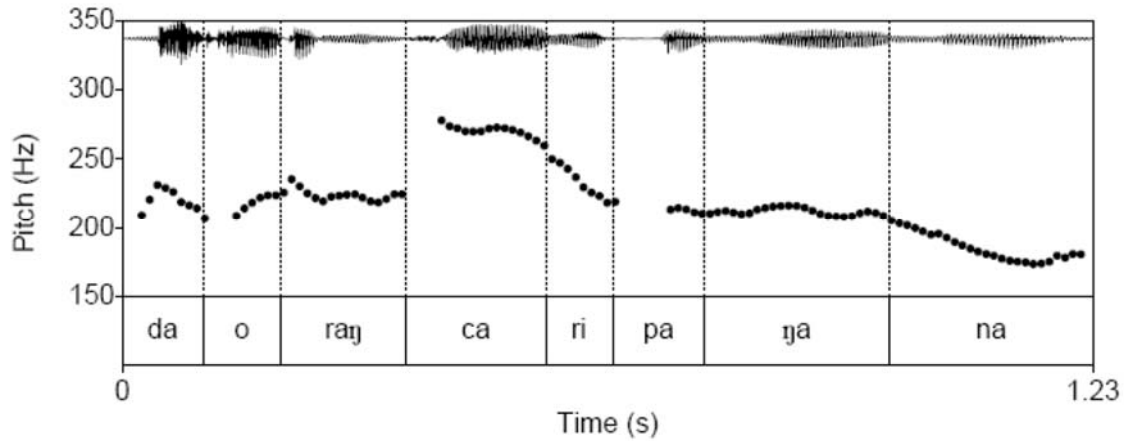


Figure 19: Example (14) with verb focus (= Figure 8 in Stoel 2007: 128)

Finally, there are examples where the *lack* of expected ip chunking appears to be influenced by information-structural considerations. This is attested in utterances where the utterance-final word carries contrastive focus as in *she does not like RED, she likes GREEN*. The Waima'a example in (15) illustrates.

(15) Waima'a [elicited]

ne de kara haru lumu  
 3s NEG like shirt green  
 ‘S/he doesn’t like the green shirt.’

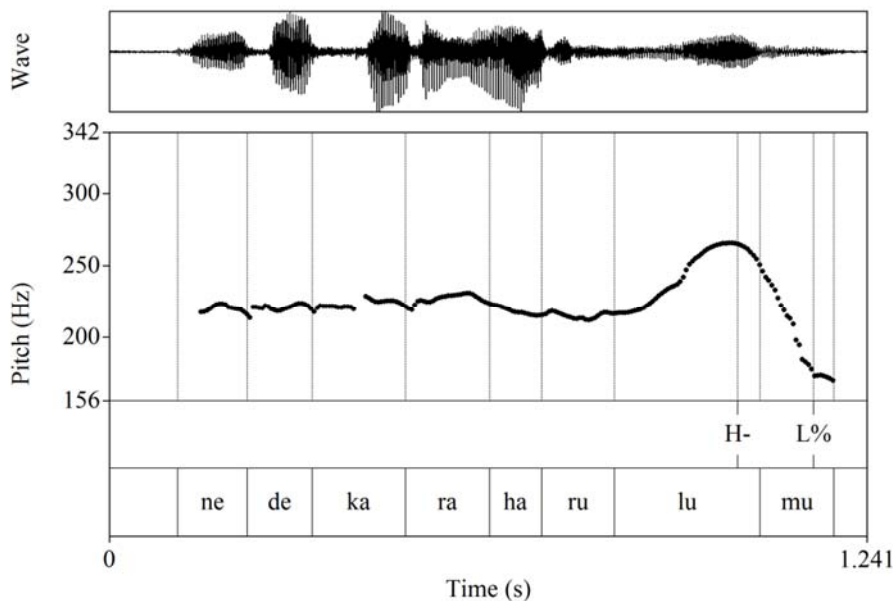


Figure 20: F0 and waveform for example (15)

However, that the lack of ip chunking may be a way to convey contrastive focus on the word it appears on is only a conjecture that needs more testing and research. Note that even if this

conjecture turns out to be true, it would not allow for a consistent marking of contrastively focussed items because the phrase accent is confined to the phrase-final word. Hence, in the current example it would not be possible to contrast the *SHIRT green* with the *SKIRT green* because the word order requires the adjective to follow the noun (cp. example (2) above).

To sum up this section, the intonational structure in Austronesian languages of Indonesia and East Timor appears to be more complicated than suggested by Figure 4, because there is an additional level of intonational structure below the IP, i.e. the intermediate phrase. The regularities obtaining for ips are not yet well understood. This concerns both their tonal structure and the factors determining ip boundaries. The major tonal target in an ip is an H tone which consistently appears on the final syllable. What is not clear yet, is whether there is also an initial tonal target (in at least some of the languages in the region) and how to analyse the different trajectories for reaching the final H\$ (continuous rise throughout the ip vs. rise over the last few syllables of the ip vs. steep rise on the final syllable only).

A further unresolved issue – briefly mentioned in connection with Figure 8 – is the question of how to analyse the (usually) final segment of an IP which follows the last H\$ and contains the phrase accent and final boundary tone. Should this segment also be analysed as an ip (as the strict layer hypothesis (Selkirk 1984) would demand)? But then, how can one explain the fact that tonal targets in this segment can be much more varied than in pre-final ips and obey different alignment regularities than the final boundary tone of ips? One option is to assume that ip-level tones are deleted at IP boundaries and overwritten by the higher-level IP edge tones, as proposed by Khan & Khan (2014: 83) for Bengali (similar proposals have also been made for other languages, as pointed out by Khan & Khan). However, it is not clear what kind of empirical evidence would support such an analysis. Furthermore, inasmuch as IP-level boundary tones are of a different type and do not include ip-level tones as a constituent (as in the model of Beckman & Pierrehumbert 1986, for example), this can still be seen as a violation of the strict layer hypothesis.

As for boundary-determining factors, it is clear that ip boundaries generally obey major syntactic phrase boundaries such as NP, VP and PP. But exceptions occur, as seen in example (5). Furthermore, there appears to be a general tendency to phrase separately preverbal constituents of various types, including nominal and prepositional phrases and single word-constituents such as question words and conjunctions. The latter clearly show that constituent length is not a primary factor in ip chunking. There is also a clear tendency to phrase clauses separately if they occur in a single IP. No regularities for the VP and following constituents have been discovered so far. Similarly, it is not clear why expected ip boundaries are occasionally missing. It may well be the case that these types of examples involve information-structural influences on ip-phrasing.

## **5. Where things get more complicated 2: Tone and stress (lexical prosody)**

In Section 3.1, it was noted that there is little or no evidence for word-prosodic distinctions in many Austronesian languages of Indonesia. In particular, there is little evidence for lexical

stress, which is of primary concern here. The present section serves to briefly point out that, though comparatively rare, the Austronesian languages of Indonesia may show considerable prosodic complexity on the word level, in particular in the eastern parts of Indonesia.<sup>18</sup> Furthermore, and more importantly, it seems that wherever there are word-prosodic distinctions they tend to be highly constrained in terms of (a) the position within the word or phrase, (b) the types of contrast allowed for, and (c) the interaction with other prosodic subsystems.

Tonal contrasts have been reported primarily for a number of Austronesian West New Guinea languages (spoken in Indonesia's two easternmost provinces Papua and Papua Barat). Remijsen (2001) and Kamholz (2014: Chapter 5) provide succinct surveys as to what is known about tone in Austronesian Papua.<sup>19</sup> The languages analysed so far show a broad variety of tonal systems. Magey Matbat, spoken on Misol, one of the Raja Ampat islands, is analysed by Remijsen (2007) as a syllable tone language with six different tones. A large part of the Magey Matbat vocabulary appears to be monosyllabic, but bi- and trisyllabic words also occur. All monosyllabic words bear tone. From the few examples provided in Remijsen (2007) it appears that at least one syllable in polysyllabic words is toneless, but the position of tone-bearing syllables is not predictable. This contrasts with Moor, a language spoken in southern Cenderawasih Bay, which is analysed by Kamholz (2014: 101–106) as disposing of four tonal patterns. Tonal marking in Moor is largely confined to the final two syllables. More importantly, and rather unusually for a tone language, “tones are realized only on phrase-final words” (Kamholz 2014: 102). It seems likely that tonal marking here interacts with the phrase-accent+boundary tone typical for IPs in Austronesian languages of Indonesia, a topic not addressed by Kamholz.

A particularly complex – and cross-linguistically unusual – word-prosodic system is found in Ma'ya, the largest of the Raja Ampat languages. Remijsen (2001; 2002) makes a convincing case for an analysis in terms of both lexical stress and lexical tone. There are three tonal contrasts which, however, are confined to the final syllable. In addition, lexical bases differ in whether they are stressed on the penultimate or ultimate syllable. That is, there are minimal pairs which differ only with regard to tone, e.g. *sa*<sup>12</sup> ‘to sweep’ vs. *sa*<sup>3</sup> ‘to climb’ vs. toneless *sa* ‘one’ (Remijsen 2002: 596). And there are minimal pairs differing only in stress, e.g. *'mana*<sup>3</sup> ‘light (of weight)’ vs. *ma'na*<sup>3</sup> ‘grease’ (Remijsen 2002: 600). Note that Remijsen (2002: 602–610) provides detailed acoustic evidence for the proposed stress difference, which includes not only duration measures, but also differences in vowel quality and spectral balance.

Unfortunately, neither Remijsen nor Kamholz discuss postlexical prosody in the languages they investigate. Hence it is unclear whether the word-level prosodies interact with postlexical tonal marking. Consequently, it is also unclear whether word-level prosodies have any role to play in conveying information-structural distinctions.

---

<sup>18</sup> This is not intended to be a comprehensive survey of word prosodies in Austronesian languages of Indonesia, for which see van Zanten et al. (2010) (to be read with the caveats found in van Heuven & van Zanten 2007 and Goedemans & van Zanten 2014).

<sup>19</sup> It is quite likely that there are more tone languages in this area than listed in Kamholz (2014). Among the Raja Empat languages, Ambel probably has tone (Arnold pc). In Yapen, current work by the author points to tone in Woi.

## 6. Conclusion

In this chapter, it has been argued that prosodic systems in the Austronesian languages of Indonesia work somewhat differently from what is known from West Germanic languages (on which most current prosodic theory is based), and should be approached accordingly. Most importantly perhaps, there is little evidence for lexical stress in many of these languages. And even if there are stress-like distinctions, it should not be presumed that these interact with the intonational system in a way similar to what has been found for Germanic languages (i.e. ‘stressed’ syllables do not necessarily serve as anchors for intonational tonal targets, recall Lindström and Remijsen’s (2005) “a language where intonation ignores stress”). Intonational targets appear to be placed with reference to the boundaries of prosodic units. Two types of units need to be distinguished, the higher-level Intonational Phrase (IP) and the lower level intermediate phrase (ip). Regularities for phrasing on both levels are not yet very well understood. Information-structural factors such as focus and activation status may play a role here, but it is unlikely that they suffice for a full account of prosodic phrasing on either level.

## Sources

Data on Totoli, Waima'a and Papuan Malay are from language documentations the author has participated in. They are all available in the *DoBeS Archive* (<http://dobes.mpi.nl/>) in the following subcollections:

Belo, Maurício C.A, John Bowden, John Hajek, Nikolaus P. Himmelmann & Alexandre V. Tilman. 2002-2006. *DoBeS Waima'a Documentation*. DoBeS Archive MPI Nijmegen.

Leto, Claudia, Winarno S. Alamudi, Nikolaus P. Himmelmann, Jani Kuhnt - Saptodewo, Sonja Riesberg & Hasan Basri. 2005-2010. *DoBeS Totoli Documentation*, DoBeS Archive MPI Nijmegen.

## Acknowledgements

I gratefully acknowledge very pertinent and helpful comments by two anonymous reviewers who thoroughly read and commented on the draft of this chapter. Very special thanks to Gabriele Schwiertz and Patrick Coenen for help with the figures and the layout. I am also very grateful for a number of grants by the Volkswagen foundation I have had the privilege to receive since 2002 and without which the research on which this chapter is based would not have been possible. Many thanks to my colleagues and friends at the Center for Endangered Language Documentation (CELD) in Manokwari for much help and support in trying to understand the languages of West Papua. Finally, I am grateful to the Center of Excellence for the Dynamics of Language (CoEDL) at the Australian National University for providing a congenial environment to carry out the major revisions of this chapter.

## Abbreviations

Conventions in the examples: each line is one IP; (=) indicates latching; pause length is given in ( ); - marks truncated IPs; < > surround false starts.

Glosses for grammatical categories: AND – andative, ASP – aspectual particle, AV – actor voice, CLF – classifier, CONT – continuative (aspect), DEM – demonstrative, DET – determiner, DIM –



diminutive, DIST – distal (demonstrative), HON – honorific, IMP – imperative, LOC – locative (preposition), N – phrase-final nasal in Waima'a with attributive function, NEG – negation, p – plural pronoun, POSS – possessive, PN – personal name marker, PRX – proximal (demonstrative), PTL – particle, REL – relative marker, RLS – realis, s – singular pronoun, and SBJ – subject.

## References

- Baumann, Stefan & Kügler, Frank (eds.). 2015. *Prosody and information status in typological perspective*. Special Issue, *Lingua* 165.
- Beckman, Mary E. & Pierrehumbert, Janet B. 1986. Intonational structure in Japanese and English. *Phonology Yearbook* 3. 255–309.
- Beckman, Mary E. & Hirschberg, Julia & Shattuck-Hufnagel, Stefanie. 2005. The original ToBI system and the evolution of the ToBI framework. In Jun, Sun-Ah (ed.), *Prosodic typology: the phonology of intonation and phrasing*, 9–54. Oxford: Oxford University Press.
- Blust, Robert A. 2013. *The Austronesian Languages*. Revised Edition. Canberra: Asia-Pacific Linguistics. (<http://hdl.handle.net/1885/10191>)
- Boersma, Paul & Weenink, David. 2015. *Praat: doing phonetics by computer*. Version 5.4.09). (<http://www.praat.org/>) (Accessed 2015-09)
- Chafe, Wallace L. 1994. *Discourse, consciousness, and time*. Chicago: The University of Chicago Press.
- Cresti, Emanuela. 1996. Speech acts units and informational units. In Fava, Elisabetta (ed.), *Speech acts and linguistic research*, 89-107. Padova: Edizioni Nemo.
- Downing, Laura J. & Pompino-Marschall, Bernd. 2013. The focus prosody of Chichewa and the stress-focus constraint: a response to Samek-Lodovici (2005). *Natural Language and Linguistic Theory* 31. 647–681.
- Féry, Caroline. 2013. Focus as prosodic alignment. *Natural Language and Linguistic Theory* 31. 683–734.
- Goedemans, Robert W.N. & van Zanten, Ellen. 2007. Stress and accent in Indonesian. In van Heuven, Vincent J. & van Zanten, Ellen (eds.), *Prosody in Austronesian languages of Indonesia*, 35–62. Utrecht: LOT.
- Goedemans, Robert W.N. & van Zanten, Ellen. 2014. No Stress Typology. In: Caspers, Johanneke & Chen, Yija & Heeren, Willemijn F.L. & Pacilly, Jos J.A. & Schiller, Niels O. & van Zanten, Ellen (eds.), *Above and beyond the segments: Experimental linguistics and phonetics*, 83-95. Amsterdam: Benjamins.
- Grice, Martine & Ladd, Robert D. & Arvaniti, Amalia. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17. 143–185.
- Himmelmann, Nikolaus P. 2010. Notes on Waima'a intonation. In Ewing, Michael & Klamer, Marian (eds.), *Typological and areal analyses: Contributions from East Nusantara*, 47–69. Canberra: Pacific Linguistics.
- Himmelmann, Nikolaus P. & Kaufman, Daniel. forthc. Prosodic systems: Austronesia. In Gussenhoven, Carlos & Chen, Aoji (eds.). *The Oxford handbook of prosody*.
- Himmelmann, Nikolaus P. & Sandler, Meytal & Strunk, Jan & Unterladstetter, Volker. forthc. *On the robustness of intonational phrases in spontaneous speech – a cross-linguistic interrater study*. Submitted ms.
- Jun, Sun-Ah. 1998. The Accentual Phrase in the Korean prosodic hierarchy. *Phonology* 15. 189-226.

- Jun, Sun-Ah (ed.). 2005. *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Jun, Sun-Ah (ed.). 2014. *Prosodic typology II: The phonology of intonation and phrasing*. Oxford: Oxford University Press.
- Kamholz, David C. 2014. *Austronesians in Papua: Diversification and change in South Halmahera-West New Guinea*. Berkley: University of California. (PhD dissertation.)
- Kanerva, Jonni M. 1990. *Focus and phrasing in Chichewa phonology*. New York: Garland.
- Khan, Sameer & Khan, Volker. 2014. The intonational phonology of Bangladeshi Standard Bengali. In Jun, Sun-Ah (ed.), *Prosodic Typology II: The Phonology of intonation and phrasing*, 82–117. Oxford: Oxford University Press.
- Ladd, Robert D. 2008. *Intonational phonology*. 2nd edn. Cambridge: Cambridge University Press.
- Lieberman, Philip. 1967. *Intonation, perception, and language*. Cambridge: MIT Press.
- Lindström, Eva & Remijsen, Bert. 2005. Aspects of the prosody of Kuot, a language where intonation ignores stress. *Linguistics* 43. 839–870.
- Maskikit-Essed, Raechel & Gussenhoven, Carlos. 2016. No stress, no pitch accent, no prosodic focus: The case of Ambonese Malay. *Phonology* 33. 353-389.
- Pawley, Andrew and Frances Syder, 2000. "The one-clause-at-a-time hypothesis". In Riegenbach, Heidi (ed.). *Perspectives on fluency*, 163-198. Ann Arbor: U. Michigan Press.
- Remijsen, Bert. 2001. *Word-prosodic systems of Raja Ampat languages*. Leiden University. LOT Dissertation Series, vol. 49. (PhD dissertation.)
- Remijsen, Bert. 2002. Lexically contrastive stress accent and lexical tone in Ma'ya. In Gussenhoven, Carlos & Warner, Natasha (eds.), *Laboratory phonology VII*. Berlin: de Gruyter, 585–614.
- Remijsen, Bert. 2007. Lexical tone in Magey Matbat. In van Heuven Vincent J. & van Zanten, Ellen (eds.), *Prosody in Austronesian languages of Indonesia*, 9–34. Utrecht: LOT.
- Schiering, René & Bickel, Balthasar & Hildebrandt, Kristine A. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46. 657-709.
- Selkirk, Elisabeth O. 1984. *Phonology and Syntax: The relation between sound and structure*. Cambridge: The MIT Press.
- Stoel, Ruben B. 2005. *Focus in Manado Malay*. Leiden: CNWS Publications.
- Stoel, Ruben B. 2006. The intonation of Banyumas Javanese. In Mixdorff, Hansjörg & Hoffmann, Rüdiger (eds.), *Proceedings of the Speech Prosody 2006 conference*, 827–830. Dresden: TUDpress.
- Stoel, Ruben B. 2007. The intonation of Manado Malay. In van Heuven, Vincent J. & van Zanten, Ellen (eds.), *Prosody in Indonesian Languages*, 117-150 Utrecht: LOT.
- Wagner, Michael & Watson, Duane G. 2010. Experimental and theoretical advances in prosody: A review. *Language and cognitive processes* 25. 905-945.
- Utsumi, Atsuko. 2011. Reduplication in the Bantik language. *Asian and African Languages and Linguistics* 6. 5-26.
- van Heuven, Vincent J. & van Zanten, Ellen. 2007. Prosody in Austronesian languages of Indonesia: Concluding remarks. In van Heuven, Vincent J. & van Zanten, Ellen (eds.), *Prosody in Austronesian languages of Indonesia*, 191–207. Utrecht: LOT.
- van Heuven, Vincent J. & van Zanten, Ellen (eds.). 2007. *Prosody in Austronesian languages of Indonesia*. Utrecht: LOT. (downloadable from LOT server)

- van Zanten, Ellen & Stoel, Ruben & Remijsen, Ruben. 2010. Stress types in Austronesian languages. In van der Hulst, Harry & Goedemans, Rob & van Zanten, Ellen (eds.). *A survey of word accentual patterns in the languages of the world*, 87–112. Berlin: de Gruyter.
- van Zanten, Ellen & van Heuven, Vincent J. 1998. Word stress in Indonesian: Its communicative relevance. *Bijdragen tot de Taal-, Land- and Volkenkunde* 154. 129–147.
- Zimmermann, Malte. 2016. Information Structure. *Oxford Bibliographies in Linguistics*. (DOI: 10.1093/obo/9780199772810-0130.)