The study of tone in languages with a quantity contrast

Bert Remijsen
University of Edinburgh

This paper deals with the study of tone in languages that additionally have a phonological contrastive of quantity, such as vowel length or stress. In such complex word-prosodic systems, tone and the quantity contrast(s) can be fully independent of one another, or they may interact. Both of these configurations are illustrated in this paper, and the phonetic pressures underlying the development of interactions are laid out. The paper pays particular attention to the challenge of investigating complex word-prosodic systems. Central to the approach advocated here is the combination of qualitative fieldwork data collection methods with instrumental analysis.

1. INTRODUCTION. In its simplest form, a tonal contrast does not interact with other parts of the phonological system. For example, in a system with Low and High tone categories, we would find the following disyllabic patterns: /màmà, màmá, mámá, mámà/. This simple constellation is the one that is used for pedagogical purposes, when students are auditorily confronted with tone for the first time in ear-training practicals. The lack of interactions which confuse the auditory impressions make such a system an ideal starting point. In practice, however, a linguist studying an un- or understudied tone system stands a good chance of confronting tone in addition to a contrast that affects the relative prominence of syllables. Consider the combination of tone with a vowel length contrast. When two levels of vowel length (short, long) are crossed with the set of tone patterns listed above, there are sixteen different logical combinations. Such a system represents a greater challenge, because both the production and the perception of tone patterns depend on duration. In other words, the same tone pattern may be implemented differently on a syllable with a short vowel as opposed to a syllable with a long vowel; and even if they are acoustically identical, they may still sound markedly different. Similar issues come up when a tone language has lexical stress, or as a function of the segmental composition of the syllable. That is, the same tone category may be acoustically or perceptually different on a stressed vs. an unstressed syllable, or on a syllable ending in a sonorant coda as opposed to one ending in a voiceless coda. Moreover, the inventory of tone patterns may not be independent of these factors.

Vowel length, stress and the segmental composition of the syllable affect tone in similar ways because all three may influence the degree of phonetic prominence of syllables. I will refer to phonological contrasts of this type generically as quantity contrasts. The critical common characteristic of quantity contrasts is duration, although vowel quality and intensity may also play a role. A first goal of this paper is to illustrate both independence and interaction between tone on the one hand and quantity contrasts on the other.[1] A second

---

[1] I leave aside phenomena where duration and F_0 do not reflect independent contrasts, as in e.g. Bur-

goal is to offer practical advice in relation to the challenge of investigating tone in such a configuration.

The paper is structured as follows. I will begin with surveying the ways tone and quantity can be combined in sound systems (Section 2). There is a basic division between configurations where tone and the quantity distinction(s) are independent phonological dimensions of contrast, as opposed to configurations where the two interact to some degree in the phonology. Section 3 considers the problem at the center of the study of languages with tone and quantity: the variation in the realization of a tone pattern as a function of a difference in quantity. Finally, Section 4 provides practical advice. In the context of other didactically oriented texts on the study of tone and prosody, such as Pike (1948), Himmelmann & Ladd (2008), Hyman (2010), and the other papers in this volume, I focus on the challenge posed by systems combining the lexical specification of tone with a quantity contrast. I will argue for the use of both qualitative and quantitative approaches.

2. TYPOLOGY OF INTERACTIONS BETWEEN TONE AND QUANTITY. When a tone language additionally has a quantity contrast, it is fairly common for the distribution of tone patterns to be sensitive to quantity. In this way, more complex tone patterns are often restricted to environments involving greater quantity. This is evident from Zhang's (2002) survey of 187 languages with contour tones, i.e., tone categories involving more than one target within a syllable. Zhang found that only 22 of these have no restriction on the distribution of contour tones. In the other languages in the sample, contour tones are restricted to syllables that offer more sonorous space, typically in terms of a long vowel, a sonorous coda, stress, or final lengthening. Level tone categories may also be restricted to these salient environments (e.g. de Lacy 2002).

The interaction of tone with vowel length is illustrated in (1), on the basis of Stegen's (2002) investigation of Rangi (Bantu, Tanzania).[2] At the syllable level, Rangi has Low, High, Fall and Rise tone categories. But the two contour tone patterns, Fall and Rise, are only found on syllables with a long vowel (Stegen 2002:132). This interaction is represented by postulating (a) that the short vowels count for one weight unit or mora, and long vowels for two such units; and (b) that the contour tones are composed of High and Low tone targets, which associate with moras. It follows then that a syllable with a long vowel can accommodate a contour (two-target) tone category, whereas a syllable with a short vowel cannot (cf. Odden 1995).

(1) mòkòvò 'navel'   ìbâandà 'hut'   ìbǎahìrà 'feather'   ìkǔungòlò 'pied crow'
     ìbátà 'duck'   bàankà 'room'   mòtîìkɔ̀ 'ladle'   mòkóʊlò 'elder brother'

mese (Wheatley 2013) and Cologne German (Gussenhoven & Peters 2006).

[2] Note on transcription. In these and following examples, long vowels are represented by doubling the vowel character, and overlong vowels by tripling it. Other suprasegmental features are transcribed above or below the first vowel character only. For example, the vowel of the Dinka word *rɔ̀oor* 'men' is breathy throughout, and the low tone mark reflects the tone pattern of the syllable as a whole, even though these features are only represented above and below the first 'o'.

Similar interactions are found between tone and stress. This is illustrated in (2) by data from Jemez (Kiowa-Tanoan, USA), one of the languages cited by Zhang. The original source is Bell (1993), from which the examples are drawn. The syllable-level distinctive tone patterns in Jemez are Low, Mid, High, and Fall. The first syllable of the word is stressed, and vowel length is contrastive in this position only. This stressed initial syllable has greater duration, even when the vowel is phonologically short. The distribution of the Fall – the only contour tone in the inventory – is limited to this stressed initial syllable.

(2) cê 'stick'   côtē 'antlers'   gídá 'pot'   hȍomȕtǽ 'shovel'
     cée 'porcupine'   cáhā 'lips'   gínȃmȕ 'dove'   wáagíshà 'cow'

In both of these languages, the distribution of contour tones is restricted to contexts that offer greater quantity: long vowels in the case of Rangi, and stressed syllables in the case of Jemez. There are several other factors that play a role in distribution of tone patterns, such as the presence of a sonorant coda, and the right edge of prosodic domains (Zhang 2002). Zhang points out that what all these contexts have in common is that they offer more sonorous space on which the tone patterns can be realized. This conclusion indicates that there is a phonetic motivation for such interactions. In Section 3 I will discuss the limitations in speech production and speech perception that give rise to such interactions.

   However, it is important to realize that the influence of sonorous space on the distribution of tone patterns is not deterministic: both vowel length and stress may represent independent dimensions of contrast in tone languages. A remarkable instance of the combination of vowel length and tone is found in Dinka (Nilo-Saharan, South Sudan). This language has three levels of vowel length in stem syllables, which are predominantly monosyllabic, and two levels of vowel length in prefix syllables (Andersen 1987, Remijsen & Gilley 2008). These contrasts are illustrated by the verb forms in (3), on the basis of data from the Twic-Malual dialect. The columns illustrate the three levels of vowel length in the stem syllable. Length in this position is distinctive both lexically – e.g. *-lèel* 'isolate\3s' vs. *-lèeel* 'provoke\3s' – and morphologically – e.g. *-lèl* 'isolate\2s' vs. *-lèel* 'isolate\3s'. Comparison between upper and bottom rows illustrates vowel length in the prefix. Here vowel length marks agreement for number with the preceding noun: the declarative prefix (DC) is short when the preceding noun is singular and long when the preceding noun in plural.

(3) ràaan   a-lèl   ràaan   a-lèel   ràaan   a-lèeel
     person DC.S-isolate\2s   person DC.S-isolate\3s   person DC.S-provoke\3s
     'You are isolating somebody.'   'He is isolating somebody.'   'He is provoking
                                                                    somebody.'

     rọ̀oor aa-lèl   rọ̀oor aa-lèel   rọ̀oor aa-lèeel
     men   DC.P-isolate\2s   men   DC.P-isolate\3s   men   DC.P-provoke\3s
     'You are isolating men.'   'He is isolating men.'   'He is provoking men.'

The inventory of tone categories differs between dialects of Dinka (cf. Andersen 1987, Remijsen & Ladd 2008, Remijsen 2013). The Twic-Malual dialect has the categories High, Low, and Fall on stem syllables.[3] If Dinka were like Rangi, the distribution of the Fall would be limited to syllables with long (VV) and overlong (VVV) vowels. To the contrary, the Fall is found on syllables with short, long and overlong vowels alike. This is illustrated in (4), which extends the minimal set in (3): each of the vowel length patterns illustrated with a Low tone on the stem syllable in (3) appears with a Fall tone on the stem syllable in (4). These examples show that the contrast between Low and Fall is orthogonally crossed with the vowel length contrasts in stem and prefix syllables.

(4)

ràaan a-lêl	ràaan a-lêel	ràaan a-lêeel
person DC.S-isolate\PAS	person DC.S-provoke\PAS	person DC.S-provoke\3S.FUG
'Somebody is being isolated.'	'Somebody is being provoked.'	'He provokes somebody to leave.'
ròor aa-lêl	ròor aa-lêel	ròor aa-lêeel
men DC.P-isolate\PAS	men DC.P-provoke\PAS	men DC.P-provoke\3S.FUG
'Men are being isolated.'	'Men are being provoked.'	'He provokes men to leave.'

As in Rangi, the vowel length contrast can be represented using moras. In Dinka, the ternary nature of the contrast in stem syllables motivates distinguishing between mono-, bi-, and trimoraic syllables. The difference between the two languages is that in Rangi the syllable-internal weight structure determines the distribution of tone patterns, but not in Twic-Malual Dinka. In terms of the phonological analysis, the tone-bearing unit is the mora in Rangi, but the syllable in Dinka.

Similarly, stress can also be contrastive in a tone language. This configuration is found among others in Ma'ya (Austronesian, Indonesia), spoken off the west coast of New Guinea. The forms in (5) illustrate the dialect spoken on the island Salawati (Remijsen 2002). As seen from (5a), stress is either on the penultimate or on the final syllable. Longer words with penultimate stress include *mu ˈtumol* 'people', and *ma ˈlyabán* 'lightning'. In addition, tone is distinctive: there are High and Rise tonemes which contrast with the absence of tonal specification, as in (5b). The lexical specification of tone is limited to the final syllable. There is one restriction in the distribution: when the final syllable has the Rise, stress is invariably on the same syllable, as in *ma ˈlŏm* 'wet' and *ka ˈlĕp* 'hole'.

---

[3] Tone contrasts in prefix syllables exist, but it is limited to High vs. non-High. I interpret non-High prefix as not specified for tone rather than Low, because a Low tone specification would trigger a different different (i.e., low level) allophonic variant of the Low toneme on the stem syllable, rather than the early-falling variant observed in the forms in (3,4).

(5)   a.

ˈtalá	taˈlá	b.	ˈsá	ˈsǎ	ˈsa
'banana'	'k.o. liana'		'climb'	'sweep'	'one'
ˈmaná	maˈná		ˈná	ˈnǎ	ˈna
'grease'	'light (weight)'		'k.o. palm'	'sky'	'belly\3'
ˈkayá	kaˈyá				
'rich'	'machete'				

Listening to the minimal sets for stress in (5a), the most salient difference between the members of each set is vowel duration: the penultimate vowel sounds longer in the penult-stressed forms, and the final vowel in the final-stressed forms. Just on the basis of pairs like these, and limited to auditory evidence, we could speculate that Salawati Ma'ya combines tone with vowel length, just like Dinka. But there is an important difference: in Dinka, vowel duration is contrastive on a syllable-by-syllable basis. For example, when in Dinka a long vowel conditions greater duration in the prefix, then the stem syllable may still have either a short, long or overlong vowel. In contrast, vowel duration in Ma'ya is part of an obligatory and culminative contrast, i.e., stress (cf. Hyman 2006). By implication of culminativity, relatively greater vowel duration in one syllable, due to stress, implies relatively shorter vowel duration on any other syllables in the word, due to lack of stress. The rejection of a vowel length analysis is strengthened by the evidence from monosyllabic words: these do not present noticeably different levels of vowel duration. Moreover, the phonetic realization of stress in Ma'ya also includes vowel quality and the distribution of intensity over the spectrum (Remijsen 2002). The same package of correlates (duration, vowel quality, relative intensity) realizes stress in various other languages (e.g. Sluijter & van Heuven 1996).

    Given that tone can co-occur with either vowel length or stress, what about a configuration in which tone, vowel length and stress all three constitute independent dimensions of contrast? A compelling instance of this configuration is presented in Michael (2011), in relation to Iquito (Zaparoan, Peru). His examples illustrating the various contrasts are cited in (6). The words in (6a) illustrate that vowel length is contrastive in a paradigmatic way: in a disyllabic word, none, either one, or both of the syllables may have a long vowel. The examples in (6b) show that primary stress is on the penultimate when both the penultimate and the final syllable have a short vowel, and that the addition of an affix shifts primary stress to the right. In phonological terms, this can be represented by postulating left-headed binary feet, assigned from the right edge of the word. Vowel length matters for stress: if the final syllable has a long vowel, it carries stress, as in ˌmaˈhuu in (6a). On this evidence, Michael infers that the metrical feet are composed not of syllables but of moras, with a long vowel corresponding to two moras. In this analysis, the final syllable in *maˈhuu* corresponds to a whole metrical foot. It is clear that the prominence of the syllables transcribed as stressed reflects a metrical structure. Finally, the examples in (6c) show that High tone distinguishes unrelated words, and that the specification of High tone is not limited to the

stressed syllable. Like stress, tone is sensitive to weight structure below the syllable level. Whereas only one pattern of tonal specification is found on syllables with a short vowel, there are two possible for syllables with a long vowel: with the peak reached earlier, as in ˈtúuku, or with the peak reached later, as in ˈtŭuku. In the phonological analysis, this configuration is represented by postulating that the mora is the tone-bearing unit, rather than the syllable. In this way, the pattern with a later-aligned high target is represented as High tone associated with the second mora.

(6)[4]   a.   ˈisɨ

‘k. o. lizard’

ˈsaapi

‘stingray’

ˌmaˈhuu

‘charichuelo (k. o. tree)’

ˌiiˈpɨɨ

‘red howler monkey’

b.   ˌmiˈnati

‘pineapple’

ˌminaˈtika

‘pineapples’

c.   ˌmáˈʃiku

‘raft’

ˌmaˈʃíku

‘paucar (k.o. bird)’

ˈtúuku

‘tumpline’

ˈtŭuku

‘ear’

Both Ma'ya and Iquito show that High tone and stress do not necessarily coincide on the same syllable. As such, they represent violations of the constraints postulated to favor their cooccurrence by de Lacy (2002). The more complex Iquito system is of particular interest for another reason, namely the phonetic realization of the stress contrast. Many acoustic studies report vowel duration as the main correlate of stress, with vowel quality or intensity-related parameters as secondary correlates (Beckman & Edwards 1994 on English; Sluijter & van Heuven 1996 on Dutch; de Jong & Zawaydeh 1997 on Arabic; Remijsen & van Heuven 2005 on Papiamentu; Ortega-Llebaria & Prieto 2011 on Spanish). Nonetheless, some languages display a divergent package of phonetic correlates, and Iquito is one of them. Michael reports lengthening of the consonant that follows the vowel of the stressed syllable as the most important phonetic parameter in the realization of stress. Greater duration of the same constituent has also been reported as a stress correlate for Washo (Yu 2008), Welsh (Williams 1985, 1986), and Zapotec (Pickett, Villalobos & Marlett 2010). All of these languages have contrastive vowel length. This suggests that the realization of stress may vary in such a way as to optimally use the available phonetic parameters (Berinstein 1979). Intriguingly, Everett (1998) reports that in Pirahã, the duration of the syllable onset consonant plays a prominent role in the phonetic realization of stress. In the light of these divergent findings on the realization of stress, we can conclude that its realization is not fixed across languages. At this point, it is not clear whether the set of correlates is predictable on the basis of the phonological system. This means that, in the study of un- or under-investigated prosodic systems, we need to be open-minded regarding the possible phonetic correlates of any culminative prominence we may encounter.

---

[4] I thank Lev Michael for providing the Iquito sound examples linked to the transcriptions in (6).

3. LINGUISTS MAY MISINTERPRET; NATIVE SPEAKERS MAY HYPOCORRECT. The transcriptions of the word-prosodic systems illustrated above separate out the constituent contrasts of tone, length and stress in an unambiguous manner. This is of course an abstraction from the data: what we perceive are strings of syllables, each with a set of phonetic characteristics. A salient pitch pattern on a particular syllable, for example, could be due to lexical specification for tone, intonational specification due to an edge tone or a pitch-accent, or it could be that no phonological specification or tone is involved at all, but that the syllable is just more prominent due to another factor, such as vowel length or stress. This last scenario is illustrated by the data from Ma'ya in (5). In forms with penultimate stress like ˈ*maná* 'light', we can perceive falling pitch on the stressed penultimate syllable, ahead of the high pitch on the final syllable. In the first descriptive analysis of Ma'ya word prosody, van der Leeden (1993) postulated a falling tone category on penultimate syllables of words with this pattern, so *mâˈná* 'light'. As seen from his transcription, he also postulated stress along with the High tone on the final syllable. An extensive evaluation of the competing analyses is presented in Remijsen (2002). The controversy stems from the fact that the correlates of stress (duration, vowel quality, relative intensity) render the F_0 pattern of stressed penultimate syllables more salient – even though F_0 is not distinctive in this position. Moreover, if intonational tone targets are a reliable indication of the location of stress in our first language, we will be biased to similarly infer culminative prominence in the study of an unfamiliar language. I speculate that, in this way, van der Leeden's native competence of Dutch led him to perceive stress prominence on the High-toned final syllable.

This example shows that, when a language has not just tone but also vowel length or stress, the number of possible analyses increases substantially. In the study of such systems, disagreement regarding the phonological constellation is more common than the scenario whereby independent descriptive analyses are in unison. For Iquito, see the discussion of Eastman & Eastman (1963) in Michael (2011); for Dinka, see the discussion of Gilley (2003) in Remijsen & Gilley (2008). Controversies like these indicate that our auditory impressions do not consistently offer us a solid foundation on which to develop a descriptively adequate analysis. I will come back to this in Section 4.
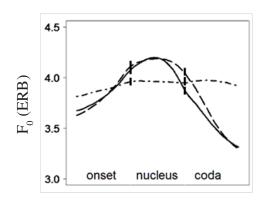
Leaving aside the challenges to the linguist, the phonetic influence of a quantity contrast on tonal realization is a driver of diachronic change. In particular, a system that has independent tone and quantity contrasts may develop phonological interactions that limit the distribution of tone categories to syllable structures of greater quantity. An examination of an instance of this development in Dinka will reveal the phonetic influence of quantity on tone. The variants at issue in this phenomenon are the Bor dialects. The main difference with the Twic-Malual dialect illustrated in Section 2 is that it has an additional tone category, i.e., the Mid toneme.

When tone and vowel length are phonologically independent, then a particular tone category can sound different depending on the level of vowel length. For example, consider the Bor Dinka forms *ā-lêl* 'DC.S-isolate\PAS' and *ā-lêel* 'DC.S-provoke\PAS'. These forms have the same phonological specification for tone on the stem syllable: the Fall, which involves a peak at about 30 to 40 milliseconds into the vowel of the stem syllable, followed by a drop in F_0 over the remainder of the syllable (Remijsen 2013). The

alignment and excursion size of the falling contour on the stem syllables of *ā-lêel* and *ā-lêl* are very similar acoustically. This can be seen from Figure 1A, which shows the averaged F_0 traces of realizations of these two words, from 13 speakers of the Bor dialects. The full trace shows the F_0 pattern of *ā-lêel*, and the interrupted trace reflects *ā-lêl*. The similarity between the traces is partly due to fact that that the traces are time-normalised: that is, the traces are shown on the same horizontal space in the graph, abstracting away from the difference in duration. In fact, the average durations for short-vowel nuclei in Dinka is about 70 ms, whereas the average duration of long-vowel nuclei is over 50% longer, at about 110 ms. This difference in duration affects the way the tone pattern is realized and perceived. Note how, by the end of the nucleus, the Fall has conditioned a much greater drop in F_0 when the vowel is long (full line) than when it is short (interrupted line). In the latter case, i.e., when the vowel is short, most of the F_0 fall is realized on the sonorant coda.
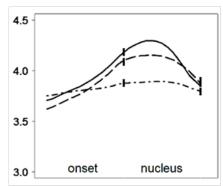
1.A Sonorant coda   1.B Voiceless coda

FIGURE 1: F_0 traces on normalized time axes, averaged across realizations of tone patterns in Dinka dialects: CV̂VC (full line, based on 13 speakers of Bor North and Bor South), CV̂C (interrupted line, based on 7 speakers of Bor North), and CV̄C (dash-dotted line, based on 6 speakers of Bor South). Separate panels by coda type (sonorant vs. voiceless plosive). The edges of the nucleus (vowel) are shown by short vertical lines. Based on data in Remijsen (2013).

The situation is more acute when the coda is a voiceless plosive, and therefore does not carry F_0. Averaged F0 traces for these patterns are shown in Figure 1.B. When the vowel is long, as in *ā-ŋôot* 'DC.S-spit\PAS' (full trace), F0 has changed substantially from the peak (at about one third into the vowel) to the end of the vowel. But when the vowel is short, as in *ā-ŋôp* 'DC.S-take.a.gulp\PAS' (interrupted trace), there has been insufficient time for the speakers to realize a substantial change in F_0. This is due to an important characteristic of the way F_0 changes are implemented, examined at length in Xu & Sun (2002). In an experimental study on the maximum speed of F_0 change, Xu & Sun found that the realization of F_0 patterns is nothing like a straight-line interpolation between targets. Instead, F_0 changes more like a bell curve, with the rate of change being much slower near the turning point as opposed to when a change is well underway. Taking into account the slow-moving

part near the turning point, Xu and Sun found that, on average, it takes 124 milliseconds to realize a fall in F_0 of around four semitones (e.g. 150 to 120 Hz). The effect of this limitation in speech production can be observed in Figure 1.B, by examining the realization of the Fall on a short vowel (interrupted trace). As noted above, we need to keep in mind that the horizontal dimension of the graph does not allow for a comparison of the duration of the short and long vowels. Given that the Fall reaches its peak at about 35 milliseconds into the vowel, then by the end of the vowel, i.e., on average about 35 milliseconds further along, F_0 has not gotten fully out of the slow-changing section near the turning point. The same is the case when the vowel is followed by a sonorant coda (Figure 1.A), but there the coda carries the fast-changing part of the curve.

Apart from this limitation on the side of speech production, another important factor here is our ability to perceive F_0 changes as pitch contours. This ability depends not only on the size of the F_0 change but also on the duration of the signal on which the change is realized. This limitation is known as the glissando threshold (Rossi 1971, Greenberg & Zee 1979, 't Hart, Collier & Cohen 1990). An F_0 change from 150 to 130 Hz can be perceived as a pitch contour when it is realized over 100 milliseconds. But when the same F_0 change is realized over 50 milliseconds, it is prone to be perceived as a level pitch pattern. As a result of this glissando threshold, then, for *ā-ŋôp* to be *perceived* with the same fall in pitch as *ā-ŋôot*, speakers would need to *produce* a greater fall in F_0, to counter the fact that the glissando threshold is higher when the signal is shorter. And, of course, it is difficult to do so, given the 'speed limit' observed by Xu & Sun (2002). In this way, the limitation on the speed of F_0 change conspires with the glissando threshold to make audible contour tone patterns more difficult to realize when sonorous space is limited. This is in line with the typological observations of Zhang (2002).

Measurements on the realization of *ā-ŋôp* in the Bor North dialect (interrupted line in Figure 1.B) suggest that this F_0 fall is below the glissando threshold, i.e., prone to be perceived as a level pitch pattern (Remijsen 2013). This interpretation is borne out by the evidence on the cognates of forms with a short stem vowel (such as *ā-lêl* and *ā-ŋôp*) in the Bor South dialect. In this dialect, the Fall contour does not occur on syllables with a short vowel. In inflections like the passive, where Bor North has the Fall both on short stems and on long stems, Bor South has the Fall only if the vowel is long. If the vowel is short, the passive is realized with the Mid toneme, i.e., *ā-lēl* and *ā-ŋōp*. This can be observed in the dash-dotted traces in both panels. This development can be interpreted as hypocorrection (Ohala 1989). The starting point was arguably the situation that can still be observed in Bor North, where the Fall category is realized both through falling and level tone percepts, depending on vowel length. The diachronic reinterpretation of the Fall toneme on a short vowel as a Mid toneme can be attributed to this 'pool of synchronic variation', in the sense of Ohala (1989). The outcome of this diachronic change is that the Bor South dialect has developed an interaction of the kind illustrated by Rangi in (1), whereby the tone inventory is restricted on short vowels.

Overall, the Dinka language is characterized by substantial between-dialect variation with respect to tone, whereas the quantity system is consistent across dialects. This state of affairs is recurrent among the world's languages: tonal phenomena appear to be less stable diachronically compared to other components of the sound system, including quantity distinctions. That is, we often find considerable variation in terms of tone and intonation among dialects or closely-related languages – e.g. Haraguchi (1977) on Japanese, and Riad

(1998) on dialects of Swedish and Norwegian. In contrast, vowel length and stress are more stable, to the extent that a particular stress pattern may predominate within an extensive genetic linkage (e.g. Goedemans 2010). This difference between tone and quantity may provide valuable external evidence in relation to complex suprasegmental phenomena: if the status of a suprasegmental contrast is challenging in a particular variety of a language, the same forms in another variety or in a closely-related language may offer valuable insights.
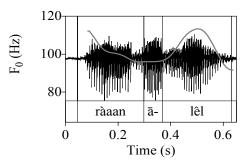
In summary, a quantity contrast tends to affect the production and perception of a tone contrast on the same constituent. We confront the data in this fused manner, so that a correct interpretation of the tone system hinges on an accurate understanding of the prosodic system as a whole. If tone and quantity contrasts are independent at the phonological level, we can expect substantial interactions in terms of the phonetic realization of the tone categories. These phonetic interactions may become phonologised diachronically, resulting in restrictions on the distribution of tone patterns to environments that offer more sonorous material (cf. Zhang 2002).

4. ADVICE ON HOW TO STUDY COMPLEX WORD-PROSODIC SYSTEMS

4.1 THE RATIONALE FOR COMBINING QUALITATIVE AND INSTRUMENTAL APPROACHES. What makes a description of a tone system adequate in a descriptive sense? The first characteristic that comes to mind is that the contrasts that are postulated have been examined in relation to relevant contextual factors and potential confounds. Snider (this volume) and Yu's (this volume) discussion of Pike (1948) present detailed advice. As for the nature of the data that are used, traditionally researchers investigating speech on the basis of primary data had expertise either in qualitative or in quantitative methods, but not both. The former were prepared to discover uninvestigated phenomena based on auditory impressions, and postulate phonological representations. The latter designed experiments to evaluate competing analyses. In recent decades this division has become less pronounced, as more and more researchers combine qualitative and quantitative methodologies (e.g. Coupe, this volume). When it comes to the study of complex word-prosodic phenomena, such as systems combining tone and quantity contrasts, this development is very welcome, because both qualitative and quantitative approaches have an important role to play in order to produce a comprehensive descriptive analysis that is maximally reliable and compelling.

At the center of the qualitative approach is ear-based transcription, applied to data collected through controlled elicitation, spontaneous speech (e.g. narratives), or participant-observation. This approach holds the potential for discovery, with respect to prosody as much as in relation to any other part of the language system. Such exploratory investigations yield all kinds of insightful data, evidence of tonal contrast, tone sandhi processes, morphological alternations, etc. In turn, quantitative methods are equally of great importance. First, acoustic examination can complement the limitations of our auditory sensitivity. Linguists differ in auditory sensitivity, and our native language affects our auditory abilities (cf. Flege 1995). For example, if our first language has word-level stress and intonational pitch-accent, as in English and closely-related languages, we are primed to weigh all of the phonetic parameters (duration, F_0, intensity, etc.) to determine the location of a strong syllable within a given constituent – irrespective of the role these parameters play in

the language under investigation. In my own auditory impression of the Bor Dinka word *āa-lêl*, the greater duration of the long vowel in the first syllable is overshadowed by the tone pattern of the second syllable, which involves high F_0. This perception is not surprising, given that intonational pitch-accents reliably cue stressed syllables in my first language (Dutch). This bias means that I cannot count on it that my auditory impressions will reveal the crucial phonetic parameters. Importantly, the long duration of the prefix is evident from the acoustic record. This can be seen from Figure 2, which shows the segmented waveform for *āa-lêl* in the righthand panel, and a comparable utterance with *ā-lêl* in the same position in the lefthand panel. The overlaid F_0 traces show the high F_0 on the final syllable, which obscures the difference in duration in my auditory impressions of these utterances.
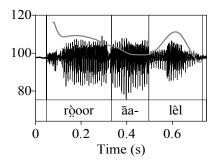
FIGURE 2: Waveforms, segmentation, and F_0 traces for *ràaan ā-lêl* 'A person is being isolated.' (left) and *rò̰oor āa-lêl* 'The men are being isolated.' (right), uttered by a speaker of the Bor North dialect of Dinka. Glosses of the target words are 'DC.S-isolate\ PAS' and 'DC.P-isolate\PAS', respectively.

In Section 3, I speculated that the same overriding influence of high F_0 led van der Leeden to transcribe final stress in Ma'ya words like ˈ*maná* 'light'. That is, the high-F_0 correlate of the High tone would have outweighed the non-F_0 stress correlates on the penultimate. This 'culminative' bias is a liability when the phonetic parameters reflect two separate phonological distinctions. Given the role of our native language experience in molding our auditory potential, biases like this one are inevitable when we confront an unfamiliar prosodic system. Examination of acoustic representation can reveal to us what we did not hear, in terms of F_0, duration, vowel quality and voice quality.

A second motivation to use acoustic analysis is that the risk of descriptive error is already factored in among our peers. That is, when a descriptive study postulates a structure that is typologically or theoretically challenging, the typology or theory is usually not straightforwardly revised. Consider, for example, the notion of three-level vowel length, discussed above in relation Dinka. This contrast was first reported for Dinka in Andersen (1987). In spite of this study and reports on three-level vowel length in a handful of other languages, many linguists continued to assume that vowel length contrasts are more restricted, with a maximum of two levels. For example, in a monograph on syllable structure, Duanmu (2009) makes no mention of three-level vowel length. The reason for such a cautious disposition becomes clear when we consider that, for Dinka as for several other languages where it was postulated, three-level vowel length was called into question in

subsequent studies based on primary data.[5] In a survey of hypothesized cases of three-level vowel length, Odden (2011) adopts alternative interpretations for most cases. In summary, the discipline is conservative when it comes to hypotheses that challenge theory and typology, and typically holds out for independent evaluation. This stance is justified by the observed level of between-researcher disagreement. In this context, quantitative analysis has an important role to play: it makes it possible to test hypotheses through inferential statistics based on instrumental measurements. Crucially, this evaluation is independent of the sensory faculties of the researcher. Below I provide some specific suggestions on how to combine qualitative and quantitative approaches in descriptive analysis of complex word-prosodic systems.

4.2 COMBINING QUALITATIVE AND INSTRUMENTAL APPROACHES IN THE EXPLORATORY PHASE. Even when an investigation is narrowly focused on the phonology or phonetics of tone and other word-prosodic contrasts in a given language, it is worthwhile to develop a thorough first-hand understanding of the lexical and morphosyntactic functions to which these contrasts are relevant. For this reason, it is good to start with extensive explorations, during which we develop a familiarity with the form and function of the word-prosodic contrasts within the grammar as a whole. We arrive at this understanding using traditional fieldwork methods such as controlled elicitation, corpus analysis and participant-observation (e.g. Newman & Ratliff 2001). At this early stage, only one or two native-speaker consultants need to be involved.

Even during the exploratory stage, it can be useful to complement ear-based transcription with visual inspection of recorded forms, using representations such as waveforms, F_0 traces and spectrograms. Recently I explored the sound system and grammar of Kwasio, a Bantu language of Cameroon, in the context of a field methods course. We started out eliciting nouns in citation form and in two contexts. Based on my auditory impressions, I distinguished four tone patterns: low, mid, high, and falling. Some of the words for which I transcribed a mid tone include *dzjɛh* 'forest' , *sa* 'feather' , and *mbvo* 'roof' . When we started recording some of the material after a few sessions, it was clear that the syllables I had transcribed with mid tone consistently started out with mid-level F_0, and then continued to rise in the latter part of the syllable. I had missed this rise completely in my initial transcriptions. From that point onwards, I developed awareness for this tone pattern. In general, knowing the typical excursion size or approximate alignment of a tone pattern can help to focus our attention on what to listen out for, or to come up with additional contexts that are worth exploring. In the case of the rising tone category of Kwasio, for example, the instrumental evidence motivates the elicitation of utterances in which this tone pattern is followed by a Low-toned syllable, i.e., a context where the high end target of the rising pattern is most salient.

Minimal sets often represent valuable bootstraps into hard-to-hear contrasts: they enable us to tune in to a distinction, even if we don't have a clear sense of its phonological

---

[5] Gilley (2003) presented an alternative analysis for Dinka, which does not make reference to three-level vowel length. Later on, Remijsen & Gilley (2008) corroborated Andersen's (1987) analysis. As for three-level vowel length contrasts postulated for other languages, competing analysis have been postulated among others for Mixe (Hoogshagen 1959, Wichmann 1995), Gaelic (Ternes 1989, 2006), and Estonian (Prince 1980; Hayes 1989; Lippus, Pajusalu & Allik 2009).

structure and phonetic realization. In Kwasio, even before the tone pattern of words like *sǎ* 'feather' was determined, it was clear that it was distinct from the High tone pattern in words *sá* 'k.o. fruit' . Going back to the example of the long vowel in the prefix of Dinka *āa-lêl* , we can develop awareness for it by listening to it alongside *ā-lêl* , where the prefix is short. Visually, the difference is obvious, as seen from the waveforms and segmentations in Figure 2. Intent listening in this way is most effective when the forms are recorded, so as not to wear out the consultant's patience. I recommend to not just listen to the forms but also to mimick them. This practice brings to mind exercises for second-language learners, instructed to listen to and repeat a target pronunciation. In a recent experimental study, Adank, Hagoort & Bekkering (2010) have demonstrated that mimicking is an effective tool to develop awareness of unfamiliar sound patterns. Obviously, it only becomes possible to focus on the difference in vowel duration between *ā-lêl* and *āa-lêl* after one has discovered that there is a meaningful difference here – in this case agreement marking. I became acutely aware of vowel length in this position when a native-speaker consultant corrected me as I mispronounced it, uttering a a short vowel where agreement required a long one.

In summary, rather than using acoustic analysis only as a stage from data collection to statistics, in a one-way-traffic sense, I advocate adding it to mix of methodologies used during the exploratory phase, in a dialectic process with traditional field methods. At the same time, I caution against spending too much time and effort on instrumental analysis early on. At this exploratory stage in the investigation, we need to get a grip on the system of distinctions and processes, and be open to alternative interpretations. Additional lexical items, word forms, or utterance contexts can greatly enrich or alter the budding analysis. It is a stage of fertile confusion, when our working hypotheses are in flux. We should not cut this phase short and move on to systematic recording too early. This temptation presents itself in particular to researchers who are in their comfort zone using acoustic analysis and statistics. As a rule of thumb, as long as the exploratory sessions lead us to alter our working hypotheses, it is best to keep going with them: who knows what the next session could yield?

Consider for example the situation whereby, in an early session, we have recorded a list of nouns in a frame. Following the session, the key task is to process and reflect on the session notes, and to listen to any recorded forms – often this can be part of storing them as individual sound files. In the process, saliently different suprasegmental patterns may reveal themselves. For some patterns the realization and phonological nature may be instantly obvious; others may be enigmatic. Rather than to attempt to determine the acoustic realization exhaustively at this point, I would recommend to simply group forms that have the same suprasegmental pattern and to spell out the salient phonetic characteristic(s) of such patterns. We are then poised to further our understanding in the next session, by getting additional cases, or by examining the patterns in a different context.

There are no reliable templates as to where in the lexicon or the morphosyntax the most insightful data are to be found during the explorations into tone. It may be verb morphology (Andersen 1993) or noun morphology (Hyman 2010), content words in context, or content words in citation form (Kutsch-Lojenga 1996). We need to explore a wide range of contexts, so that we can consider many alternative hypotheses. Analysis of spontaneous speech is particularly insightful to ensure that the description covers the system. For

example, Western Nilotic languages such as Dinka and Shilluk present sequential clause types, which are marked through vowel length and tone of the verb. If data collection is limited to controlled elicitation of single-clause utterances, this area of the suprasegmental morphophonology will be overlooked completely. In this way, the transcription of spontaneous speech material such as narratives can be crucial to reveal unexpected dimensions of the sound system and the grammar.

How much help the language consultant can offer in figuring out the suprasegmentals varies greatly. Speakers of tone languages are typically not aware that they use tone contrastively, just as most speakers of English will be unable to tell how many vowel categories their native language has. Still, speakers of some tone languages effortlessly develop awareness of their tone categories. Investigating Matbat, a tone language of New Guinea, I found that my consultant, Absalom Jemput, had no difficulty in adding dozens of instances of each of the six lexical tone categories. In contrast, the Dinka language consultants I have worked with did not have this awareness of tone, presumably because tone in Dinka is determined to a large extent by the morphology rather than in the lexicon.

When the hypotheses are becoming solid, it is worthwhile to put them to the test informally with the reference speaker(s). This can be done by resynthesizing one member of a minimal set, modifying the parameter that is hypothesized to be primary correlate, so as to yield category that is most similar it in phonetic terms. If the reference speaker's interpretation of the resynthesized form confirms our prediction, then we know we are on solid ground.

4.3 SYSTEMATIC QUANTITATIVE ANALYSIS. At the end of the exploratory phase, we have firm working hypotheses on the suprasegmental distinctions that are at issue, on the lexical and morphosyntactic forms in which they can be found, and on contextual processes that play a role. Now we are in a position to document the system and examine the phonetic realization in a systematic manner, by recording illustrations of the various suprasegmental patterns from a sample of speakers. Ideally, this sample is big enough to make statistical inferences regarding the language population as a whole. Collecting data from eight speakers is usually safe. Importantly, their dialect background should match that of the reference speaker(s). This is particularly important in relation to tone, where divergence between language varieties is likely (see Section 3 above). By collecting data from various speakers, our data offers a window on variability within the speech community. In contrast with the exploratory phase, which can extend over many months with daily sessions, systematic data collection does not take much time. After piloting the elicitation with one speaker, it is often possible to collect data from eight speakers in a matter of days.

The suprasegmental patterns can best be collected in minimal sets, or otherwise in words that are comparable in their segmental composition. Different classes of segments present their own advantages and potential problems in relation to acoustic analysis. For example, plosives are easy to segment (Turk, Sugahara & Nakai 2006). So are nasal consonants and /l/, and these are also optimal for F_0 analysis. However, F_0 often shows a spike at the boundary between these segments and adjacent vowels, and such spikes should be trimmed (Xu 1999:61-63). These are just a few of the relevant considerations: a full overview of the design of production studies is beyond the scope of this paper.

It is worthwhile to include in the design not only the prosodic contrast(s) under investigation, but also any contexts that are likely to influence the prosodic parameters that act

as correlates for the phonological contrast(s). For example, when vowel duration is affected by the distinction at issue, it may be insightful to collect the target words both in utterance-medial and in utterance-final position, so as to examine how the distinction interacts with utterance-final lengthening (Nakai, Kunnari, Turk, Suomi & Ylitalo 2008). When stress is involved, the interaction with utterance-level prominence is a key consideration (Huss 1978; van Zanten, Goedemans & Pacilly 2003, Himmelmann & Ladd 2008). In general, whatever the nature of the word-prosodic contrasts, intonation is likely to play some role as well (e.g. Liu & Xu 2005; Karlsson, House & Svantesson 2012). There is no template for the range of contexts which a comprehensive acoustic analysis of an unfamiliar should cover. Instead, the exploratory phase is key to ensuring that the quantitative production study is designed so that the results are as insightful as possible, leading to increased understanding of complex word-prosodic systems.

If the working hypotheses are correct, then the measurements will yield consistency across speakers in the descriptive statistics. If they do not, we have missed out something. Often the results hint at a dimension we have not controlled, i.e., we did not fully understand one of the phonological, morphological or intonational factors that are involved. It is a situation I have found myself in several times. This new confusion is the gateway to subsequent collection of exploratory or systematic data, which will ultimately yield a compelling analysis.

REFERENCES

Adank, Patti, Peter Hagoort & Harold Bekkering. 2010. Imitation improves language comprehension. *Psychological Science* 21. 1903-1909.

Andersen, Torben. 1987. The Phonemic System of Agar Dinka. *Journal of African Languages and Linguistics* 9. 1-27.

Andersen, Torben. 1993.Vowel quality alternation in Dinka verb inflection. *Phonology* 10. 1-42.

Beckman, Mary E. & Jan Edwards. 1994. Articulatory evidence for differentiating stress categories. In Patricia A. Keating (ed.), *Phonological Structure and Phonetic Form: Papers in Laboratory Phonology III*, 7-33. Cambridge: Cambridge University Press.

Bell, Alan. 1993. Jemez tones and stress. *Colorado Research in Linguistics* 12. 26-34.

Berinstein, Ava E. 1979. *A Cross-Linguistic Study on the Perception and Production of Stress*. (UCLA Working Papers in Phonetics 47). Los Angeles: University of California.

Coupe, Alexander R. Strategies for analyzing tone languages. *Language Documentation and Conservation.* 8. 490-524

de Jong, Kenneth & Bushra A. Zawaydeh. 1999. Stress, duration and intonation in Arabic word-level prosody. *Journal of Phonetics* 27. 3-22.

de Lacy, Paul. 2002. The interaction of tone and stress in Optimality Theory. *Phonology* 19. 1-32.

Everett, Keren M. 1998. The acoustic correlates of stress in Pirahã. *Journal of Amazonian linguistics* 1(2). 104-162.

Duanmu, San. 2009. *Syllable Structure: The Limits of Variation*. Oxford: Oxford University Press.

Flege, James E. 1995. Second-language speech learning: Theory, findings, and problems. In Winifred Strange (ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*, 229-273. Timonium, MD: York Press.

Gilley, Leoma. 2003. The feature of stress in Nilotic. In Rose-Juliet Anyanwu (ed.), *Stress and Tone – the African Experience. Frankfurter Afrikanistische Blätter* 15, 99-120.

Goedemans, R.W.N. 2010. An overview of word stress in Australian Aboriginal languages. In Harry van der Hulst, Rob Goedemans & Ellen van Zanten (eds.), *A survey of word accentual patterns in the languages of the world*, 55-85. Berlin: Mouton de Gruyter.

Greenberg, Steven & Eric Zee. 1979. On the perception of contour tones. *UCLA Working Papers in Phonetics* 45. 150-164.

Gussenhoven, Carlos & Jörg Peters. 2004. A tonal analysis of Cologne Schärfung. *Phonology* 21. 251-285.

Haraguchi, Shosuke. 1977. *The tone pattern of Japanese: an Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.

Hayes, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20. 253-306.

Himmelmann, Nikolaus & D. Robert Ladd. 2008. Prosodic description: an introduction for fieldworkers. *Language Documentation & Conservation* 2. 244-274.

Hoogshagen, Searle. 1959. Three Contrastive Vowel Lengths in Mixe. *Zeitschrift für Phonetik und Sprachwissenschaft* 12. 111-115.

Huss, Volker. 1978. English word stress in the post-nuclear position. *Phonetica* 35. 86-105.

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

Hyman, Larry M. 2010. How To Study a Tone Language, with exemplification from Oku (Grassfields Bantu, Cameroon). *UC Berkeley Phonology Lab Annual Report* (2010). 179-209.

Karlsson, Anastasia, David House & Jan-Olof Svantesson. 2012. Intonation adapts to lexical tone: the case of Kammu. *Phonetica* 69. 28-47.

Kutsch-Lojenga, Constance. 1996. Participatory research in linguistics. *Notes on linguistics* 73. 13-27.

Lippus, Pärtel, Karl Pajusalu, Jüri Allik. 2009. The tonal component of Estonian quantity in native and non-native perception. *Journal of Phonetics* 37. 388-396

Liu, Fang & Yi Xu. 2005.Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica* 62. 70-87.

Nakai, Satsuki, Sari Kunnari, Alice Turk, Kari Suomi & Riikka Ylitalo. 2008.. Utterance-final lengthening and quantity in Northern Finnish. *Journal of Phonetics* 37. 29-45.

Michael, Lev. 2011. The interaction of tone and stress in the prosodic system of Iquito (Zaparoan, Peru). *Amerindia* 35, 53–74.

Newman, Paul & Martha Ratliff 2001. *Linguistic Fieldwork*. Cambridge: Cambridge University Press.

Odden, David. 1995. Tone: African languages. In John A. Goldsmith (ed.), *The Handbook of Phonological Theory*, 444-475. Cambridge (MA): Blackwell Publishers.

Odden, David. 2011. The Representation of Vowel Length. In Marc van Oostendorp, Colin J. Ewen, Elizabeth Hume, & Keren Rice (eds.) *The Blackwell Companion to Phonology*, 465-490. Wiley-Blackwell.

Ohala, John J. 1989. Sound change is drawn from a pool of synchronic variation. In Leiv

E. Breivik & Ernst H. Jahr (eds.), *Language Change – Contributions to the Study of its Causes* (Trends in Linguistics 43), 173-198. Berlin: Mouton de Gruyter.

Ortega-Llebaria, Marta & Pilar Prieto. 2011. Acoustic Correlates of Stress in Central Catalan and Castilian Spanish. *Language and Speech* 54. 73-97.

Pickett, Velma, B., María Villalobos Villalobos & Stephen A. Marlett. 2010. Isthmus (Juchitán) Zapotec. *Journal of the International Phonetic Association* 40. 365-372.

Prince, Alan. 1980. A metrical theory of Estonian quantity. *Linguistic Inquiry* 11. 511-562.

Pike, Kenneth L. 1948. *Tone Languages: A Technique for Determining the Number and Type of Pitch Contrasts in a Language, with Studies in Phonemic Substitution and Fusion*. Ann Arbor: University of Michigan Press.

Remijsen, Bert. 2002. Lexically contrastive stress accent and lexical tone in Ma`ya. In Carlos Gussenhoven & Natasha Warner (eds.), *Laboratory Phonology 7*, 585-614. Berlin: Mouton De Gruyter.

Remijsen, Bert. 2013. Tonal alignment is contrastive in falling contours in Dinka. *Language* 89(2), 297-327.

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

Remijsen, Bert & Leoma Gilley. 2008. Why are three-level vowel length systems rare? Insights from Dinka (Luanyjang dialect). *Journal of Phonetics* 36. 318-344.

Remijsen, Bert & D. Robert Ladd. 2008. The tone system of the Luanyjang dialect of Dinka. *Journal of African Languages and Linguistics* 29. 149-189.

Riad, Tomas 1998. Towards a Scandinavian accent typology. In Wolfgang Kehrein & Richard Wiese (eds.), *Phonology and Morphology of the Germanic Languages*, 77-109. Tübingen: Max Niemeyer Verlag.

Rossi, M. 1971. Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 23. 1-33.

Sluijter, Agaath M.C. & Vincent J. van Heuven. 1996. Spectral tilt as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100(4). 2471-2485.

Snider, Keith. On Establishing Underlying Tonal Contrast. *Language Documentation and Conservation.* 8. 707-737

Stegen, Oliver. 2002. Derivational processes in Rangi. *Studies in African Linguistics* 31. 129-153.

Ternes, Elmar. 1989 [2nd edition], 2006 [3rd edition]. *The Phonemic Analysis of Scottish Gaelic. Based on the Dialect of Applecross, Ross-shire*. Hamburg: Helmut Buske Verlag.

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

Turk, Alice, Satsuki Nakai & Sugahara, Mariko. 2006. Acoustic segment durations in prosodic research: a practical guide. In Stefan Sudhoff, Denisa Lenertová, Roland Meyer, Sandra Pappert, Petra Augurzky, Ina Mleinek, Nicole Richter & Johannes Schliesser (eds.), *Methods in Empirical Prosody Research*, 1-28. Berlin: De Gruyter.

van der Leeden, A.C. 1993. *Ma'ya: A Language Study: A Phonology*. Jakarta: Lembaga Ilmu Pengetahuan Indonesia & Rijksuniversiteit Leiden.

van Zanten, Ellen, Rob Goedemans & Jos Pacilly. 2003. The status of word stress in Indonesian. In Jeroen van de Weijer, Vincent J. Van Heuven & Harry van der Hulst (eds.), *The Phonological Spectrum / Volume II: Suprasegmental Structure*, 151-175. Amster-

dam: John Benjamins.

Wheatley, Julian. 2013. Burmese. In Randy J. LaPolla & Graham Thurgood (eds.). *Sino-Tibetan Languages*. Routledge.

Wichmann, Søren. 1995. *The Relationship Among the Mixe-Zoquean Languages of Mexico*. Salt Lake City: University of Utah Press.

Williams, Briony. 1985. Pitch and duration in Welsh stress perception: the implications for intonation. *Journal of Phonetics* 13. 381-406.

Williams, Briony. 1986. An acoustic study of some features of Welsh prosody. In Catherine Johns-Lewis (ed.), *Intonation in Discourse*, 35-51. London: Croom-Helm.

Xu, Yi. 1999. Effects of tone and focus on the formation and alignment of f_0 contours. *Journal of Phonetics* 27. 55-105.

Xu, Yi & Xuejing Sun. 2002. Maximum speed of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America* 111. 1399-1413.

Yu, Alan C.L. 2008. The phonetics of quantity alternation in Washo. *Journal of Phonetics* 36. 508-520.

Yu, Kristine. The experimental state of mind in elicitation: illustrations from tonal fieldwork. *Language Documentation and Conservation.* 8. 738-777

Zhang, Jie. 2002. *The effects of duration and sonority on contour tone distribution: Typological survey and formal analysis*. New York: Routledge.

Bert Remijsen
b.remijsen@ed.ac.uk