Mental representation of tonal spreading in Bemba: Evidence from elicited production and perception

Nancy C Kula[1]* and Bettina Braun[2]
[1]*Department of Language and Linguistics, University of Essex, Colchester, Wivenhoe Park, UK*
[2]*Fachbereich Sprachwissenschaft, Universität Konstanz, Germany*
*Corresponding author, email: nckula@essex.ac.uk

Abstract: Previous research has shown that listeners from tonal languages are better at processing tone compared to speakers from non-tonal languages. However, most of this research has tested Asian tone languages, particularly those which have many tonal contrasts and a dense tone-to-syllable association. In this paper we investigate the mental representation of derived tones in Bemba, a Bantu language that has a two-way tone contrast but which shows robust tone spreading patterns. Specifically, we test ternary high-tone spreading, a process that is unique from a phonological perspective. In a production task we test whether ternary spread can be extended to non-words. We complement this with an AX discrimination task comparing binary vs ternary spread, which are phonologically contrastive, on the one hand, with a tonally similarly salient but non-phonologically relevant contrast, on the other. We show that in both the production and perception of non-words, ternary spread is distinct from binary spread, suggesting that derived tone is equally mentally represented as lexical tone is in Asian tone languages.

Introduction

Evidence suggests that speakers of tone languages, who utilise pitch to signal lexical differences, have different representations for tonal aspects than speakers of non-tonal languages (e.g. Van Lancker and Fromkin 1973, 1978; Gandour 1983, see below for more detail; Repp and Lin 1990; Lee et al. 1996; Ye and Connine 1999; Bent et al. 2006; Zheng et al. 2007; Francis et al. 2008; Braun and Johnson 2011; Chiao et al. 2011). These studies, conducted on Mandarin Chinese, Cantonese, Taiwanese and Thai, have a number of findings of significance, in particular: (i) different lateralisation for lexical (tone) and non-lexical (intonation) pitch processing (Van Lancker and Fromkin 1973, 1978); (ii) more integral processing of segments (and in particular vowels) and tones in tone language speakers (Braun and Johnson 2011; Repp and Lin 1990); (iii) an advantage for vowel over tone information in vowel and tone monitoring tasks (Ye and Connine 1999); and (iv) more categorical processing (in particular identification) of tone by tonal listeners (Sun and Huang 2012) but no difference in tone discrimination in AX tasks with short inter-stimulus intervals (Cutler and Chen 1997). (See Francis et al. 2003 for a more critical view on the categorical perception of tones).

However, tone languages are not homogenous and thus while the findings above hold of the Asian tone languages investigated, they cannot be assumed to also hold for Bantu tone languages which differ in many critical respects from Asian language tone (Yip 2002). First, Bantu tone languages have fewer tonal contrasts than seen in Mandarin, Cantonese, Taiwanese and Thai (2–3 as compared to 4–7). Second, in Bantu tone languages, only a few vowels/morphemes are lexically specified for tone, while nearly all vowels/morphemes are tonally specified in the type of Asian tone languages investigated. Third, in Bantu tone languages, the surface tonal realisation is derived from intricate tone-spreading rules (see e.g. Clements and Goldsmith 1984; Hyman and Kisseberth 1998, and below), while there are considerably fewer tonal (sandhi) processes in Mandarin, Cantonese, Taiwanese and Thai (Yip 2002: 173ff). While it is a legitimate assumption that tone is mentally represented as part of the vowel that carries it in the Asian tone languages investigated, this assumption is likely not tenable for Bantu tone languages, at least not for derived tones. If this is the case we expect tonal processing for speakers of this type of Asian language and

Bantu tone languages to differ significantly. However, to date there are basically no studies on the processing of tone in Bantu—let alone any African tone languages—to test this hypothesis. This study provides a first step towards filling this gap, by investigating the mental representation of tone in a Bantu language. We will particularly focus on how derived tones—resulting from tonal spreading processes—are produced and perceived by native speakers of Bemba.

Before we lay out the phonology of tone in Bemba in more detail, we will briefly discuss the literature on the processing of tone sandhi. These studies are highly relevant to our research question as they deal with cases of tonal alternations. Not surprisingly, previous studies show that tone sandhi processes (resulting in allotones) affect prelexical and lexical processing. For instance, a high falling /53/ tone in Cantonese (which has a high level /44/ allotone) was judged to be perceptually similar to both level and contour tones, although the allotone was not part of the experiment (Gandour 1981). In Mandarin Chinese, Tone2 and Tone3 are also highly confusable, probably due to the sandhi rules that turn a Tone3 into Tone2 before an adjacent Tone3 (Huang 2001).

Background on Bemba tone

Bantu languages generally have two tones (High and Low) and are renowned for having intricate tone spreading rules. The Bantu literature generally assumes that the H vs L tonal contrast can be lexically expressed as H vs Ø with low tones inserted late in the derivation (see e.g. Kisseberth and Odden 2003). We do not take a strong position on this here as it has no real bearing on the issues discussed. Our experiments will be on Bemba, a central Bantu language of Zone M in Guthrie's classification (Guthrie 1948; Guthrie 1967–1971). In particular, we focus on the verbal domain. Bemba, as other Bantu languages, is an agglutinative language that allows the addition of a number of morphemes to verbal stems. Lexical tone (High or Low) can be associated with the verb stem or with affixes preceding the verb stem. A typical Bantu verb structure consists of Subject Marker + TAM (tense-aspect-mood) + (Object Marker) + Verb Stem. Bemba tonology is discussed in Sharman and Meeussen (1955) and Bickmore and Kula (2013) where tone is shown to play both a lexical and a grammatical role. There are two main High tone (H henceforth) spreading processes central to the tonology of Bemba; unbounded spreading and bounded spreading. Unbounded spreading spreads a H rightwards up to the end of the verb form, targeting all following toneless moras in a phrase-final word. The examples in (1) show unbounded H spreading in a verb form where the initial mora of the class 2 subject marker is lexically H-toned (1a–b) and the following future marker -*ka*- and verb are lexically toneless. This contrasts with (1c) where the 1st plural subject marker is low-toned and the verb form therefore surfaces as all low.[1]

(1a) bá-ká-lúk-á
      2SM-FUT3-plait-FV
      'They will plait'
(1b) bá-ká-lóóndólól-á
      2SM-FUT3-explain-FV
      'They will explain'
(1c) tù-kà-lòòndòlòl-à
      1PL-FUT3-explain-FV
      'We will explain'

Unbounded H spreading contrasts with bounded H spreading which does not spread a H to the end of the verb form. There are two contexts where bounded H spreading applies. Examples in (2a–b) illustrate one of these—when the verb is followed by another constituent, here an adverb. Bounded H spreading contrasts two Bemba dialects, namely Copperbelt and Northern Bemba (CB and NB, respectively, henceforth) which differ with respect to the domain of bounded H spreading as ternary in CB and binary in NB.[2] The examples in (2a–b) thus illustrate CB and only differ from (1a–b) in not having unbounded H spreading on the verb form. To show that this is not influenced by the tone of the following constituent both a low-toned adverb (2a) and one with an initial H (2b) are used. (2c) shows that there is no H spreading when the subject marker is low-toned.

(2a) bá-ká-lúk-à   bwìínò
      2SM-FUT3-plait-FV   well
      'They will weave well'
(2b) bá-ká-lóòndòlòl-à   sáàná
      2SM-FUT3-explain-FV   a.lot
      'They will explain a lot'
(2c) tù-kà-lòòndòlòl-à   bwìínò
      1PL-FUT3-explain-FV   well
      'We will explain well'

The other context where bounded H spreading applies is when two Hs are separated by a number of toneless moras sufficient to allow bounded H spreading. This is best illustrated by a preceding lexical H that is followed by a final H as provided by, for example, the lexically H-toned post-verbal enclitic =kó.[3] Examples using the same lexically toneless verbs and future marker as in (1) and (2) are given in (3).

(3a) bá-ká-lóòndòlòl-à=kó
      2SM-FUT3-explain-FV-17LOC
      'They will also help to explain'
(3b) bá-ká-lúk-ìl-à=kó
      2SM-FUT3-plait-appl-fv-17LOC
      'They will plait in there'
(3c) tù-kà-lòòndòlòl-à= kó
      1PL-FUT3-explain-FV-17LOC
      'We will also help to explain'

The lexical H of the subject marker in (3a–b) spreads in bounded fashion (ternary spread for CB) because of the following H on the final mora of the verb form, provided by the enclitic. As in the examples in (2), bounded H spreading has a specified domain and does not continue to spread the H even when there are potential target toneless moras.

In terms of phonology, processes involving ternary spread are atypical because they are suggestive of counting, which is generally not found in grammatical patterns. Ternary spread is also disfavoured on account of its violation of contiguity and locality between the trigger and target of a process (see e.g. Odden 1994). Thus a number of analyses considering ternary spread of tone in Bantu have proposed two-staged analyses involving two separate but related processes rather than a single rule of ternary spread.[4] Some examples can be found in Myers (1987) for Shona dialects (Zezuru, Korekore and Northern Karanga); Casimjee (1986) for Tshivenda; and Hyman and Mathangwane (1998) for iKalanga, among others, but see Johnson and Paster (2012) for an analysis utilising a TRIPLING constraint in the Manyika dialect of Shona, and Hyman (1993) for a case possibly eluding a two-staged analysis. The Bemba (CB) case has also been argued as indicative of an analysis involving two processes which are independently motivated in Bickmore and Kula (2013). This means that first language learners of CB, who have no NB input or experience, are able to create the two separate rules in their L1 phonological system because they will encounter evidence for doing so. The other question that arises with ternary spread is whether it is merely a phonetic effect (phonetic overshoot), given work that some claimed cases of H spreading are really just due to the phonetic implementation of f0 where the high pitch simply extends a bit beyond its target (see e.g. Myers 1999). By looking at contrastive minimal pairs in CB, Bickmore and Kula (2013) argue that ternary spreading is phonological even though it is achieved by two independent processes which apply in tandem. This paper follows up on this work by investigating how ternary spread is processed and whether speakers are able to distinguish binary from ternary spread so as to establish whether a ternary tone span is cognitively represented and overall establish the status of derived tone as opposed to lexical tone.

As noted earlier the binary vs ternary spread pattern is determined by dialect. Note that the

contrast that is phonological in the verbal domain in each dialect is the robust distinction between bounded and unbounded spread and not the more fine-grained difference between binary and ternary spread. Since binary and ternary spread are both good exemplars of bounded spreading, we can question whether it matters to speakers which of the two is used, i.e. how categorical bounded spreading is in each dialect. If it is not categorical then we would expect that listeners are not very sensitive to the tonal difference between a binary and a ternary spread; whereas if it is, speakers will be very sensitive to the difference.

In this paper we focus on speakers from CB where bounded spreading is ternary. We aim to test whether this ternary spread also applies to non-words (suggesting inferential, rule-based processes (cf. discussion in Pinnow and Connine (2014)), and whether the domain of this bounded spreading is clearly specified, such that items with a binary spread are perceived as different. In Experiment 1, we investigate the application of bounded (ternary) spreading processes in the production of non-words, using a variant of the *Wug* test (Berko 1958; Ratner and Menn 2000). Specifically, participants heard real and non-word verb forms in one conjugation and had to conjugate them into another form (1st pl. to 3rd pl.). This task involves the application of a ternary spreading process in their variety. If ternary spreading is a rule-based process, we expect speakers to produce ternary spreading in non-words as well. If it is lexically encoded, we expect fewer ternary spreads in the non-word condition than in the real-word condition.

In Experiment 2, we complement the production data by behavioural results in a speeded AX discrimination task in which nonce stems with binary vs ternary spread are contrasted with a long inter-stimulus interval (ISI) of two seconds. Such a long ISI allows us to tap into phonological representations (Crowder and Morton 1969; Cowan and Morse 1986; Babel and Johnson 2010). If ternary spread is categorically specified, listeners should be sensitive to the difference between a binary and ternary spread. On the other hand, if ternary spread is just an allophonic variant of binary spread or is simply the result of phonetic temporal overshoot, we predict that listeners will have difficulties in perceiving the difference. We used two kinds of control conditions, one with a tonal difference that has an equally strong f0-difference but involves a paradigmatic difference (High vs Mid, Mid vs Low) instead of the syntagmatic difference between a binary and ternary spread.[5] The second control condition involves real words with a binary vs ternary spread to address whether participants perform a dialect discrimination (in which case participants' sensitivity to the tonal difference should be the same in the non-word and word condition) or discriminate the tonal patterns more phonologically (in which case we would expect no or a smaller difference in the real word condition, due to lexical activation, than in the non-word condition). Following Guenther and Gjaja (1996), who argue that linguistic experience warps the auditory cortex and hence enhances between-category perception, we predict that native speakers of Bemba are more sensitive to the binary-ternary difference than to the paradigmatic difference (High vs Mid, Mid vs Low) in non-words, despite similar acoustic differences.

In an earlier pilot study (Kula and Braun 2013) we used a more traditional *Wug* test including pictures with unknown activities. In that experiment, 26 participants heard a nonce infinitive verb form (e.g. úku-sena, produced as úkú-séná, i.e. with all syllables having high tones) and had to produce two altered forms, each starting with a lexically high-toned prefix subject marker. The first was a 3rd singular future form (i.e. á-ka-sena) to test unbounded spreading and the second was a 3rd plural future form with a following high-toned post-verbal clitic -ko (i.e. bá-ka-sena-kó) to induce bounded spreading. There was no consistent pattern in the results, leading us to conclude that the task was too difficult for participants. We attribute this to at least three factors: (i) participants had to indirectly access the information that the non-word verb stem was lexically low-toned by 'undoing' the unbounded spread; (ii) participants failed to treat the post-verbal clitic -ko as a following constituent owing to its close relation with the verb stem; and (iii) the production of two items in a row may have falsely led subjects to treat the second form as a 'following constituent' and therefore failed to show the expected unbounded spreading in the first form. The current experimental design remedied these problems and presented stimuli with all low-toned syllables, used a following independent word (rather than an enclitic) as subsequent constituent and subjects only had to produce one target form. We also included a sufficient number of real-word stems to make the task easier to comprehend.

Experiment 1

In Experiment 1 we first tested which kind of high-tone spreading patterns native speakers of Bemba from ternary spreading dialectal areas produce when conjugating real and non-word verbal forms in a bounded spreading setting. To this end, the verbal forms had to be complemented with a following constituent whose purpose was to trigger bounded spreading.

Specifically, the task involved the production of a novel form that required changing one morpheme of the verb form. The experiment hence tested participants' ability to produce ternary spread on novel forms as well as on existing, known forms. This paradigm taps into participants' representations of ternary spread. If ternary spread of lexically specified high tones is stored in the mental representation of the verb itself—through frequent experience of the forms (Johnson 1997; Goldinger 1998)—we expect more forms with a ternary spread in the real-word stems than the non-word stems. On the other hand, if ternary spread is generated by rule, however complex ternary spread may be for phonological rules to generate (Odden 1981; Myers 1987; Hewitt and Prince 1989; Hyman 1993; Creissels 1998; Hyman and Mathangwane 1998; Johnson and Paster 2012; Bickmore and Kula 2013), we expect a high proportion of ternary spread forms also in the non-word condition.

Methods

Participants

Twenty-three native speakers of Bemba (aged between 18 and 66 years, average 37.7 years, SD 14.3 years, 15 male, 8 female) participated for a small fee. Participants originated from three main Bemba-speaking areas (Northern Province, Luapula Province and Copperbelt Province) but had all spent a significant amount of time (between 1–34 years, average = 16.5 years, SD = 10.9 years) in the Copperbelt area (ternary dialect) and none resided in the northern area (binary dialect) at the time of testing. Some participants had spent time in areas that were not originally Bemba-speaking areas (Lusaka, Kabwe, Kaphiri Mposhi) but in which Bemba (more predominantly of the Copperbelt variety) is nowadays commonly used as a lingua franca. They had no prior training in intonation or intonational phonology and were unaware of the purpose of the experiment.

Materials

For the production task, we selected eight experimental stimuli each for the non-word and real-word conditions. Each item consisted of a trisyllabic verb stem (e.g. sakula, 'comb'). Four additional stimuli were used for the familiarisation phase (two with non-word stems and two with real-word stems). Furthermore, we selected ten filler items, which consisted of eight real verb stems and two non-word forms. These filler items all had four-syllabic verb stems. All real verb stems were lexically low-toned. All verb stems were presented in the habitual aspect, using the lexically low-toned first person plural subject marker *tu-* followed by the low-toned independent adverb *lyonse* (all the time). Stimuli were thus of the shape *tu-sakula lyonse* 'we comb all the time'.

All 30 items (16 experimental, 10 fillers, 4 familiarisation) were produced by a single female native speaker of Bemba from the ternary dialect and were recorded with a Zoom H4 recorder with a sampling rate of 44.1 kHz and a resolution of 16Bit. She first recorded the real words, with all syllables low-toned. She then recorded the non-word stems with the same prosody, keeping them similarly low-toned. The speaker was instructed to produce the items as naturally as possible and not to vary the pitch contour across items. Figure 1 shows example pitch tracks of the real and non-word verb stimuli. Acoustic analyses of the non-word and real-word stems showed no difference in f0 on the trisyllabic verb stem (see Table 1).

Table 1: Average f0 values (and standard deviations) of the four syllables in the stimuli of the real-word and non-word condition

	Average f0 (Hz)			
	1st syllable	2nd syllable	3rd syllable	4th syllable
real-word	140.5 (6.2)	132.4 (4.6)	126.0 (5.3)	117.0 (3.8)
non-word	133.4 (11.2)	127.0 (8.1)	124.1 (9.0)	119.7 (7.8)
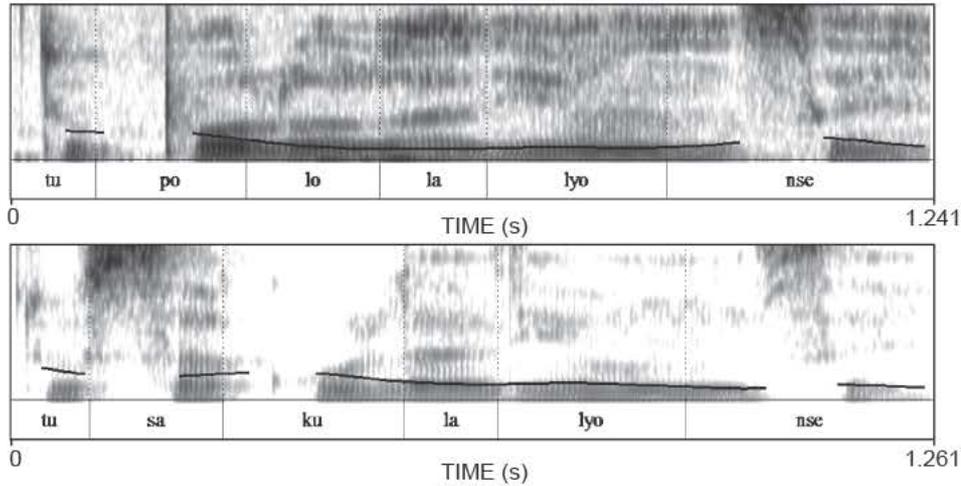
Figure 1: Example pitch tracks of stimuli (nonce word stem in top panel, real word stem in bottom panel) in Experiment 1. F0-range is shown between 100 and 300 Hz

Procedure

Participants were instructed to replace the low-toned subject marker *tu-* with the high-toned subject marker *ba-* (class2 ѕм) and retain the rest of the verb form, i.e. to use the same trisyllabic verb stem as in the recording they heard, also followed by the adverb *lyonse* 'all the time'. Target words in the *Wug* test would thus have the form *basakula lyonse* 'they comb all the time'. As earlier noted this context with a constituent following the verb was intended to induce bounded spreading, i.e. predominantly ternary spread.

Participants were tested individually in a quiet room in two locations at Nortec College in Ndola and Nkrumah University in Kabwe. They were seated in front of a laptop computer and were recorded using a Zoom H4 recorder (44.1 kHz, 16Bit). They were instructed orally in English and in Bemba that they would hear a verb form and would have to produce one form where they only changed the initial (CV-) subject marker. The experimenter was a native speaker of Bemba of the Copperbelt dialect.

Each of the verbal forms for the *Wug* task was set up using Microsoft PowerPoint. The participants saw an image of a loudspeaker on the screen that the experimenter clicked in order to play the recording. The sound files were played via headphones. There was no predefined time interval in which participants had to respond but they generally responded within 1 500 ms. Participants were allowed to listen to the recording a second time if they wished, an option that was used in only 10% of cases. The experiment was preceded by four practice trials whose responses were not recorded. The experiment lasted approximately 12 minutes. After the experiment, the participants filled out a questionnaire on their origin and their language background.

Results

The recordings were tonally coded per vowel in each syllable. The tonal realisation in the three syllables of the verb stem were annotated as H(igh), L(ow) or F(alling). The data were annotated by a student assistant who was trained in tone annotation and who was unaware of the purpose of the

Table 2: Average percentages of the different types of prosodic realisations for real and non-word stems

Stem type	Prosodic realisation			
	Unbounded	Ternary	Binary	Other
Real	5.7	88.9	4.2	1.2
Non-word	5.7	90.0	4.1	0.2

experiment, and by the first author, a native speaker of Bemba with training in tonal transcription. In case of disagreement, a consensus annotation was established by consulting a third evaluator with data discarded if no agreement could be reached. This happened only in a minority of cases (about 2% of cases). Next, the prosodic realisation of the items was classified as unbounded (all three syllables following *ba-* were realised with an H), ternary (spread of H until two syllables after *ba-*), binary (spread of H on one syllable following *ba-*) and other (in particular an L on the first syllable following *ba-* and then high on the next syllable after *ba-*: LHH or LHL). Average percentages of the prosodic realisations for real and non-words are shown in Table 2. The most frequent pattern is clearly a ternary H spread both in the real and non-word items. Binary spreading was very rare, but still present in some instances. Furthermore, there was a small, but non-negligible number of cases with unbounded H-tone spreading. These cases of unbounded spreading only occurred when there was a significant pause between the verb form and the following adverb. The few participants who produced two phrases with unbounded H-tone spreading in the verbal domain showed the same patterns in both real and non-word contexts.

The raw data were analysed using logistic mixed effects regression models (Baayen 2008; Baayen et al. 2008; Jaeger 2008) with *stem type* (real vs non-word) as fixed factor, ternary spread (yes vs no) as dependent variable and *participants* and *items* as crossed random factors (allowing for random intercepts and slopes). The results showed no effect of stem type, i.e. the model did not deteriorate by removing stem type. Overall, there were significantly more items with a ternary spread than items without binary spread ($p < 0.001$).

Discussion

The results from the real-word performance demonstrate a clear preponderance of the ternary spread pattern. The same tone spreading pattern is found for the non-word verb stems. This suggests that a linguistically meaningful tonal pattern (ternary spreading) can be generalised to novel forms. Thus, from an input that consisted of all low tones, participants were able to produce novel forms with a high-toned initial syllable that then showed ternary spread. As there was no high tone at all in the input, the high tones after the high-toned subject marker can only be interpreted as the result of the application of tonal spreading rules. In this sense these results differ from the findings on Mandarin, Cantonese, Taiwanese and Thai tone where tone is seen as a lexical property of a vowel. The Bemba production results thus support an autosegmental characterisation of tone, i.e. of tone being independent of the host segment and therefore able to spread to adjacent positions.

The non-words were framed in a context that included real subject and aspectual markers as well as a following independent, real adverb (lyonse 'all the time'). As the adverb itself was also low-toned, our results support the analysis of bounded spreading as occurring when a verb form is not final in its phrase. Note that this account of bounded spreading is also supported by the cases in which participants showed unbounded spreading since this pattern only occurred when participants produced a significant pause between the verb and the following adverb, thereby placing the two words in different (phonological) phrases. The presence of a phonological phrase boundary following the verb then induced unbounded spreading of the rightmost high tone. (See Kula and Bickmore 2015 for detailed discussion of phrasal tone in Bemba).

Note that the use of trisyllabic verb stems throughout the experiment and the high proportion of real-word stems (60%) may have led to a prosodic priming effect, so that participants used the same tonal patterns in the non-word stems without applying phonological knowledge (see e.g. Bock (1986) for syntactic priming). This interpretation would imply that the ternary spread in the non-words was not present early in the experiment but became more and more frequent over the course of the experiment. This is not what we find, however, with ternary spreading already produced in the first trials and remaining equally frequent throughout the experiment. Nevertheless, future experiments will have to use more varied phonological forms in the stimuli so as to be able to attribute the findings to phonological processing alone.

Experiment 2

In Experiment 2 we tested how specific the mental representation of derived ternary spread is. To

this end we conducted a speeded AX (same-different) task with a 2 000 ms inter-stimulus interval, which has been argued to tap into phonological representations (Crowder and Morton 1969; Cowan and Morse 1986; Babel and Johnson 2010). We compared participants' sensitivity to the syntagmatic tonal difference between a binary and a ternary spread (i.e. HHLLL vs HHHLL, henceforth *binary-ternary*) to an acoustically similar, but paradigmatic pitch difference. This control condition involves the tonal difference between H and mid-level (HHMM vs HHHH, henceforth HM-HH) or between L and mid-level (HHMM vs HHLL, henceforth HM-HL). Since mid-level tones do not occur in Bemba, we hypothesise that participants are more sensitive to the syntagmatic tonal difference in the binary-ternary context compared to the paradigmatic tonal difference in the control condition. Since such a result is also compatible with the alternative interpretation that participants represent and store a dialect difference in this task (see introduction), we included a further control condition with real-word stimuli and the same binary-ternary difference. If the sensitivity to the binary-ternary spread is due to dialect discrimination, we should find the same sensitivity in the real and non-word conditions. If the difference between the binary-ternary spread is represented and stored at the sub-lexical level, we expect higher sensitivity to the binary-ternary difference in the non-word than in the real-word condition. This hypothesis stems from findings that show that the activation of lexical representations makes it difficult to perform sub-lexical decisions (e.g. McQueen 2005). We provide a summary of the aim and predicted findings of this rather complex experiment in (4) and Table 3 below for ease of reference.

(4)   Experiment 2 summary
        Hypothesis: Do Bemba speakers have a mental representation of a ternary tone span?
        Task: AX-discrimination

Methods
Participants
Participants were the same as in Experiment 1. All participants first completed Experiment 1 and then conducted Experiment 2.

In addition to the target Bemba participants, a group of Mandarin Chinese speakers were tested as a way of evaluating the psychoacoustic similarity of the different tonal patterns across the stimuli. The inclusion of these participants is thus used to further validate the stimuli. Mandarin Chinese speakers were chosen because they speak a tone language in which tone is specified for each vowel. There were 16 native speakers of Mandarin (aged between 19 and 27 years, average 22.8 years, SD 2.4 years, 13 female, 3 male), who were all students at the University of Essex at the time of testing, and participated for a small fee. They were tested individually in a quiet room (see description below). None of the participants had prior training in intonation or intonational phonology and they were all unaware of the purpose of the experiment.

Materials
For the AX task, we selected eight stimuli each for the real/non-word binary vs ternary spread

Table 3. Experiment 2 conditions and stimuli

Conditions	Description	Stimuli shape	Poor discrimination implies
Experimental condition	Binary vs Ternary (non-words)	HHLLL vs HHHLL (5 syllables each)	No phonological mental representation of a ternary span, M-H contrast nor M-L contrast
Control condition 1	(a) Mid vs High (non-words)	HHMM vs HHHH (4 syllables each)	
	(b) Mid vs Low (non-words)	HHMM vs HHLL (4 syllables each)	
Control condition 2	Real vs Non-words (binary vs ternary)	HHLLL vs HHHLL (5 syllables each)	Lexical effect – real words block tone discrimination

condition, as well as eight stimuli each for the non-word HM-HH and HM-HL control conditions. Furthermore, we chose 16 non-word stimuli as fillers and 4 for the practice phase.

The stimuli for the non-word binary-ternary spread condition all consisted of two real prefixes á- (3SG.SM, used four times) or bá- (2SM, used four times) followed by -ka- (FUT3, used four times, combined twice with á- and twice with bá-) or -la- (HAB, combined twice with á- and twice with bá). The stem was a non-word form consisting of a sequence of three syllables (e.g. tekeba) followed by a real post-verbal clitic -kó ('a bit/on behalf of/locative'), used in the verb form of each test item. The stimuli in the real-word binary-ternary spread condition had the same distribution of prefixes but had a real trisyllabic stem (e.g. nenuka 'become a bit loose'). The stimuli for the nonce words for HM-HH and HM-HL also had two real prefixes (á- and -ka- used four times in each of the HM-HH and HM-HL conditions). The stems in these 16 stimuli were always non-word bisyllabic stems (e.g. fusa). For the fillers, we constructed 16 stimuli that consisted of four-syllable non-word forms (e.g. falimasa). For the practice items, two of the four practice trials were similar in structure to the binary-ternary condition, one to the HM-HH condition and one to the filler condition.

The 32 experimental stimuli (8 non-word binary-ternary, 8 real binary-ternary, 8 HM-HH and 8 HM-HL), 16 fillers and the 4 practice stimuli were read twice in isolation by a female native speaker of Bemba. We used a single speaker for all the stimuli to avoid listener judgments from being influenced by voice characteristics. Recordings took place in a sound-attenuated cabin at the University of Essex. Data were directly digitised using a Zoom H4 recorder with a sampling rate of 44.1 kHz and a resolution of 16Bit.[6]

The speaker was instructed to produce the different stimuli in one condition as consistently as possible. For the binary-ternary spread condition, the difference was in the spread of the initial high tone. It lasted only until the second syllable in the binary spread condition (and was hence low or falling on the stem-initial syllable); the remainder of the syllables were produced with a low tone. In the ternary spread condition, the initial high tone lasted until the third syllable. Example pitch tracks of the binary-ternary spread condition are shown in Figure 2.
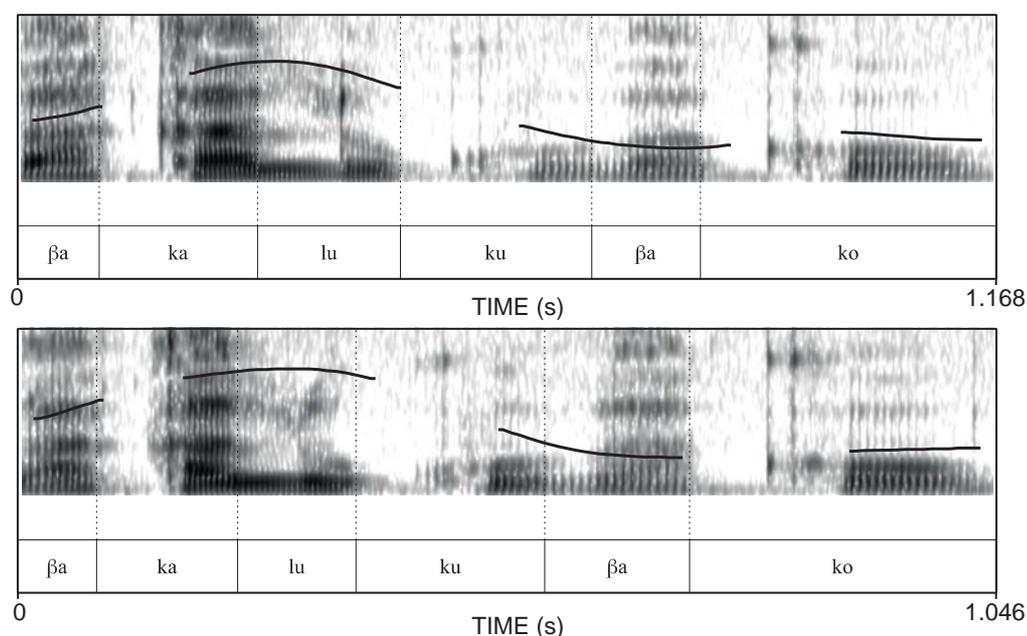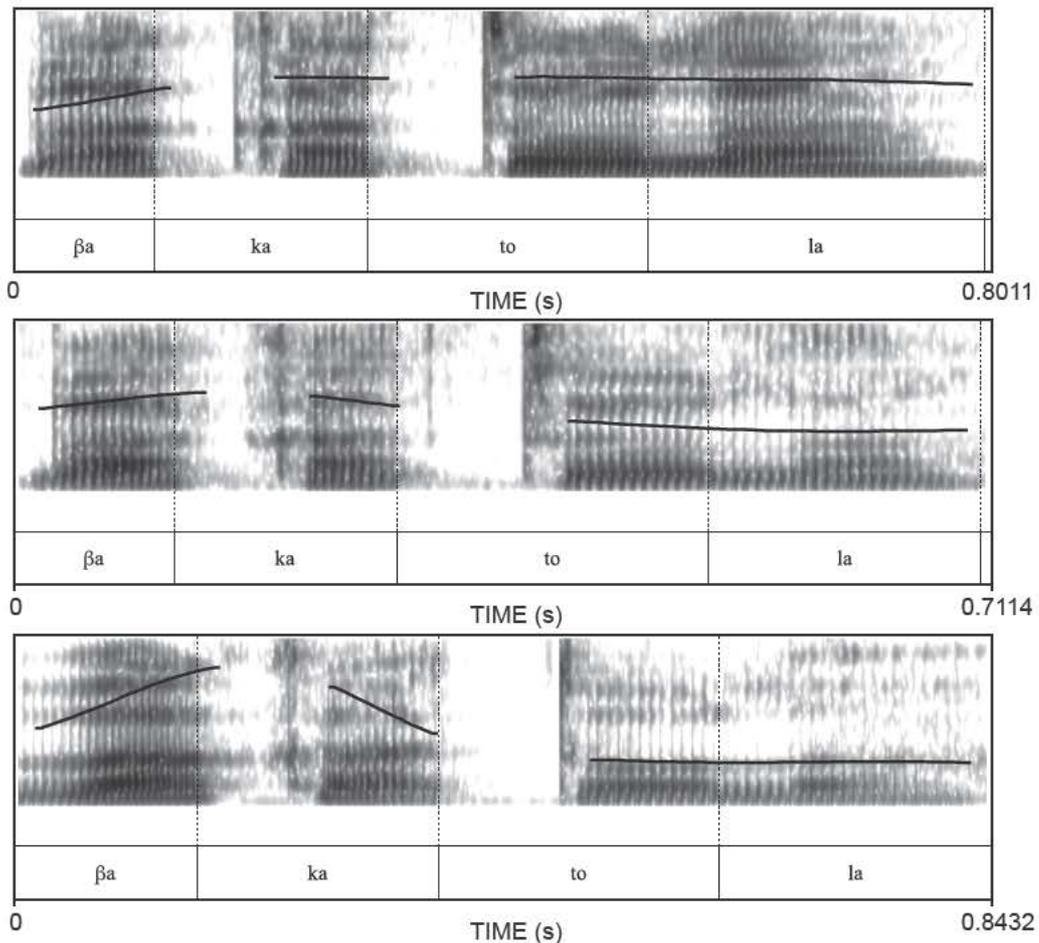

Figure 2: Example pitch tracks of stimuli in the experimental binary-ternary spread condition with non-words. The upper panel shows the binary spread (tone on third syllable is already falling), the lower panel the ternary spread (tone on third syllable is still high). F0-values are shown between 80 and 200 Hz

For the paradigmatic tonal comparisons, the HH condition had all four syllables produced with a high tone. The HL condition had the first two syllables produced with a high tone and the last two with a low tone. And in the HM condition, the first two syllables were produced with a high tone and the last two with a mid tone. Example pitch tracks of the paradigmatic tone difference are shown in Figure 3. The filler stimuli were produced with a high flat pitch (all high), a low flat pitch (all low) or a linearly declining high-to-low pattern (falling).

The recordings of all conditions (binary, ternary, HM, HH, HL) were analysed acoustically to ensure that the pitch difference was realised as intended. Three items were excluded from statistical analysis for non-ideal f0 differences of the pairs: one non-word and one real word in the binary-ternary condition and one non-word in the HM-HH and HM-HL conditions.

To quantify the distinction between a ternary and a binary spread, we measured the f0-value in the middle of the vowel of the third syllable (i.e. the first syllable of the stem) of all tokens with a binary and ternary spread, words and non-words (32 each). This difference amounted to 21.6 Hz. The tokens with a ternary spread had an average f0-value of 181.4 Hz (SD = 9.0 Hz), and the tokens with a binary spread had an average f0-value of 159.8 Hz (SD = 10.7 Hz). Note that this difference is also stable for the word and non-word group separately (average f0-difference for words: 30.2 Hz, average

Figure 3: Example pitch tracks of stimuli in the HH (upper panel), HM (medial panel) and HL conditions (lower panel). F0-values are shown between 80 and 200 Hz

f0-difference for non-words: 13.5 Hz). For the tokens with HM-HH and HM-HL, we also analysed the f0-value in the middle of the vowel of the third syllable. It was on average 130.8 Hz for the HM condition (SD = 9.0 Hz), 152.4 Hz for the HH condition (SD = 5.1 Hz) and 115.4 Hz, in the HL condition (SD = 4.1 Hz). Importantly the f0-differences were very similar in the three non-word conditions of interest.

Procedure

For the 'same' trials, two tokens of one type were presented together (different recordings of the same stimulus). For the 'different' trials, recordings with a pitch difference were combined. In each list, we presented 48 'same' trials and 48 'different' trials, preceded by the four practice trials (two as 'same' and two as 'different' trials). Participants hence heard each stimulus four times, twice in a 'same' condition (in two stimulus orders) and twice in a 'different' condition (in two stimulus orders). Trials were pseudo-randomised to generate the experimental lists, such that no more than three response types (same or different) followed each other and such that two stimuli of the same condition were separated by at least one trial of a different condition.

A trial started with a beep (500 Hz, 300 ms). After a period of silence of 1 000 ms the first stimulus was played. The second stimulus followed after an inter-stimulus interval of 2 000 ms. Participants entered their responses by pressing one of two buttons of a button-box (right button for right-handed participants, left button for left-handed participants) or different (opposite assignment of keys). Responses and response times were recorded for a period of 2 000 ms. After that, the next trial started. The experiment was controlled using Presentation software (Neurobiological Systems).

Participants were tested individually in a quiet room at Nortec College in Ndola and Nkrumah University in Kabwe. They were seated in front of a laptop computer and wore headphones. They were instructed orally in English and in Bemba to press the right button if the two stimuli they heard were identical (for left-handed participants, the button box was reversed and they pressed the left button for 'same' trials) and the left button otherwise. They were told to decide as quickly as possible. The experiment lasted approximately 20 minutes. After the experiment, they filled in a questionnaire on their origin and their language background.

Results

Participants' responses (excluding the three items mentioned above) were converted to *d'* scores, a measure of sensitivity for the detected difference (Macmillan and Creelman 2005). *d'* scores are based on hits (correct detection of an existing difference) and false alarms (detection of a difference that is not in the signal). We calculated *d'* scores for each participant for each of the four conditions (word and non-word binary-ternary, paradigmatic HM-HH and HM-HL). In our dataset a *d'* score larger than 1.4 signals a significant sensitivity to the respective difference. *d'* scores were analysed in *R* using linear-mixed effects regression models with *condition* as fixed factor and *participant* as random factor (Baayen 2008; Baayen et al. 2008; Cunnings 2012). *P*-values were calculated using the Satterthwaite approximation in the R-package lmerTest (https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf).

Bemba participant results

For the Bemba participants, we will first compare participants' sensitivity to the binary vs ternary spread non-word condition to the HM-HH and HM-HL non-word conditions. In a second step, we analyse whether the lexicality of the stimuli in the binary-ternary spread condition affected participants' sensitivity to this tonal difference.

To investigate whether participants are sensitive to the difference between a binary and ternary spread, we compared their sensitivity to this difference to their sensitivity to the difference between a mid tone and a high vs low tone (HH-HM and HM-HL). Results showed a significant main effect of *condition*: participants were significantly more sensitive to the distinction between a binary vs ternary spread (average *d'* 1.47) than to the difference between HM and HH (average *d'* 1.07, ß = 0.39, SE = 0.10, t = 3.7, p < 0.0005) or to the difference between HM and HL (average *d'* 1.12, ß = 0.34, SE = 0.10, t = 3.2, p < 0.05). The difference between HM-HH and HM-HL was not significant (p > 0.6). Average *d'* scores from the statistical model are shown in Figure 4.

Regarding the difference between real and non-word stimuli with a binary-ternary spread, we finally compared participants' sensitivity to these two conditions. Results showed a main effect of *condition*: Participants were less sensitive to the difference in tone spreading when the stimuli were real words (average *d'* 0.90) than when they were non-words (average *d'* 1.47, ß = 0.56, SE = 0.13, $t = 4.4$, $p < 0.0001$) (see Figure 5).
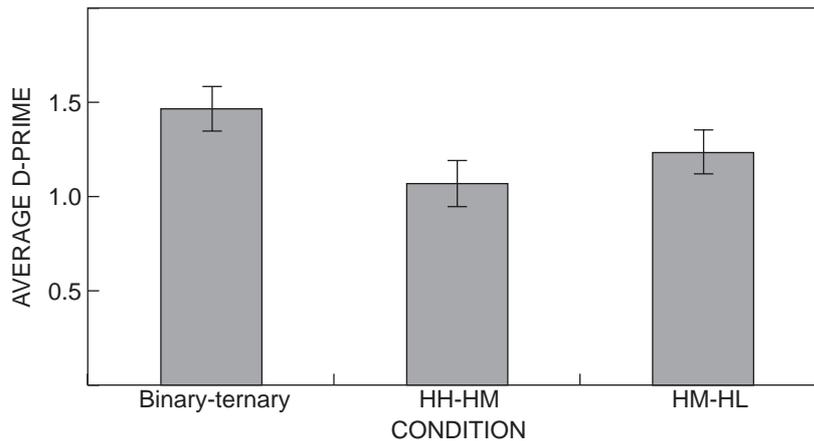
Figure 4: Averaged scores in the non-lexical conditions (binary-ternary vs HH-HM and HM-HL). Average values are based on the statistical model. Whiskers represent ± 1 standard error.
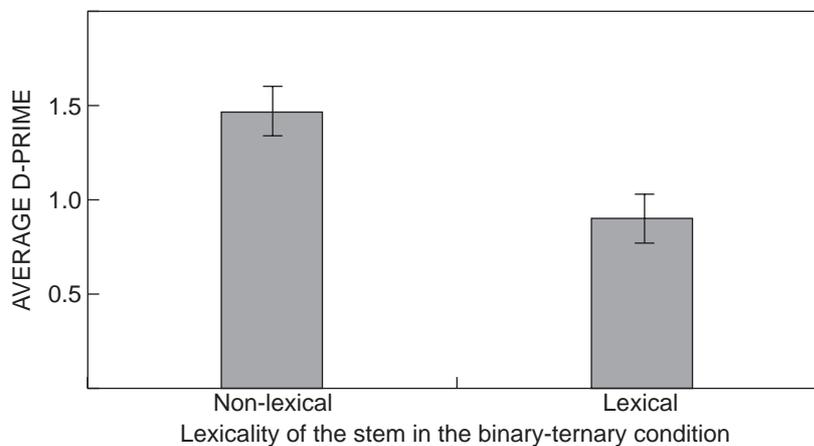
Figure 5: Average d' scores in the lexical vs non-lexical binary-ternary spread condition. Average values are based on the statistical model. Whiskers represented standard errors

Table 4: Average d' scores (and standard deviations) of the two language groups in the four conditions

| Language | Condition | | | |
	non-word binary/ternary	real-word binary/ternary	HM-HH	HM-HL
Bemba	1.47 (0.64)	0.90 (0.77)	1.07 (0.67)	1.12 (0.53)
Chinese	1.47 (0.65)	1.22 (0.50)	1.70 (0.60)	1.56 (0.34)

Chinese participant results

The data of the Chinese participants showed no effect of condition in the non-word conditions (binary-ternary vs HH-HM and HM-HL), resulting in a significant interaction between condition and language ($p < 0.01$). The Chinese participants hence showed an equal sensitivity to all four conditions (on average 1.49, see Table 4), in line with the acoustic analyses that showed equal f0 differences in the four conditions.

Discussion

Listeners were sensitive to the difference between a binary and ternary spread in the non-word condition. Furthermore, our data show a clear difference in performance between the binary-ternary difference and the equivalent tonal difference in the paradigmatic control condition (HM-HH and HM-HL pairs). Participants were significantly more sensitive to the binary-ternary difference than to the difference in the control condition. Given the long inter-stimulus-interval of 2 seconds, this pattern of results suggests that the ternary spread is not just an allophonic (or overshoot) variant of a binary spread but is clearly phonologically different from it. The rather poor performance in the real-word binary-ternary condition suggests that the good performance in the corresponding non-word condition cannot be solely explained by dialect discrimination. If participants represented the items in the binary-ternary spread condition as signalling different dialects, we would expect listeners to be equally sensitive to this sociolinguistic difference in the non-word and real-word conditions. This was not the case, though. Our discrimination results therefore provide evidence that derived tones have clearly specified spreading properties and that even small deviations from that pattern are detected since we observe that a ternary tone span is distinguished from a binary tone span.

Furthermore, the data from the paradigmatic control condition involving mid-level tones suggests that the mid-level tones were indistinguishable from both the high and the low tones in this language, suggesting a perceptual magnet (Kuhl and Iverson 1995). Mapping of two non-native categories onto a single native sound category is often found in the context of non-native/L2 listening (Best 1995; Cutler et al. 2006; Chiao et al. 2011). What is particularly striking in participants' behaviour in the control condition is that the mid-level tone appears indistinguishable from both the high *and* the low tone. These data suggest that a mid-level tone is functionally equivalent to a high or a low tone, depending on the context and that the tonal difference between two adjacent syllables ought to be large enough (i.e. at least the 15 Hz difference we tested) in order to perceive a change in tone.

The lexicality effect in the binary-ternary condition replicates findings from the segmental domain that show that lexical activation makes sub-lexical processing more difficult. This is the first study to show that this lexical dominance also affects tonal information, i.e. lexical activation appears to override tonal information.

General discussion

We tested the mental representation of bounded spreading (the ternary spread of a lexically specified H tone) in both non-word production and AX discrimination with a long inter-stimulus interval, comparing participants' sensitivity to a paradigmatic tonal difference (HM-HH vs HM-HL) and to Chinese listeners. The validity of ternary spread as linguistically meaningful in Bemba was supported by the fact that speakers are able to extend their mental knowledge of ternary spread in the production of non-words (Experiment 1). In conditions where speakers only heard low-toned forms they were able to replace an initial low-toned prefix with a high-toned one and consistently show ternary spread when the verb was not in final position, i.e. had a following constituent. This was further supported by the perceptual discrimination of non-words in the binary-ternary condition which was treated as more salient in comparison to phonologically non-distinctive pairs of words (HM-HH, HM-HL) that could not be distinguished with the same accuracy (Experiment 2). Both results suggest that Copperbelt Bemba speakers have a higher sensitivity to linguistically relevant phonological information (ternary spread) than to non-linguistically relevant information (binary spread).

In future work, whether the attested lexical effect is categorical or whether it can be modulated by an incremental 'wordiness' of the non-words needs to be tested. This is easily achievable, particularly

in a high affixing agglutinative language like Bemba, in which the number of real affixes can be manipulated. Such effects of wordiness have been shown to affect phonemic processing. Connine et al. (1997), for instance, showed that listeners were slowed down more in phoneme-monitoring the less the target resembled a real word (e.g. domestic, tomestic, fomestic). The prediction would be that the more real the word gets in terms of segments, the less sensitivity to tonal information there will be (as long as the forms do not involve lexically contrastive words or part-words, an issue that would have to be tested separately).

Note that the acoustic tonal difference in the binary-ternary condition was rather small, which may provide an explanation for why the sensitivity was not at ceiling. In future work, we plan to use resynthesised stimuli to be sure the difference results from binary vs ternary spread and not from other differences in the stimuli.

Conclusion

The goal of this paper was to elaborate on the status of derived tone in Bemba, which was evaluated by how well native speakers of Copperbelt Bemba attended to ternary high-tone spread in phonological processing. A *Wug* test showed that the ternary spreading pattern is readily applied to non-words, validating the status of derived tone. Hence, even though ternary spreading is not easily accounted for in phonology, it is readily applied by native speakers of Bemba. Furthermore, a perception experiment showed that listeners were more sensitive to the tonal difference in a binary vs ternary spread—which is linguistically relevant—than to an acoustically similar tonal difference that did not signal phonologically relevant information. However, the sensitivity to the binary-ternary spreading difference was absent when the stimuli were real words of Bemba. The lexical activation seemed to immediately override the tonal composition, suggesting that the tonal specification is weaker than it is in some Asian tone languages at the sub-lexical level (Braun and Johnson 2011). Overall, the paper provides strong evidence for the phonological representation of ternary tone spans in Bemba. Furthermore, it shows that derived tone has an independent mental representation despite not being lexically associated with vowels in the way that lexical tone is.

While this paper makes a contribution in an area of research that remains largely untapped in Bantu languages, future work will have to be conducted and, in particular, improvements in methodology must follow in order to confirm and substantiate the current findings. For a more thorough comparison between particular Asian and Bantu languages, syllable duration and word length as well as constancy in phonetic values for tone will have to be considered. For the evaluation of results in the experiments, less subjective evaluation and categorisation of data, perhaps relying on more objective acoustic measurements, is likely to produce more reliable results. The relative nature of Bantu tone might, however, suggest that perhaps the best methodology is indeed the use of independent judges (as was partially achieved in this study), but with some prior calibration that would make judges comparable. An increase in the number of test items in the production study would also improve reliability. And in the perception study, apart from the need for synthesised data, within-group variation of f0 in the HH, HM and HL conditions would have to be controlled. Although there are a number of issues and challenges still to be overcome in this area of research in Bantu languages, we hope through this initial study to have contributed to the development of experimental work in African languages and demonstrated the viability of the undertaking.

Notes

[1] The following abbreviations are used in the data: SM = subject marker; OM = object marker; HAB = habitual; FV = final vowel; LOC = locative; PL = plural; FUT = future; F1/2/3 refer to different futures. In all examples, high tone is marked by an acute accent and low tone with a grave accent. Underlying lexical high tone is underlined.

[2] CB is spoken in central parts of Zambia and NB in the Northern Province. Historically CB descends from NB whose inhabitants are considered to be located in the Bemba heartland with migrations starting around the 1920s to as recently as the 50s and 60s. Tonal differences between these two dialects are further discussed in Bickmore and Kula (2013) from which the current work draws. Other dialects also exist but much less linguistic work has been conducted on them.

References

Baayen HR. 2008. *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen HR, Davidson DJ, Bates DM. 2008. Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59(4): 390–412.

Babel M, Johnson K. 2010. Acessing psycho-acoustic perception and language-specific perception with speech sounds. *Laboratory Phonology* 1: 179–205.

Batibo H. 1991. The two directional tone melody spread in Sukuma. In: Hubbard K (ed.), *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistic Society: Special Session on African Language Structures.* Berkeley: BLS.

Bent T, Bradlow AR, Wright BA. 2006. Influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance* 32: 97–103.

Berko J. 1958. The child's learning of English morphology. *Word* 14: 150–177.

Best CT. 1995. A direct realist view of cross-language speech perception. In: Strange W (ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*. Baltimore: York Press. pp. 171–204.

Bickmore L, Kula NC. 2013. Ternary spreading and the OCP in Copperbelt Bemba. *Studies in African Linguistics* 42(2): 101–132.

Bock JK. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18: 355–387.

Braun B, Johnson EK. 2011. Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics* 39(4): 585–594.

Cassimjee F. 1986. *An autosegmental analysis of Venda tonology*. PhD dissertation, University of Illinois, Urbana-Champagne.

Chiao W-H, Kabak B, Braun B. 2011. *When more is less: Non-native perception of level tone contrasts*. Paper presented at the Psycholinguistic Representation of Tone Conference, Hong Kong

Clements N, Goldsmith J. 1984. *Autosegmental studies in Bantu tone.* Dordrecht: Foris Publications.

Connine CM, Titone D, Deelman T, Blasko D. 1997. Similarity mapping in spoken word recognition. *Journal of Memory and Language* 37: 463–480.

Cowan N, Morse P. 1986. The use of auditory and phonetic memory in vowel discrimination. *Journal of Acoustic Society of America* 79: 500–507.

Creissels D. 1998. High tone domains in Setswana. In: Hyman LM, Kisseberth C (eds), *Theoretical Aspects of Bantu Tone*. Stanford: CSLI Publications. pp. 195–229.

Crowder RG, Morton J. 1969. Precategorical acoustic storage (PAS). *Perception & Psychophysics* 5: 365–373.

Cunnings I. 2012. An overview of mixed-effects statistical models for second language researchers. *Second Language Research* 28(3): 369–382.

Cutler A, Chen HC. 1997. Lexical tone in Cantonese spoken-word processing. *Perception & Psychophysics* 59: 165–179.

Cutler A, Weber A, Otake T. 2006. Asymmetric mapping from phonetic to lexical representations in second-language listening. *Journal of Phonetics* 34(2): 269–284.

Fivaz D. 1970. *Shona Morphophonemics and Morphosyntax*. Johannesburg: University of the Witwatersrand Press.

Francis A, Ciocca V, Ng BKC. 2003. On the (non)categorical perception of lexical tones. *Perception & Psychophysics* 65(7): 1029–1044.

Francis AL, Ciocca V, Ma L, Fenn K. 2008. Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics* 36: 268–294.

Gandour J. 1981. Perceptual dimensions of tone: Evidence from Cantonese. *Journal of Chinese Linguistics* 9: 20–36.

Gandour J. 1983. Tone perception in Far Eastern languages. *Journal of Phonetics* 11: 149–175.

Goldinger SD. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105(2): 251–279.

Goldsmith J. 1985. On tone in Sukuma. In: Goyvaerts D (ed.), *African Linguistics: Essays in Memory of MWK Semineke*. Amsterdam/Philadelphia: John Benjamins.

Guenther FH, Gjaja M. 1996. The perceptual magnet effect as an emergent property of neural map formation. *Journal of the Acoustical Society of America* 100: 1111–1121.

Guthrie M. 1945. *The tonal structure of Bemba*. PhD dissertation, University of London.

Guthrie M. 1948. *The Classification of Bantu Languages*. London: Oxford University Press.

Guthrie M. 1967-1971. *Comparative Bantu* (4 vols.). Farnborough: Gregg.

Hewitt M, Prince A. 1989. The N. Karanga verb. In: Fee E, Hunt K (eds), *Proceedings of the West Coast Conference of Formal Linguistics.* Stanford: Stanford Linguistics Association. pp. 176–191.

Huang T. 2001. The interplay of perception and phonology in Tone 3 sandhi in Chinese Putonghua. In: Hume E, Johnson K (eds), *Studies on the Interplay of Speech Perception and Phonology. OSU working papers in linguistics* 55: 23–42.

Hyman LM. 1993. Structure preservation and postlexical tonology in Dagbani. In: Kaisse E, Hargus S (eds), *Studies in Lexical Phonology*. New York: Academic Press. pp. 235–254.

Hyman LM, Kisseberth C. 1998. *Theoretical aspects of Bantu tone*. Stanford: CSLI Publications.

Hyman LM, Mathangwane JT. 1998. Tonal domains and depressor consonants in Ikalanga. In: Hyman LM, Kisseberth C (eds), *Theoretical Aspects of Bantu Tone*. Stanford: CSLI Publications. pp. 195–229.

Jaeger TF. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4): 434–446.

Johnson K. 1997. Speech perception without speaker normalization: An exemplar model. In: Johnson K, Mullennix JW (eds), *Talker Variability in Speech Processing*. San Diego: Academic Press.

Johnson M, Paster M. 2012. Manyika verbal tonology. Unpublished manuscript, Pamona College.

Kisseberth C, Odden D. 2003. Tone. In: Nurse D, Philippson G (eds), *The Bantu Languages. London*: Routledge. pp. 59–70.

Kuhl PK, Iverson P. 1995. Linguistic experience and the 'Perceptual Magnet Effect'. In: Strange W (ed), *Speech perception and linguistic experience: issues in cross-language research*. Baltimore: York Press. pp. 121–154.

Kula NC, Bickmore L. 2015. Prosodic phrasing in Copperbelt Bemba. *Phonology* 32(1): 147–176.

Kula NC, Braun B. 2013. *Tonal harmony in Bantu: An experimental study on the mental representation of tone.* Paper presented at the 21st Manchester Phonology Meeting, Manchester, UK.

Lee Y-S, Vakoch DA, Wurm LH. 1996. Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research* 25(5): 527–542.

Macmillan NA, Creelman CD. 2005. *Detection Theory: A User's Guide* (2nd edn). Mahwah: Lawrence Erlbaum Associates.

Mann M. 1977. An outline of IciBemba Grammar. In: Kashoki ME (ed.), *Language in Zambia: Grammatical Sketches.* Lusaka: Institute for African Studies. Reprinted 1999, Lusaka: Bookworld Publishers.

Marten L, Kula NC. 2014. Benefactive and substitutive applicatives in Bemba. *Journal of African Languages and Linguistics* 34(1): 1–44.

McQueen JM. 2005. Speech perception. In: Lamberts, KGR (ed.), *The Handbook of Cognition.* London: Sage Publications. pp. 255–275.

Myers S. 1987. Tone and the structure of words in Shona. PhD dissertation, University of Massachusetts, Amherst.

Myers S. 1999. Tone association and F0 timing in Chichewa. *Studies in African Linguistics* 28(2): 215–239.

Odden D. 1981. *Problems in tone assignment in Shona.* University of Illionois.

Odden D. 1994. Adjacency parameters in phonology. *Language* 70: 289–330.

Patin C. 2009. Tone shift and tone spread in the Saghala Nominal Phrase. *Faits de Langue - Les Cahiers* 1: 230–247.

Pinnow E, Connine CM. 2014. Phonological variant recognition: Representations and rules. *Language and Speech* 57(1): 42–67.

Ratner NB, Menn L. 2000. In the beginning was the wug: Forty years of language elicitation studies. In: Menn L, Ratner NB (eds), *Methods for Studying Language Production.* Mahwah: Lawrence Erlbaum Associates. pp. 1–26.

Roberts R. 1992. A non-metrical theory of Sukuma tone. *OSU Working Papers in Linguistics* 41: 135–148.

Repp B, Lin, H-B. 1990. Integration of segmental and tonal information in speech perception: a cross-linguistic study. *Journal of Phonetics 18*: 481–495.

Sharman JC. 1956. The tabulation of tenses in a Bantu language (Bemba: Northern Rhodesia). *Africa* 26: 29–46.

Sharman JC, Meeussen AE. 1955. The representation of structural tones, with special reference to the tonal behavior of the verb in Bemba, Northern Rhodesia. *Africa: Journal of the International African Institute* 25(4): 393–404.

Sun K-C, Huang T. 2012. A cross-linguistic study of Taiwanese tone perception by Taiwanese and English listeners. *Journal of East Asian Linguistics* 21(3): 305–327.

van Lancker D, Fromkin V. 1973. Hemispheric specialization for pitch and "tone": Evidence from Thai. *Journal of Phonetics* 1: 101–109.

van Lancker D, Fromkin V. 1978. Cerebral dominacne for pitch contrasts in tone language speakers and in musically untrained and trained English speakers. *Journal of Phonetics* 6: 19–23.

Westphal E. 1962. Venda: Tonal structure and intonation. *African Studies* 21: 50–69, 123–173.

Ye Y, Connine CM. 1999. Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes* 14: 609–630.

Yip M. 2002. *Tone.* Cambridge: Cambridge University Press.

Zheng HY, Tsang PW-M, Wang WS-Y. 2007. *Categorical perception of Cantonese tone in context: A cross-linguistic study.* Paper presented at the Proceedings of Intersspeech, Antwerp, Belgium.