

Strategies for analyzing tone languages¹

Alexander R. Coupe

Nanyang Technological University

This paper outlines a method of auditory and acoustic analysis for determining the tonemes of a language starting from scratch, drawing on the author's experience of recording and analyzing tone languages of north-east India. The methodology is applied to a preliminary analysis of tone in the Thang dialect of Khamniungan, a virtually undocumented language of extreme eastern Nagaland and adjacent areas of the Sagaing Division Myanmar (Burma). Following a discussion of strategies for ensuring that data appropriate for tonal analysis will be recorded, the practical demonstration begins with a description of how tone categories can be established according to their syllable type in the preliminary auditory analysis. The paper then uses this data to describe a method of acoustic analysis that ultimately permits the representation of pitch shapes as a function of absolute mean duration. The analysis of grammatical tones, floating tones and tone sandhi are exemplified with Mongsen Ao data, and a description of a perception test demonstrates how this can be used to corroborate the auditory and acoustic analysis of a tone system.

1. INTRODUCTION. This paper outlines a method of auditory and acoustic analysis for determining the tonemes of a language starting from scratch, drawing on my experience of recording and analyzing the tonal Tibeto-Burman languages spoken in the mountains of Nagaland, north-east India. The following approach is consequently oriented towards tonal analysis for the purposes of language documentation and description, rather than purely to theoretical issues concerning the abstract phonological representation of tone; however, it will also touch on aspects of theoretical analysis where this is conducive to accounting for patterns of tonal realization.

The methodology is exemplified by applying it to a preliminary analysis of tone in the Thang dialect of Khamniungan (ISO 639-3 code: kix), a virtually undocumented Bodo-Konyak-Jinghpaw language of extreme eastern Nagaland that is also reported to be

¹ The two decades of research that informs this paper has been made possible by a succession of scholarships and grants, including an Australian National University Masters Scholarship, an Australian Postgraduate (doctoral) Award, an Australian Research Council Postdoctoral Fellowship (DP0450943), and more recently, by a Nanyang Technological University School of Humanities & Social Sciences startup grant and a Singapore Ministry of Education Tier 2 research grant (MOE2012-T2-1-100). I gratefully acknowledge the generosity of all these institutions for funding my work on the languages of Southeast and South Asia. I thank two anonymous reviewers, my editors Larry Hyman and Steven Bird, and Alexis Michaud for their helpful comments and suggestions on an earlier draft, and Denis Creissels, Frantisek Kratochvil, Randy LaPolla, John Newman, Phil Rose, David Sandong and Joanna Sio Ut Siong for some informative discussions on tone languages and tonal analysis. I alone bear responsibility for any remaining errors of interpretation or analysis in this paper.

spoken in adjacent areas of the Sagaing Division of Myanmar (Burma).² The discussion is further supported by insights gained from recording and analyzing Ao (ISO 639-3 code: njo), Chang (ISO 639-3 code: nbc) and other Tibeto-Burman languages of Nagaland in which the tonal domain is the syllable. Given that these languages present a smorgasbord of features that are also found in African languages (level tone systems, floating tones, tone sandhi, polysyllabic word structure) as well as prototypical East Asian languages (potentially rich tonal inventories including contour tones, fewer phonological oppositions in codas than in onsets), they offer a hitherto relatively unexploited source of information for advancing our general understanding of the typology of the world's tone languages.

Before embarking upon the topic of this paper, it is important to mention that languages of the Tibeto-Burman (hereafter TB) family diverge significantly in the nature of their tonal profiles. The tone bearing unit is typically the syllable in TB languages spoken in mainland Southeast Asia, in common with the Sinitic, Tai-Kadai and Hmong-Mien languages of this region. In contrast, TB languages spoken on the Tibetan Plateau and in the Himalayan region demonstrate a great deal of variation in the nature of their prosodic systems. Some of these are characterized by the domain of tone being the word, e.g. Tamang (Mazaudon 1974, 2005, this volume), Kham (Watters 2003) and Manangbe (Hildebrand 2007), while for others the tone bearing unit is recognized as the syllable, as in Naxi and other Naish languages (Michaud & He 2007, Michaud 2011). Lexically contrastive tone may be restricted to just polysyllabic members of a certain word class, such as the polysyllabic nouns of Balti Tibetan and Lepcha (Sprigg 1966), or tone systems may demonstrate a complex interrelationship with phonation contrasts, rhyme durations and laryngeal coda consonants, as reported in varieties of Tibetan (Sun 2003), and in Tamang (Mazaudon, this volume). Some northern varieties of Qiang have not phonemicized tone at all, while southern varieties demonstrate systems of tonal contrasts in various stages of nascence (e.g. Evans 2001, LaPolla 2003), thus demonstrating how even closely related TB languages can have radically different prosodic systems. TB languages spoken in the mountain ranges dividing the Brahmaputra River valley from the Chindwin River valley generally have tone systems similar to the monosyllabic tonal languages of mainland Southeast Asia, except that contour tones occur less frequently in their tonal inventories than level tones. Many also diverge in demonstrating moderately synthetic and agglutinative word formation characteristics that can result in polysyllabic concatenations of two to six syllables, each bearing its own phonemic tone. The juxtaposition of multiple syllables in word formation potentially results in complex tone sandhi perturbations that can present a considerable challenge to phonological representation.

The paper has the following structure. Section 2 considers preparations for recording a word list and the steps that can be taken to ensure that the recorded corpus of words will facilitate the auditory analysis of the tone system. It also discusses in detail a number of strategies that can be adopted to prevent or minimize intonation interfering with the elicitation of citation tones, as this can present a significant obstacle to the analysis of tone in the TB languages of north-east India. Section 3 presents a demonstration of the preliminary tonal analysis of Khamniungan. In Section 4 the Khamniungan data is again used to dem-

² Khamniungan speakers report that there are more Khamniungan villages located in Myanmar than in India, but this remains to be confirmed.

onstrate how instrumental techniques can be used to arrive at a quantified phonetic model of the tonemes by plotting their fundamental frequencies as a function of absolute mean duration. The main purpose of Section 5 is to demonstrate the phonological analysis of grammatical tones, tone sandhi and floating tones, using data from the Mongsen dialect of Ao. Section 6 outlines how a perception test can be used to test the accuracy of the analysis and the ability of native speakers to recognize tonal contrasts in the absence of contextual cues, and Section 7 concludes with a summary of the main points of the paper.

2. RECORDING THE DATA FOR TONAL ANALYSIS. Recording a word list is the first important step towards analyzing the tone system of a language. This essentially involves recording the target words uttered within a predetermined carrier sentence rather than recording repetitions uttered in isolation, although the latter may also prove useful for comparison. For most field workers, determining the number and nature of the tonal contrasts is just one of many tasks undertaken in documenting and analyzing a language, so the choice of lexical items for recording the word list might take into account previous work in the language family, and possibly also future uses for the corpus. I incorporate word lists from earlier research, such as those found in Benedict 1972 and Matisoff 2003, because these publications offer reconstructed proto-forms for hundreds of Tibeto-Burman etyma. In addition to furnishing a useful list of basic vocabulary to record, they provide an excellent foundation for doing subsequent comparative research and historical reconstructions in the language of investigation.

Before making a recording I find it helpful to sort the lexical items into semantic fields and word classes to minimize misinterpretations on the part of the consultants. Confusions can easily arise if the list randomly jumps from nouns to verbs or from a kinship term to an abstract noun during elicitation, simply because the words happen to be arranged in alphabetical order. Organizing lexical items according to their word classes has the added advantage of facilitating the early recognition of recurring morphology, such as the nominalizing affix that is often attached to the citation forms of verbs in many Tibeto-Burman languages. Also, some languages may use tone to encode grammatical categories associated with a particular word class. Tone is frequently reported to mark case on nouns or tense and aspect on verbs in African languages,³ therefore such grammatical uses of tone are more readily recognized if the data to be elicited is appropriately organized. Because of the crucial grammatical role played by tone in many languages of Africa, Snider (this volume) strongly advocates organizing the data in such a way that tones are only compared within lexical categories, not across them. In contrast, Morey (this volume) notes that lexical categories are not rigidly defined in Tai languages, and that the full range of tonal contrasts occurs in Tai Phake without regard for word class. This observation is also true of many

³The use of tone for encoding inflectional and derivational categories of grammar appears to have first been reported in Christaller's grammar of Akan (Tshi), published in 1875. Pike (1948: 22–24) gives examples of the inflectional uses of tone in Mixteco and Mazateco, two Oto-Manguean languages of southwestern Mexico. He also provides references to publications on Eastern Sudanic, Niger-Congo, Chadic, Muskogean, Na-Dene, and Bantu languages that demonstrate similar grammatical functions of tone. The non-lexical use of tone generally has a much more marginal function in the languages of Asia – see Section 5 below for some rare documented examples.

other tonal languages of Southeast Asia, thus reflecting differences both in the typology of tone languages and in approaches to their analysis.

Having decided on an appropriate word list, the next step is to consult with speakers to choose a suitable substitution frame within which each word is to be uttered.⁴ Some native speakers will balk at or refuse to utter something that is semantically nonsensical,⁵ so I find it expedient to choose a sentence like ‘I write ____’ or something equivalently acceptable, e.g.

- (1) Mongsen Ao
nì *nḁ* _____ *tḁ* *sā-Ø*
 1SG AGT thus say-PST
 ‘I said ____.’⁶

A judiciously chosen carrier sentence permits the insertion of every lexical item in the word list, and the insertion point for the target word is appropriately located in the middle of the sentence rather than at the end, where a following pause can trigger the realization of a contour tone in some languages. This occurs in the Khonoma dialect of Angami (Angami-Ao cluster, Nagaland), as we discovered in a field methods class taught at Nanyang Technological University in 2010. The citation forms of monosyllabic words uttered in isolation were found to have one of four falling pitches, whereas the pitches of the same syllables were pronounced with one of four level registers when concatenated with a following syllable.

In addition, it is known that vowels have intrinsic pitch that may affect the pitch of a syllable, and that high vowels have a higher fundamental frequency (henceforth F_0) than low vowels (Lehiste 1970: 68ff.; Maddieson 1997: 622–623), therefore high vowels should not flank the target word in the carrier sentence, so as to control for the chance of their inadvertently affecting the pitch of a syllable. For similar reasons, [+high] consonants such as palatals might also be avoided in the proximity of the target word.

A battery of substitution frames with different pitches on syllables flanking the target word might be used to reveal tone sandhi patterns and possibly other information useful

⁴ Pike 1948 provides an informative overview of the use of different substitution frames for identifying tonemes and exploring tone sandhi perturbations. See Yu (this volume) for extended discussion of Pike’s methodology in the context of experimental design.

⁵ The reluctance of some native speakers to be recorded uttering nonsensical constructions potentially has consequences for the utility of some forms of elicitation methodology, such as that proposed in Hyman 2007.

⁶ Abbreviations used in glosses conform to the Leipzig Glossing Rules, except for the following: AGT agentive case marker; NZP nominalizing prefix, RL relational noun prefix, SUP superlative derivational suffix. For some of the languages discussed in this paper I represent tones with the following diacritics: \acute{o} high tone, \grave{o} low tone, \hat{o} falling tone, \check{o} rising tone and \bar{o} mid tone. For other languages I use combinations of tone numerals, e.g. σ^{55} high-level tone, σ^{51} high-falling tone etc. (after Chao 1930). Intonation patterns superimposed over the tones of individual syllables forming words, phrases and clauses are notated with \nearrow for a rising intonation and \searrow for a falling intonation.

to the description of a language. After eliciting a basic word list in the Konyak language Chang using a carrier sentence equivalent to that of (1), I explored tone sandhi patterns by recording consultants uttering a selection of verb roots marked by the declarative mood suffix in the carrier sentence ‘I VERB’, e.g. *ŋo¹¹ haʋ⁵⁵-kəɪ¹¹* (1SG.ABS GO-DECL) ‘I went’.⁷ This was useful not only for revealing that the low tone of the declarative mood suffix exerts a sandhi perturbation on the high tone of a preceding syllable, but also that the transitivity status of the verb consistently determines the case of the pronominal argument, thus confirming my suspicions that Chang has formalized an ergative-absolutive pattern of case-marking alignment (Coupe 2011).

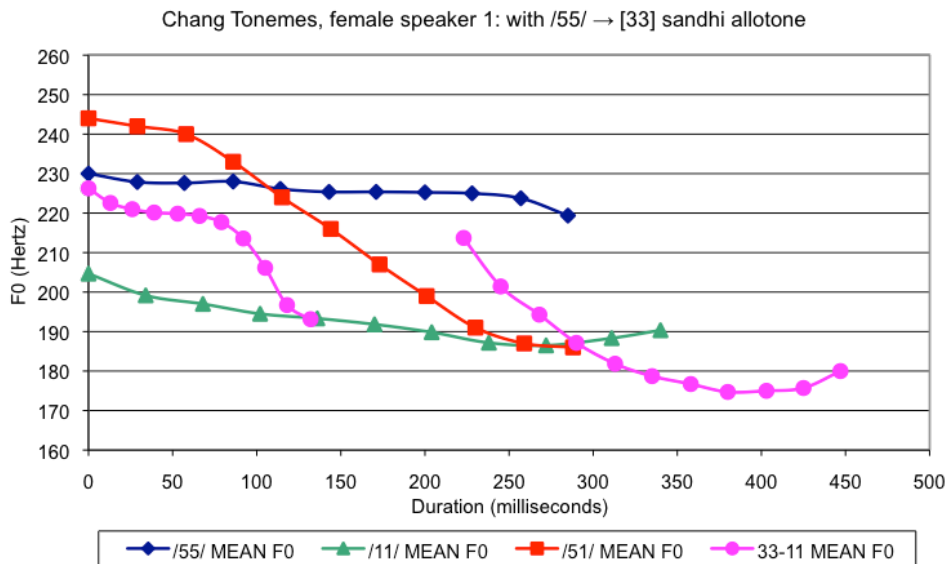


FIGURE 1. Tone sandhi effects in Chang disyllables

Chang is found to have a high /55/ toneme, a high-falling /51/ toneme, and a low /11/ toneme, as respectively represented by the blue, red and green F_0 contours of Figure 1 above. However, when eliciting the word list I also heard a mid-level pitch in many polysyllabic environments. Figure 1 additionally demonstrates how the /55/ toneme is realised phonetically with a [33] allotone in the environment of a [11] tone in disyllabic words such as /lou⁵⁵-kəɪ¹¹/ ‘come-DECL’, /thu⁵⁵-kəɪ¹¹/ ‘dig-DECL’ and /we⁵⁵-kəɪ¹¹/ ‘count-DECL’, all of which are consistently uttered with [33-11] output tones. The figure permits us to see how the high tone of the initial syllable is downstepped to a mid-level pitch in the environment of a low tone in these disyllabic words (represented schematically by the pink lines with circles).

⁷ In common with its distantly related neighbour Ao, the past tense appears to be the morphologically unmarked category in Chang (Coupe 2007, 2013). The zero morpheme of the past tense could therefore be represented by $-\emptyset$ in a morphological representation.

When eliciting basic word lists I usually request consultants to utter two elicitations of a target word in isolation first, and then have them utter the word in the carrier sentence. That gives me three opportunities to listen to the word when I am transcribing without having to replay that portion of the sound file. Furthermore, any variations in tonal realization between a word uttered in isolation and the same word uttered in the carrier sentence should immediately be revealed by a comparison of the tokens heard in succession.

The use of carrier sentences serves two important advantages over having a consultant utter words in isolation. Firstly, they provide a yardstick against which an elicited word's tone(s) can be determined, particularly in a language that only has terrace-like level tones.⁸ To illustrate, the agentive case marker *nə* in (1) above is consistently uttered with a mid tone in Mongsen Ao, so it can be used to gauge whether a following syllable is higher, lower, or equivalent in pitch. If every elicited word is uttered in a carrier sentence, it eventually becomes considerably easier to hear the tonal contrasts in the language being studied, even if the investigator is initially not sure what the pitches of all the syllables in the carrier sentence might be.

Secondly, a carrier sentence can help to minimize the effects of “list” intonation when, for example, three tokens of a word are uttered in succession, as is standardly used for recording word lists, e.g.

- (2) Mongsen Ao: three elicitations of ‘back’ demonstrating list intonation
 [təʃɪn ↗ təʃɪn ↗ təʃɪn ↘] (1)

While three repetitions of a word may be sufficient for working on the phonology of a non-tonal language, such a methodology can run into problems in tone languages that also make use of intonation. Under the effects of list intonation, the first two tokens of (2) were pronounced with a rising intonation by a Mongsen Ao speaker, while the final token was pronounced with a falling intonation (respectively represented by the notation ↗ and ↘). This word actually has underlying mid-mid phonemic tones in isolation.

The influence of intonation on the realization of lexical tones appears to be especially salient in the TB languages of north-east India (e.g. cf. Morey's Singpho data in this volume), and is something that must be controlled for if one is to elicit a valid corpus for tonal analysis. In 1996 when I first began working on Mongsen Ao I was initially unsure whether the language had contrastive tone or not, due to the very strong effects of intonation on the individual tones of syllables. And if tone was indeed contrastive in the language, then it remained to be determined if the pitch shapes were of the contour type or the level register type, as these were concealed by the overarching intonation patterns. Subsequent research eventually revealed that this language has three register-like contrastive tones, but the tones of syllables are often overridden by the imposition of intonation patterns that serve to delineate phrasal and clausal boundaries in running speech (Coupe 2007: 73–75; Coupe 2012). A clause-final rising intonation that trumps the individual syllables of words

⁸ It can be significantly easier to identify tonal contrasts in languages with contour tones, because the pitch shape serves to provide an additional cue to determining a syllable's tone. In contrast, the tone of a syllable in a level-tone language can often only be confidently determined by comparison to another tone in the initial stages of analysis.

is commonly found to signal non-finite converb clauses in clause chaining, and this intonation pattern is also frequently used in other types of clauses – even finite main clauses – to indicate topic continuity.

To illustrate, example (3) below presents a Mongsen Ao speaker's self-introduction in phonemic transcription, and the corresponding pitch trace of Figure 2 demonstrates the dramatic impact that intonation can have on the lexical tones of individual syllables in these clauses. Observe that the high rising intonation occurs at the edges of phrasal and clausal boundaries preceding the final word of the speaker's introduction, then on the last word the intonation steeply falls to indicate that the speaker is about to introduce a new topic. Thus, while tone is phonemic in this language, a superimposed high rising intonation is applied independently of the lexical tone system to indicate topic continuity and to delineate phrasal or clausal boundaries, or to signal the non-final status of those constituents. It goes without saying that it is essential to tease apart these two dimensions of prosody when doing tonal analysis, and this is best done at the start when recording the word list.

(3) Mongsen Ao: self-introduction in narrative text

nì *ĩmtʃə̀n* *tʃà-ɿ*↗ *mā̀ŋmə̀tũ̀ŋ* *nĩ̀ŋ-ə̀ɿ*↗
 1SG clan.name offspring-NMLZ village.name name-NMLZ

tə̀-ə̀ɿ *nì* *tə̀-nĩ̀ŋ*↗ *ĩmtĩ* *lũ̀ìn*↘
 thus-CVB 1SG RL-name personal.name

'I am a descendent of the Imchen phratry, a Mangmetong villager, and my name is Imti Loin.'

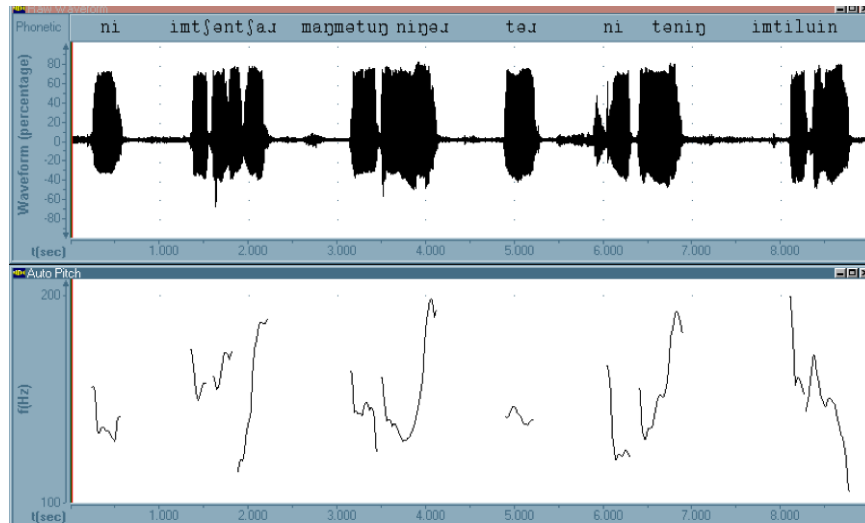


FIGURE 2. The impact of intonation on the lexical level tones of Mongsen Ao⁹ 🔊

⁹ See Coupe, Alexander R. (2012) 'Referent tracking strategies in Mongsen Ao and Chang discourse' for additional sound files and transcriptions demonstrating the impact of intonation on lexical tones in Chang and Mongsen Ao, available online at <http://jseals.org/seals22/seals22.htm>.

When attempting to elicit citation tones, one strategy for making speakers aware that they are unconsciously using list intonation is to play back a recording so that they can listen to their own pronunciations. They can then be asked if a word such as ‘back’ should be pronounced in isolation as [tə¹¹ʃin¹³], or as [tə⁵⁵ʃin³¹], or as something different. As it turned out, on replay the tokens of (2) above were rejected as not being representative of ‘back’ if the citation form of this noun is not uttered with level [33] tones on both syllables in isolation. Once speakers become cognizant of the fact that their intonation patterns are affecting their pronunciation of citation tones on lexical items, they are in a better position to control or minimize its effect. Without this awareness, the investigator will be unable to record representative data that will permit an accurate analysis of the tone system.¹⁰

Despite one’s best efforts to control for it, sometimes the entire carrier sentence falls victim to the list intonation effect, as I found with one speaker of the Thang dialect of Khamniungan. At the beginning of recording the word list, the speaker uttered the carrier sentence with the typical non-final rising intonation and then respectively produced rising and falling intonations on the second and third tokens of *khə³¹tha³³* ‘above’, a disyllabic word with underlying phonemic mid-falling and mid-level tones. This is demonstrated in example (4) below: (a) provides a phonetic transcription of the speaker’s intonation patterns and their concomitant influence on the lexical tones of individual syllables, and (b) provides a phonemic representation of the same sentence showing the underlying tonemes of those syllables.

- (4) Thang dialect of Khamniungan [Bodo-Konyak-Jinghpaw, eastern Nagaland]
- a. [ŋə: ↗ k^hət^haʔ tə ɲətʃe ↗ k^hət^ha ↗ k^hət^haʔ ↘]
- b. ɲə³¹-əʔ³⁵ khə³¹tha³³ tə³³ ɲu³³-tʃe³³ khə³¹tha³³ khə³¹tha³³
- 1SG-AGT above thus say-PRES above above
- ‘I say “above.”’ ◀▶

This raises an important question to ponder: should one attempt to influence a speaker to minimize the effects of intonation on citation tones? If it is the case that the particular intonation patterns exemplified by examples (2) and (4) are indeed an artefact of eliciting repetitions, then minimizing its effect seems to be the only way of ensuring that valid data is recorded. The analysis of lexical tone is more easily accomplished if we treat tone and intonation as belonging to separate domains of prosody that happen to interact through the common medium of F_0 . Tone is intrinsic to the phonological representation of the syllable in Mongsen Ao and Khamniungan, whereas the domain of intonation may be a single word, a phrase, or a whole clause that subsumes all these constituents, as we observe in Figure 2 and the accompanying transcription of example (3). It seems fundamentally impossible to tease out the lexical tones of the language if the task is obfuscated by not


¹⁰ Alexis Michaud (personal communication, 3/11/2013) suggests that uttering a citation form in isolation might be a strange thing for speakers to do, and for that reason the use of a contextually relevant carrier sentence is a more appropriate way to investigate tone. While I essentially agree with this, the fact that speakers readily identify which are the correct tones of a word uttered in isolation suggests that this method of elicitation still has a place in tonal analysis.

distinguishing between these lexical and postlexical prosodic domains. Intonation extending over chunks of sentential constituents is certainly something that is deserving of study, especially when it is found to interact with tone in tone languages, but logically it must follow the analysis of tone lexically assigned to individual syllables or words, whatever the tonal profile of the language may be.

The strategy for dealing with the problem of superimposed intonation in carrier sentences is the same as that proposed above for the elicitation of citation tones in individually uttered words. One can only attempt to make the speaker aware of what they are unconsciously doing by playing back the recording in the hope that its recognition will result in elicitations that are representative of the pitch contrasts of the language, rather than the superimposed list intonation patterns.

Despite one having the best intentions to control for external influences on the production of pitch contrasts, other factors may conspire to complicate the gathering of reliable data for tonal analysis. For example, the speaker may shift their pitch range up by one or two semitones,¹¹ with the result that the investigator inadvertently hears a word uttered with a [33] tone as having a [44] tone, thus representing what is interpreted as an additional pitch level. This is a third reason why using a carrier sentence is important, because pitch in a language's tone system is always relative and never absolute (Pike 1948: vi, 20). The effect of random upward or downward shifts in the pitch range of a speaker will always be relative to the pitches of words used in the carrier sentence. The relativity of pitch in tone languages is also why an auditory analysis necessarily precedes the instrumental analysis, as a pitch extraction algorithm can only provide an absolute value.

In addition to recording a substantial word list, I find it helpful to request from consultants minimal sets of words that differ only by pitch. My experience is that speakers of unwritten tonal languages are often unaware that their languages have lexically contrastive tone, therefore a demonstration of how tone is used to distinguish meaning in another language can be illuminating. I rely on minimally-contrastive sets of words in Thai or some other language I know to demonstrate the use of lexically contrastive pitch, e.g.

(5) a. Mongsen Ao	b. Chang	c. Standard Thai
<i>ní</i> QPTCL	<i>són</i> 'breast'	<i>khāa</i> 'be stuck, lodged in'
<i>nī</i> 'day'	<i>sən</i> 'hail'	<i>khàa</i> 'galangal'
<i>nì</i> 'I' (1SG)	<i>sən</i> 'we' (1PL.INCL)	<i>khāa</i> 'kill'
		<i>kháa</i> 'trade'
		<i>khāa</i> 'leg'

¹¹ In the Western music tradition, a semitone is the smallest interval between two adjacent notes in a 12-tone scale and can be equated to one fret on the neck of a guitar.

A demonstration of lexical tone contrasts can prompt a consultant to think of some tonal contrasts in their own language before the investigator even gets around to comparing the data of word lists. If consultants have trouble thinking of such sets of minimally contrastive words off-the-cuff, then another method is to ask for words formed from simple syllables, such as *pa*, *ta*, *ka* etc. to see if there are any with identical segmental phonology but different suprasegmental phonology and meanings. By the time a substantial word list has been recorded, the investigator hopefully will have encountered other minimal pairs or triplets of words differing only in tone, and these might be further expanded upon with the help of a native speaker's knowledge.

One should not automatically assume that a previous analysis done on a language called X necessarily represents the same variety of language X. Gedney (1972: 193) writes of students of Lao coming up with different tonal analyses depending upon which side of the capital Vientiane their consultants came from. My personal experience of doing linguistic research in north-east India is that every village that ostensibly speaks the same uncodified language will have subtle to substantial differences in phonology, morphology and perhaps also tone systems, so it is possible that the particular variety of a language investigated will differ in some respects from those described by previous investigators. Remarkably, even in the same village there can be significant differences between the varieties spoken in different wards.

There is also the chance that previous researchers may have got it wrong. Weidert (1987) claimed that Khamniungan is the most tonally complex language of north-east India because he analyzed it as having six contrastive tones, including two rising tones, which is an extremely rare contour type in the Tibeto-Burman languages of north-east India; even one rising tone is a rarity in the languages of this region.¹² If not for the strangely coincidental fact that Weidert and I both worked with the same consultant with an intervening hiatus of approximately thirty years, my finding just five contrastive tones could perhaps be explained away by varietal differences. Another possibility is that our mutual consultant lost a tonal contrast in the intervening years, but that is unlikely.

Lastly, it may prove to be extremely helpful to video speakers when recording a word list. A language with challenging phonology can be much more easily transcribed when one no longer has access to native speakers if it is still possible to observe the speaker as they articulate their speech sounds in a video clip. In an ideal situation all the transcriptions will be checked and confirmed while one is in the field; however, in reality an accurate phonological analysis usually involves a great deal of checking and re-checking of the data, and errors will be frequently encountered in the preliminary attempts at analysis. When one is ensconced back in the office and belatedly realises that something is amiss in the analysis, it is comforting to have a video of the speaker to check in addition to the audio recording. Furthermore, a video recording provides an emergency backup of the data, and depending upon the compression algorithm used by the video camera, it may be possible to extract lossless sound files from the video if one's digital audio recorder malfunctions in

¹² Of the eight Tibeto-Burman languages of Nagaland on which I have personally done fieldwork (Mongsen Ao, Chungli Ao, Chang, Khamniungan, Lotha, Yimchungrü, Sangtam and Khonoma Angami), Khamniungan is the only language found to have a rising contour tone amongst its lexical tone contrasts.

the field. Naturally to ensure a good audio recording, it is essential to connect an external microphone to the video camera.

3. AUDITORY ANALYSIS. Once a substantial word list has been recorded, the data can be transcribed and each syllable impressionistically assigned a pitch value. From my personal experience of researching Tibeto-Burman and Tai-Kadai languages, at this preliminary stage it is best to work with Chao numerals or letters (Chao 1930), as these permit a fairly fine-grained representation of the phonetic pitch values of syllables.¹³ Later one may decide that two pitches transcribed with separate values are not contrastive and should be collapsed into one tone category, but it is better to err on the side of caution at first and transcribe pitch narrowly, just in case a pitch level that turns out to be contrastive is missed through employing too broad a transcription. If possible it is easiest to work first with monosyllables, if one is fortunate enough to find a sufficient number in the data. Once the monosyllables have been dealt with, the pitches of various other combinations of polysyllabic words can be investigated. Disyllables may involve tone sandhi perturbations, which justifies delaying their analysis until the tonal contrasts on monosyllables have been determined.

Some tones may be limited to particular syllable types; for example, Tai-Kadai languages belonging to the south-western branch of the Tai family have fewer contrasts on short syllables with a stop coda than on unchecked syllables – typically no more than two or three – as demonstrated by the data of Brown 1985.¹⁴ This possibility warrants sorting the transcribed words into syllable types according to their rhyme duration and other relevant phonotactic characteristics, so as to facilitate the recognition of restricted distributions and to control for phonotactic influences on the realization of pitch.

Laryngeal oppositions in syllable-final consonants and the nature of a syllable's initial consonant (e.g. whether voiced or voiceless) are known to correlate with the evolution of phonemicized tone contrasts in many Southeast Asian languages. Specifically, tones are believed to have evolved via the transphonologization of distinctive features associated with consonants (Haudricourt 1954, 1972; Gedney 1972; Mazaudon 1977, 2012, this volume), so the vestiges of tonogenesis may possibly be demonstrated by different syllable types in the language under investigation. Laryngeal gestures can also be associated with particular tones, such as the creaky tone of Burmese (Thurgood 1981, 2002). A glottalized termination is often manifested in reflexes of the third historical tone of Tai languages (i.e. Proto-Tone C), as found in Shan, Central Thai and many other Tai dialects. Watters (2003:

¹³ Instead of employing superscript integers 1 to 5, an alternative is to use a graphic representation of a pitch level or contour, especially when the transcription of multiple diacritics results in messy tiers of stacked glyphs. Some IPA fonts, such as SIL Doulos, permit pitch to be represented by tone letters corresponding to levels or transitions between levels, e.g. ɿ for a high level [55] tone and † for a mid level [33] tone etc.

¹⁴ In the Tai descriptive tradition, syllables with an oral or glottal stop coda are referred to as 'checked' or 'dead', whereas syllables terminating in sonorant consonants or vowels are termed 'smooth', 'free' or 'live' (e.g. Gedney 1972: 193). These two types are respectively referred to as คำตาย *khām tāaj* 'dead word' and คำเป็น *khām pēn* 'live word' in Standard Thai and are the most significant determinant of the types of tonal contrasts that a syllable can bear.

687) notes the influence that voice register has on the tone system of Kham (Kiranti branch of TB, Nepal) and observes that breathiness and laryngeal laxity are associated with lower frequencies in comparison with the modal register. In view of this, the possible influence of phonation type might additionally be taken into consideration when analysing and sorting the data. On the other hand, phonation types and tonal contrasts may exist independently in a language. Monsen Ao has a phonemic creaky voice ~ modal voice contrast on the low central vowel /a/ that does not appear to interact with tone at all, as creaky voice co-occurs with both low and mid tones, and possibly also with the high tone (Coupe 2003: 43–45).

An approach that recommends methodically organizing the data to facilitate tonal comparisons is repeatedly echoed in many of the papers in this volume (e.g. Cruz & Woodbury, Mazaudon, Rice, Snider), and for good reason – it is essential to ensure that the data is analogous to minimize the number of variables that can potentially trigger a particular tonal realization.

In Coupe 2003 I isolated checked monosyllables, unchecked monosyllables, and a third group of monosyllables ending in nasals in preparation for analyzing the tone system of Mongsen Ao. But with the benefit of hindsight it may be unnecessary to distinguish a nasal coda syllable type, save for excluding syllables with velar nasals. Rose 1992 found that a syllable-final velar nasal coda caused significant differences in both the duration of a tone and the F_0 height of its syllable in the Zhenhai dialect of Chinese, so syllables with velar nasal codas should at first be quarantined from other types of unchecked syllables until the effect of a velar nasal coda can be determined. If the language has a phonemic contrast in vowel length or has diphthongs, then it is often necessary to further sort monosyllables according to this criterion as well. Thai has particular tonal contrasts in checked syllables that are conditioned by vowel length, and the vast majority of falling contour tones of Chang occur on syllables with long vowels or diphthongs. Presumably there is a diachronic reason for this in Chang, the most likely one being that the high and low tones of adjacent syllables historically merged to form a single syllable with a high falling contour.

Having sorted the monosyllables into categories according to their relevant phonotactic structures, the next step is to further separate each group according to its pitch characteristics. In the preliminary auditory analysis of the Thang dialect of Khamniungan I initially recognized six distinct pitches occurring on unchecked and checked monosyllables. The identified tone categories were:

1. high level tone [55]
2. high falling tone [51]
3. mid level tone [33]
4. mid falling [31]
5. low falling [21]
6. mid rising [35]

The mid falling [31] and low falling [21] tones were reviewed and later collapsed into one tone category. It was suspicious that they shared an almost identical contour, differing only in the pitch height at the start of their trajectories, and subsequent investigations with native speakers (discussed below) suggested that my preliminary transcription was too narrow and thus unlikely to reflect a phonemic difference. The differences perceived

between monosyllables with [31] and [21] pitches may have been attributable to the speaker slightly adjusting his pitch range during the elicitation session, or possibly could have been caused by the intrinsic pitch of vowels or consonants exerting a discernible effect. Such miscalculations are likely to occur in the initial stages of analysis when the investigator is struggling to attune their perception to possible pitch contrasts. Speakers will not necessarily produce exactly the same pitch on every word belonging to one and the same tone category, but it takes some time to familiarize one's ear as to how much leeway should be given to the range of each pitch in the system of contrasts. In some respects, tonological analysis compares with the phonological analysis of a vowel system vis-à-vis determining the extent of a particular phoneme's vowel space. The same might be said for determining a particular toneme's pitch parameters.

Examples of tone categories found on unchecked and checked syllables in the Thang dialect of Khamniungan are respectively presented in Table 1 and Table 2 below.

[55] tones	[33] tones	[31] tones
sou ⁵⁵ 'shadow'	lou ³³ 'ashes'	pau ³¹ 'grandfather'
ho ⁵⁵ 'tooth'	ou ³³ 'axe'	tʃi ³¹ 'grandmother'
tʃhi ⁵⁵ 'work'	war ³³ 'bamboo tie'	khun ³¹ 'dream'
	pem ³³ 'body'	mia ³¹ 'ant'
[51] tones	hon ³³ 'crab'	ʃɛ ³¹ 'blood'
ŋou ⁵¹ 'honey bee'	kou ³³ 'door'	u ³¹ 'bone'
ʃru ⁵¹ 'stinking bug'	le ³³ 'in-law'	tʃhiu ³¹ 'plate'
ar ⁵¹ 'cane'	ʃi ³³ 'shield'	kha ³¹ 'face'
ar ⁵¹ 'garden'	pau ³³ 'snake'	khay ³¹ 'head'
mia ⁵¹ 'corpse'	ju ³³ 'word'	
[35] tones		
tʃi ³⁵ 'dog'		
sa ³⁵ 'dao' (hatchet)		
ki ³⁵ 'exist'		
kay ³⁵ 'father-in-law'		
iy ³⁵ 'oil'		
ʃi ³⁵ 'shout'		
lai ³⁵ 'spear'		

TABLE 1. Khamniungan tone categories on unchecked monosyllables

[55] tones tsak ⁵⁵ ‘shadow’ ʃap ⁵⁵ ‘bear’	[33] tones tʃhiŋ ³³ ‘food’ hok ³³ ‘leg’ phaŋ ³³ ‘hand’ kop ³³ ‘jaw’ tshap ³³ ‘hoe’	[31] tones ɲauŋ ³¹ ‘breast’ ɲeɪŋ ³¹ ‘flesh’ ŋouŋ ³¹ ‘fish’ hru ³¹ ‘louse’ sauŋ ³¹ ‘rat’ iŋ ³¹ ‘breath’
[51] tones luk ⁵¹ ‘six’		
[35] tones ʃeŋ ³⁵ ‘poison’ tsau ³⁵ ‘thorn’ tʃeɪ ³⁵ ‘urine’ ʃiŋ ³⁵ ‘dung’		

TABLE 2. Khamniungan tone categories on checked monosyllables

Ideally, syllables such as [tʃhi⁵⁵] ‘work’, [ʃi³³] ‘shield’ and [iɣ³⁵] ‘oil’ etc. in Table 1 should be omitted from any subsequent acoustic analysis, as these syllables’ palatal onsets and/or high vowels may skew the results when their F_0 values are measured and plotted as a function of time in the subsequent acoustic analysis (see Section 4 for further discussion). Table 2 below demonstrates the paucity of checked syllables belonging to the [51] tone category, and how sometimes the analysis might be limited by a lack of all syllable types, even in a corpus of circa 700 lexical items. As in segmental phonemic analysis, rarely occurring examples need to be carefully checked to make sure that they are not actually allotonic realizations of some toneme.

Having decided on how many tone categories are represented in the preliminary analysis of the transcribed data, the recorded utterances of each monosyllabic word can then be combined in a single sound file according to tone category; this permits their pitches to be compared to see if all members respectively conform to the various tone categories that have been established.¹⁵ Alternatively, the investigator can request a native speaker to pronounce the words in each tone category, and any that are found to diverge from the common pitch contour or height can be reassigned to their appropriate tonal category. It was at this point that we decided that the difference between [31] and [21] open syllables was insufficient for the two tone groups to be considered contrastive. These were combined into one tone category of [31] open syllables, as that appeared to be the most representative contour for the category. Once the odd members of each group have been reassigned by trial and error, the process can be repeated until the investigator and consultants are satisfied with the sorting and membership of each tone category. Consultants may also be able to offer additional examples to ensure that there is a sufficient number of tokens for each tone category within each syllable type.

¹⁵ Naturally this assumes that there has not been major upward or downward shifts in the pitch range of the speaker during the recording session, otherwise this will need to be taken into account when listening to elicited words and deciding on their tone category membership.

If other members of the speech community are available, now is a good time to record more words belonging to each of the posited tone categories in preparation for the instrumental analysis. Opinions as to what constitutes a sufficient number of speakers on which to base a tonal analysis may well vary from investigator to investigator, but I assume that a minimum of six to twelve words should provide a representative sample of each tone category of the different types of monosyllables, and the combined data of six to twelve speakers should provide a substantial and representative corpus for the next step of acoustic analysis.

4. ACOUSTIC ANALYSIS. This section describes a methodology for plotting in a line chart each tone's F_0 contour as a function of absolute mean duration. It requires the use of instrumental analysis software such as Praat (Boersma & Weenink 2013) to facilitate the creation of spectrograms from which F_0 can be calculated.¹⁶

The domain that constitutes the tone bearing unit (hereafter TBU) – whether it is the vowel, the mora, the entire syllable, or just the rhyme – has been much discussed in the literature. Pike's original notion of a tone language was one 'having lexically significant contrastive, but relative pitch in each syllable (1948:1), so clearly for him the syllable constituted the TBU. Goldsmith (1990: 45) takes a more restricted view and asserts that the vowel being the TBU is a *fact* about tonal systems, and independent of autosegmental approaches to their analysis. For Chao (1968: 19), the domain of tone is the voiced part of the syllable, thus it is assumed to include an onset if it is voiced. If this is valid, then it is likely to be limited to sonorants. This is due to physiological limitations on the control of F_0 during the articulation of oral stops (Michaud & Kühnert 2006). A possible situation in which a sonorant onset might contribute a perceptually relevant cue to the perception of pitch is when one occurs medially in a disyllable. This is suggested by the F_0 height of the medial nasal at its transition from the preceding syllable's rhyme in *tə-māŋ* 'PROH-believe' in Figure 4, and is discussed in further detail in section 6.

In an important experimental study, Howie (1974) demonstrated that the tones of Mandarin are coextensive only with the vowel and following voiced segments, and that the onset carries tonally irrelevant, anticipatory adjustments, even if it consists of a voiced segment. In consideration of this finding, it is recommended that F_0 only be calculated over the duration of the tone bearing unit of the syllable, i.e. the rhyme.

To measure a F_0 contour and its duration, I first make a broadband spectrogram of the syllable using the standard range of 0–5000Hz. A broadband spectrogram resolves the formants, which can be used to help determine precisely the beginning and termination of the rhyme. The wave form may also provide a useful cue for deciding the point at which the rhyme begins and ends, depending on the type of initial consonant if the syllable onset is filled. I then expand the rhyme to the full width of the monitor's window so that I have consistent reference points for sampling F_0 over the duration of the rhyme of this and all other syllables to be analyzed. The Praat manual recommends that the pitch range setting be adjusted according to whether the speaker is a male, female or a child. The default range is 75Hz to 500Hz, but this can be set so that it is just below the floor and just above the ceil-

¹⁶ e.g. see <http://www.fon.hum.uva.nl/praat/>, and <http://savethevowels.org/praat/> for a useful introduction by Will Styler titled *Using Praat for linguistic research*.

ing of the speaker's pitch range to improve the accuracy of the F_0 calculation. I then sample F_0 at 10% increments of the rhyme if there are contour tones, starting at 0% and ending at 100%. For languages with level tones, the sampling frequency can be more parsimonious. I also note the duration of the rhyme, and all data is entered into a spreadsheet to facilitate the analysis and eventual graphic representation. The process is repeated for each monosyllable of each tonal category, and their mean F_0 at each incremental sampling point, their standard deviations and overall mean durations are calculated.

Sometimes the automatic pitch extraction algorithm of software applications gives spurious results, or is unable to extract an F_0 measurement for some parts of the rhyme. This is particularly likely to happen when there is creaky voice phonation in the signal or some degree of glottalization immediately preceding a glottal stop closure. Styler (2013: 15–17) describes a number of settings that can be tweaked in Praat to improve the accuracy of the pitch tracking function, but ends with the caveat that 'relying on it to give you sane measures is not wise, especially in scripts'. When I find that the pitch tracking function of Praat gives a dubious calculation of F_0 over part of the rhyme because of creaky voice or some limitation with the software, I instead make a narrow band spectrogram of the rhyme with a range of less than 1000Hz. This provides a decent resolution of the first seven or eight harmonics. Calculating the F_0 of the pitch at any point in the rhyme is then a simple matter of measuring the F_0 of the n^{th} harmonic divided by n (Baken 1987: 140). This manual method is somewhat tedious, but ensures that accurate and reliable measurements can be made of all the data at hand.

The acoustic measurements of the Khamniungan data permitted me to once more review the assignment of each monosyllable to a tonal category. Any monosyllables whose F_0 was significantly aberrant was checked with a native speaker and reassigned if its pitch was found to diverge from the parameters of a posited tone category. The minimal differences I found between the normalized [31] and [21] pitch shapes provided yet further evidence corroborating our auditory perceptions that these two pitches actually constituted a single tone category, and justified collapsing the two populations into one. This single category is represented by the [31] pitch shape of Figure 3 below.

The following tables present calculations of F_0 and rhyme durations of each tone category for one male speaker of the Thang dialect of Khamniungan, based on acoustic measurements of the data presented above in Table 1. For the purposes of this demonstration, monosyllables with high vowels and high consonants have been included, but with more data it would be preferable to exclude such tokens to minimize their potential impact on the results.

[55] F_0	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
sou ⁵⁵ 'shadow'	163	152	150	147	145	143	141	138	137	135	128
ho ⁵⁵ 'tooth'	122	123	124	123	121	115	110	107	105	102	99
tʰi ⁵⁵ 'work'	139	133	127	125	123	121	120	118	113	104	97
[55] MEAN F_0	141	136	134	132	130	126	124	121	118	114	108
[55] STDEV	20.6	14.7	14.2	13.3	13.3	14.7	15.8	15.7	16.7	18.5	17.3
[55] MEAN DUR	0	0.027	0.053	0.08	0.107	0.134	0.16	0.187	0.214	0.24	0.267

TABLE 3. F_0 values (in Hz.) and rhyme durations (in ms.) for unchecked [55] monosyllables

[51] F ₀	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
ŋou ⁵¹ 'honey bee'	156	152	138	124	112	99	91	84	80	70	68
ʃru ⁵¹ 'stinking bug'	153	149	143	130	111	94	83	75	75	71	76
ai ⁵¹ 'cane'	131	133	136	132	125	114	106	96	89	76	68
ai ⁵¹ 'garden'	162	138	132	127	119	110	99	88	80	74	74
mia ⁵¹ 'corpse'	140	140	136	129	121	112	103	94	83	76	65
lei ⁵¹ 'tongue'	148	148	143	132	120	109	99	88	79	72	66
51 MEAN F₀	148	143	138	129	118	106	97	88	81	73	70
[51] STDEV	11.3	7.4	4.3	3.1	5.4	8.0	8.4	7.5	4.7	2.6	4.5
51 MEAN DUR	0	0.031	0.063	0.094	0.125	0.157	0.188	0.219	0.251	0.282	0.313

TABLE 4. F₀ values (in Hz.) and rhyme durations (in ms.) for unchecked [51] monosyllables

[35] F ₀	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
tʃi ³⁵ 'dog'	117	115	118	122	131	135	137	141	141	131	123
sa ³⁵ 'dao' (hatchet)	110	109	113	116	120	123	127	127	127	116	107
ki ³⁵ 'exist'	124	116	120	123	128	132	139	140	136	134	132
kay ³⁵ 'father-in-law'	116	121	120	121	124	127	130	129	129	128	120
iɣ ³⁵ 'oil'	121	119	121	127	135	142	148	141	135	129	110
ʃi ³⁵ 'shout'	128	129	129	127	132	136	140	146	148	127	112
lar ³⁵ 'spear'	103	104	112	117	118	118	123	125	123	111	103
[35] MEAN F₀	117	116	119	122	127	130	135	136	134	125	115
[35] STDEV	7.9	7.5	5.2	4.0	5.9	7.7	8.0	7.7	8.0	7.8	9.4
[35] MEAN DUR	0	0.034	0.069	0.103	0.138	0.172	0.207	0.241	0.276	0.31	0.345

TABLE 5. F₀ values (in Hz.) and rhyme durations (in ms.) for unchecked [35] monosyllables

[33] F ₀	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
lou ³³ 'ashes'	134	129	128	125	122	118	116	115	113	112	105
ou ³³ 'axe'	129	119	119	118	117	117	116	113	114	108	104
wai ³³ 'bamboo tie'	118	122	119	117	113	109	110	109	107	104	94
pem ³³ 'body'	130	125	122	120	119	116	113	113	111	111	105
hon ³³ 'crab'	123	124	121	118	115	114	115	114	112	104	91
kou ³³ 'door'	130	119	114	112	111	111	112	112	110	104	99
le ³³ 'in-law'	127	119	116	115	114	112	112	111	111	108	105
ʃi ³³ 'shield'	124	123	121	117	115	114	113	114	115	112	105
[33] MEAN F₀	127	123	120	118	116	114	113	113	112	108	101
[33] STDEV	5.0	3.5	4.2	3.8	3.5	3.1	2.1	1.9	2.5	3.6	5.7
[33] MEAN DUR	0	0.032	0.065	0.097	0.129	0.161	0.194	0.226	0.258	0.291	0.323

TABLE 6. F₀ values (in Hz.) and rhyme durations (in ms.) for unchecked [33] monosyllables

[31] F ₀	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
pau ³¹ ‘grandfather’	122	116	110	105	99	93	90	86	80	73	67
tʃi ³¹ ‘grandmother’	119	115	111	106	98	89	82	78	76	72	65
khun ³¹ ‘dream’	119	116	112	103	96	86	79	75	71	69	64
mia ³¹ ‘ant’	114	112	108	103	98	97	93	89	84	78	76
ʃe ³¹ ‘blood’	123	112	107	101	95	92	87	81	76	73	73
u ³¹ ‘bone’	119	118	111	104	97	90	84	87	88	86	84
tʃhiu ³¹ ‘plate’	115	109	105	99	95	91	87	81	78	75	73
kha ³¹ ‘face’	110	113	112	107	101	94	89	86	83	78	72
khay ³¹ ‘head’	109	115	114	108	101	96	91	85	80	74	68
[31] MEAN F₀	117	114	110	104	98	92	87	83	80	75	71
[31] STDEV	5.0	2.7	2.8	2.9	2.3	3.5	4.5	4.6	5.1	4.9	6.2
[31] MEAN DUR	0	0.025	0.05	0.076	0.101	0.126	0.151	0.176	0.202	0.227	0.252

TABLE 7. F₀ values (in Hz.) and rhyme durations (in ms.) for unchecked [31] monosyllables

Once all the values have been calculated, the mean F₀ and mean rhyme durations can be plotted to display the pitch contours as a quantified phonetic model of the language’s tonemes. Because the primary purpose of this article is to demonstrate a methodology for tonal analysis, and because the analysis and documentation of Khamniungan is still a work in progress, the data provided in the tables above represents that of just one male speaker and only looks at one syllable type, viz. unchecked monosyllables. As mentioned above, the process is ideally repeated with the data of six to twelve speakers and the results for each syllable type combined to arrive at a truly representative picture of the tonal contrasts in the language of investigation. Due to the significant differences in F₀ between females and males, the data of male and female speakers should be calculated separately.

After the tonemes of monosyllables have been identified, the process can be applied in an identical manner to an acoustic analysis of the rhymes of checked syllables, and then to the rhymes of disyllables.

Figure 3 (below) plots the data presented above in Tables 3–7. Observe that the [55] and [33] tones fall quite significantly after 150ms, yet these falls are not perceptible, so their pitches sound quite level. This may be attributable to their F₀ trajectories at this duration of the rhyme being within the glissando threshold, which determines how much change in F₀ is required for a pitch to be heard as having a contour pattern (for further discussion see Remijsen, this volume, and papers cited therein).

The most unexpected result in this data is the pitch shape of the [35] tone. Even though the F₀ contour demonstrates that its pitch begins to fall at around 250ms after rising from a mean F₀ of 120Hz to nearly 150Hz, the auditory impression is that it rises throughout its entire duration.¹⁷ There may be two interrelated explanations for this. This first is that

¹⁷This mid rising-falling pitch shape is attested in the acoustic data of another male speaker of the Thang dialect of Khamniungan, so I am confident that the pattern is representative of the /35/ toneme of this language.

pitch variations are known to occur due to changes in airflow as well as adjustments of the laryngeal muscles (Ladefoged 2003: 86–88). It is expected that airflow diminishes at the end of an utterance, so this is likely to be accompanied by a fall in F_0 due to aerodynamic reasons. The second reason may be that the most salient and therefore most perceptible part of a tone's contour in Khamniungan is the first 200ms, and the [35] contour has already hit its high target at this duration of the rhyme. In fact, if we consider just the first 200ms, then all of the tonemes of Khamniungan contrast considerably in both contour and pitch height over this duration, despite the fact that they all decline to some extent over their durations, again, probably due to aerodynamic reasons. The only partial exception to this is the [35] tone, which does not fall until the last 75ms of its contour.

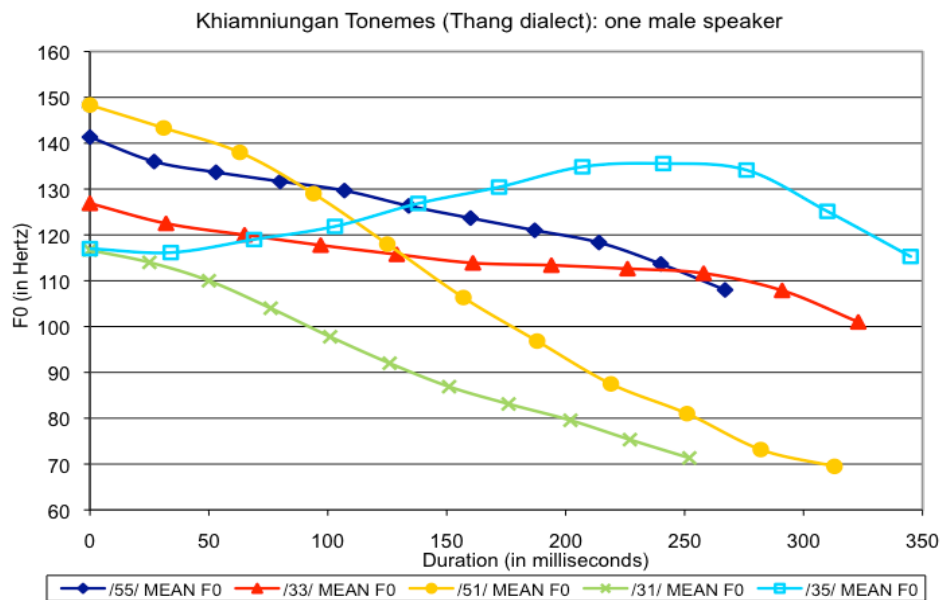


FIGURE 3. Mean /55/, /51/, /35/, /33/ and /31/ shapes for one male speaker of Khamniungan (Thang dialect).

I tested the ability of linguists and linguistics students in my department who are either native speakers of various tone languages or who have done research on tone languages (specifically, Mandarin, Cantonese and Rawang) to listen to and describe the pitch shape of the [35] tone, the examples of which are listed above in Table 1. All my consultants heard this as a rising pitch, and not one was able to perceive the fall that occurs at approximately 275ms. These *ad hoc* perception tests confirm my suspicion that the latter portion of the TBU of the [35] tone does not contribute significantly to the recognition of its contour over the last 75ms of its duration. What perhaps is the most salient feature of the [35] tone is the high target. Once that is achieved, the intensity is seen in Praat spectrograms to fall off at the same time that the pitch begins to fall, and arguably that is why my consultants were unable to discern the fall that is observable in Figure 3. The fall is therefore most likely to be the insignificant artefact of the terminal drop in intensity. That being said, the overall

duration may also be important to the [35] tone, the average of which is considerably longer than any of the other tones.

A related issue concerns which part of the rhyme should be considered the most perceptually salient, or indeed if this is a universally shared property of tone languages. Some cues for the identification of tones are almost certainly more prominent at different durations of the TBU, but there may also be significant cross-linguistic differences with respect to what constitutes a perceptually relevant feature of a tone.

This is suggested by Brunelle's (2009) experimental research on Vietnamese, the northern dialect of which has two glottalized tones: one with mid-laryngealization, and the other manifesting final glottalization. Brunelle used resynthesized stimuli to test the perceptions of native speakers and concluded that voice quality was a more salient cue than pitch for these two tones. It stands to reason that if the most important feature for the identification of a tone is the perception of voice quality, and this is produced late in the rhyme, then interlocutors must need to focus on the terminal part of the TBU, as that portion conveys the linguistically significant cue. Conversely, a language with level register tones that retains a relatively constant pitch height over the total duration of its rhyme may achieve a perceptually relevant cue at the very beginning of the TBU, in further contrast to a contour tone that varies its pitch height over time. I therefore surmise that there must be considerable language-to-language and within-language variation with respect to which part of the TBU conveys tonally relevant cues for distinguishing tonal contrasts.

5. GRAMMATICAL TONES, FLOATING TONES AND TONE SANDHI. In addition to encoding lexical contrasts, some languages use tone for grammatical purposes, although this is considerably more commonly found in languages of Africa and the Americas than in Asian languages.¹⁸ Benedict (1972: 88) reports a rare example in Modern Burmese, which uses the creaky tone (ḡ) to mark nouns with 'specialized force' and diminutives:

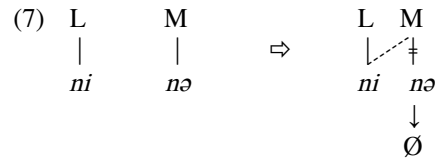
(6)	<i>lyà</i>	'thin'	<i>lyá</i>	'flimsy'
	<i>khà</i>	'bitter'	<i>khá</i>	'bitterish'
	<i>lu</i>	'man'	<i>lú</i>	pejorative
	<i>ne</i>	'sun'	<i>né</i>	'day'

The creaky tone of Burmese is additionally used for subordinating pronouns and proper nouns, e.g. *ŋa* 'I' ~ *ŋá* 'mine', and is used in a variety of derivational and valency changing functions, e.g. *lań* 'revolve, turn around' ~ *hláń* 'turn around, make revolve', and *ńi* 'to be even' ~ *hńi* 'make even' (ibid.).

In the Mangmetong Village variety of Mongsen Ao, floating tones that encode grammatical categories have developed from the loss of the segmental material with which their particular lexical tones are associated. For example, the agentive case marker *nə̄* carries a mid tone, and the first person singular pronoun *nì* carries a low tone. Occasionally speakers optionally omit the segmental representation of the agentive case marker when narrating texts, which dissociates its mid tone. The floating mid tone then reassociates to the TBU

¹⁸ For discussions of grammatical tones in African, Athabascan and Chatino languages in this volume, the interested reader is respectively directed to Snider, Rice, and Cruz & Woodbury.

of the preceding first person singular pronoun *nì*. The output is a low rising [13] tone realized on that monosyllable, the unique contour of which encodes the agentive case of its pronominal noun phrase. This process of tone spreading can be illustrated using a standard autosegmental representation (Goldsmith 1976):

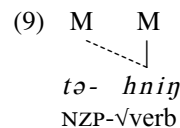


Tonal coalescences resulting in contour tones are known to be common in African tone languages, but this phenomenon is considerably rarer in Asian tone languages. It is extremely limited in Mongsen Ao and has only been attested on the first person singular pronoun. Tonal phenomena such as this are difficult to elicit and are more likely to be encountered serendipitously when working on naturalistic narrative data. On the other hand, grammatical uses of tone applying to derivational morphology and word formation processes may be more readily revealed in the general course of documenting the language of investigation. The rest of this section discusses examples of nominal derivation, floating tones, and the characteristics of tone spreading in Mongsen Ao word formation processes.

The Mangmetong village variety of Mongsen Ao appears to have tonally unspecified nominalizing prefixes. These typically copy the tone of the initial syllable of their verb roots to derive a deverbal adjective.¹⁹

- (8) *tə-pətī* (NZP-big) 'big'
tə-hnīŋ (NZP-ripe) 'ripe'
tə-hnák (NZP-black) 'black'

The leftward spread of the tone from the initial syllable of the verb root to the tonally unspecified nominalizing prefix in *tə-hnīŋ* (NZP-ripe) is represented autosegmentally in (9).



Adjectives expressing the comparative degree are further derived by the addition of a nominalizing suffix, which consistently has an output mid tone in the comparative derivation as part of its phonetic representation. Note that the derivation of 'blacker' additionally results in the mid tone of the nominalizing suffix spreading leftward and replacing the high tone of the verb root *hnák* 'be black'. This logically follows assignment of the high tone to the

¹⁹ However, exceptions are not infrequently encountered in other varieties of Mongsen Ao. The variety spoken in Waromung Village, for example, has examples of words in which the nominalizing prefix appears to be independently specified for tone, e.g. *tə-màŋ* 'NZP-dark'. This is found to contrast with segmentally identical words in the perception test described in Section 6 below.

nominalizing prefix in the preceding cycle of derivation illustrated by (8–9), since that morpheme retains the high tone originally copied from the tone of the verb root.

- (10) *tə-pətí-pā?* (NZIP-big-NMLZ) ‘bigger’
tə-hníŋ-pā? (NZIP-ripe-NMLZ) ‘riper’
tə-hnāk-pā? (NZIP-black-NMLZ) ‘blacker’

Adjectives expressing the superlative degree encode this derivation morphologically with a mid-toned emphatic suffix *-thí* (glossed SUP); this is affixed between the root and the nominalizing suffix. The superlative derivation triggers a further interesting manifestation: the syllable immediately preceding the emphatic suffix is now realized with a high tone.

- (11) *tə-pətí-thí-pā?* (NZIP-big-SUP-NMLZ) ‘biggest’
tə-hníŋ-thí-pā? (NZIP-ripe-SUP-NMLZ) ‘ripest’
tə-hnāk-thí-pā? (NZIP-black-SUP-NMLZ) ‘blackest’

One way of accounting for this tone sandhi is to assume that the emphatic suffix carries a floating high tone, viz. $\acute{t}hí$, and that in the process of word formation, the floating high tone spreads leftward and displaces the tone of the preceding syllable, e.g.

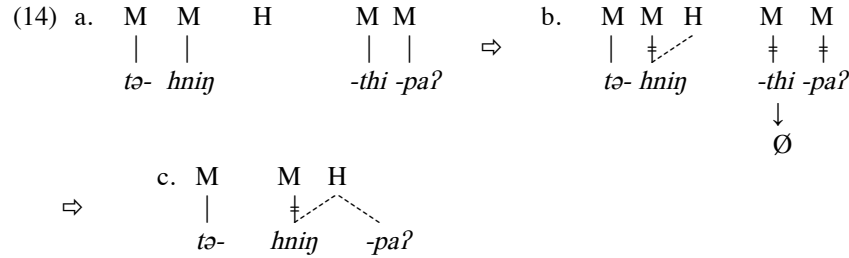
- (12) M M H M M \Rightarrow M M H M M
| | | | | | | |
tə- hníŋ -thí -pā? *tə- hníŋ -thí -pā?*

The emphatic suffix and its mid tone can also be omitted entirely, with the result that its floating high tone is now dissociated from its host syllable. The dissociated floating high tone subsequently spreads leftward, attaching to the preceding syllable and displacing its tone. This reassociated floating high tone thereby becomes the principal means of encoding the superlative meaning of this derivation.

- (13) *tə-pətí-pá?* (NZIP-big: SUP-NMLZ) ‘biggest’
tə-hníŋ-pá? (NZIP-ripe: SUP-NMLZ) ‘ripest’
tə-hnāk-pá? (NZIP-black: SUP-NMLZ) ‘blackest’

The remaining tone sandhi change to account for now is the realization of the high tone on the stem-final nominalizing suffix. We might infer that the floating high tone has a perseverative effect on the output tone of the nominalizing suffix and exerts its influence in a rightward direction as well, displacing the mid tone of that stem-final suffix. The general tendency we have seen in the data of this section is for tone changes to spread from right to left, so this seems to go against the grain. However, maintaining a high pitch over the rest of the stem seems prosodically consistent with the emphatic semantics of the superlative derivation, as pitch is easily exploited to express emphasis.

An attempt to capture this schematically using an autosegmental analysis is presented in the figure of (14) below.



6. PERCEPTION TESTS. A perception test is a useful way of determining the extent to which the tones of a language are perceptually recognizable to a native speaker when the words that carry them are isolated from contextual clues. In the following discussion I outline the results of a perception test applied to Waromung Village Mongsen Ao data and reported in Coupe 2003. This was used as a definitive measure of the significance of lexically contrastive pitch in this language, and the test additionally served to check the accuracy of the tonal analysis.

The stimuli were prepared as follows. First, some minimally contrastive sets of words were recorded – these happened to be disyllables, as at the time I had not encountered minimal sets of monosyllables such as those presented above in (5a). The disyllables were then quasi-randomly mixed with approximately 200 recorded utterances of other lexical items in a single recording (including other minimally contrastive sets of words), resulting in a word list of 225 items containing nine elicitations of each test word.

The following segmentally identical Mongsen Ao words were used. All except *təmáŋ* ‘all’ have a grammatical prefix,²⁰ but this has no impact on the realization of their tones, in contrast to grammatical affixes in African languages (e.g. Snider, this volume).

- (15) [təmáŋ] ‘all’
 [təmāŋ] ‘body’ (RL-body)
 [təmāŋ] ‘don’t believe’ (PROH-believe)
 [təmàŋ] ‘dark’ (N郑-dark)

The F_0 contours of these words were plotted in a line chart using the methodology described for the acoustic analysis of Khiamniungan tonemes in §4, and their pitch shapes are represented below in Figure 4. Note that the prefix of *təmāŋ* ‘N郑-dark’ is not copied from the tone of the verb root, so tone may have lexicalized on the prefix of this word (cf. the examples of [8] above). One of the questions this test was designed to address was whether the very short initial syllable carried tonally relevant pitch. The results of the perception test indicated that this could be the case despite the fleeting nature of their rhyme durations, which in most cases amounted to no more than three or four glottal pulsations.

²⁰ The relational prefix *tə-* and the nominalizing prefix *tə-* are distinguishable grammatically by their distributions, and also by its derivational functions in the case of the nominalizing prefix.

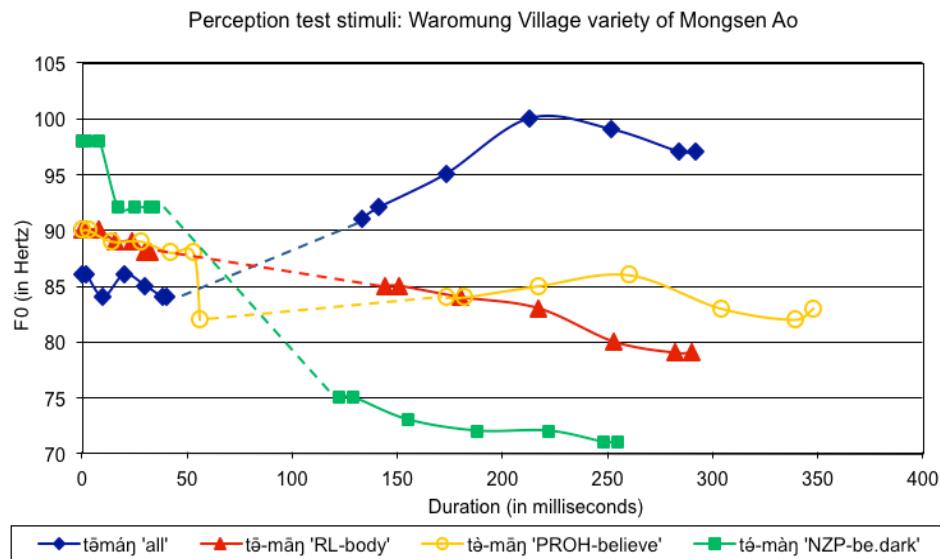


FIGURE 4. Individual F_0 shapes and rhyme durations for the test words of example (15) above. The broken lines represent an interpolated nasal consonant (adapted from Coupe 2003: 97).

The entire recording was played to a consultant, who was invited to respond to each stimulus with a translation of the token he heard, and his responses were recorded as being correct or incorrect. In Table 2 below, the numbers in the shaded cells indicate the number of correct responses to the nine stimuli he heard. To illustrate, 9/9 in the bottom right cell indicates that the consultant correctly recognized examples of *tā-māŋ* 'NZP-dark' on each of the nine occasions he heard this stimulus, but he had some confusions between the low-mid tones of *tāmāŋ* 'PROH-believe' and the mid-mid tones of *tāmāŋ* 'RL-body'. That is, four out of the nine times that he was played *tāmāŋ* 'PROH-believe', he heard it as *tāmāŋ* 'RL-body'. It is significant that the confusions were unidirectional, because it suggests that they are not attributable to chance. If they were, then we would expect to find confusions with *tāmāŋ* 'PROH-believe' being wrongly identified when the consultant heard *tāmāŋ* 'RL-body', but the results show 100% intelligibility for this stimulus-and-response pair.

Given the similar F_0 height and contours of the confused words (represented by the red- and mustard-coloured lines in Figure 4), it is not so surprising that confusions between these two words occurred, as their initial syllables are almost identical. It appears that the main perceptual cue for the low tone of *tāmāŋ* 'PROH-believe' is actually the interpolated nasal, the F_0 of which drops to 82Hz at the transition from the rhyme of the initial syllable to the sonorant onset. In contrast, *tāmāŋ* 'RL-body' has an F_0 of 88Hz at this point. The F_0 values of the nasal onsets were not calculated and are represented by broken lines in Figure 4, but these were measured and found to maintain a 5-6Hz difference for a considerable portion of their durations. This difference, together with the late drop to 82Hz seen at end of the minor syllable of *tāmāŋ* 'PROH-believe', are likely to be the main cues for conveying the contrastive low tone. In view of this, it may well be the case that an interpolated

sonorant could carry linguistically significant pitch in disyllables (contra Howie 1974), but this is something requiring further investigation.

Stimulus→ ↓Response	<i>t̄máŋ</i> (MH) 'all'	<i>t̄māŋ</i> (MM) 'NPF-body'	<i>t̄māŋ</i> (LM) PROH-believe	<i>t̄-màŋ</i> (HL) 'NPF-dark'
<i>t̄máŋ</i> (MH) 'all'	9/9			
<i>t̄māŋ</i> (MM) 'NPF-body'		9/9	4	
<i>t̄māŋ</i> (LM) PROH-believe			5/9	
<i>t̄-màŋ</i> (HL) 'NPF-dark'				9/9

TABLE 8. Confusion matrix for the Mongsen Ao test words of example (15) above.
(Coupe 2003: 96)

Overall the perception test revealed a high rate of intelligibility (89%), with confusions only occurring between the low-mid and mid-mid tokens, thus indicating that tone has a lexically contrastive function in Mongsen Ao. The results additionally suggest that at least some of the minor syllables of disyllabic words also carry lexically contrastive pitch in this variety.

Some refinements could be made to the methodology of this perception test. Firstly, only a single elicitation was recorded; this was then dubbed a total of nine times onto the test recording. It would be better to have nine separate tokens rather than repeating the same one nine times, so as to minimize the chance of one performance error in tone production affecting the results of the entire experiment. Secondly, it is probably best to test a speaker's ability to recognize the tones of their own elicitations initially, because these words are heard in isolation and the consultant may struggle with establishing the pitch range of the tones of a speaker unfamiliar to them, especially if the test word is uttered in the absence of other contextual cues. However, it might also be revealing to have a mix of other speakers provide the stimuli, so as to see how much impact this might have on intelligibility.

7. CONCLUDING COMMENTS. This paper has outlined a methodology for establishing tonal contrasts in a language using auditory techniques, then expanded on ways of corroborating the analysis by using instrumental techniques and perception tests. Perhaps the most crucial step in this process is ensuring that the data is not affected by the imposition of intonation on the individual syllables of words. The best way to control for this is to use carrier sentences, but the investigator should remain vigilant to the fact that even a well-chosen frame can be subjected to global intonation patterns that nullify the value of the recorded data for tonal analysis.

REFERENCES

- Baken, R.J. 1987. *Clinical measurement of speech and voice*. Boston: College-Hill Press.
- Benedict, Paul K. 1972. *Sino-Tibetan: a conspectus*. (Contributing editor James A. Matisoff). Cambridge: Cambridge University Press.
- Boersma, Paul & David Weenink. 2013. Praat: doing phonetics by computer. [Computer program]. Version 5.3.51, retrieved 2 June 2013 from <http://www.praat.org/>
- Brown, J. Marvin. 1985. *From ancient Thai to modern dialects and other writings on historical Thai linguistics*. Bangkok: White Lotus Co. Ltd.
- Brunelle, Marc. 2009. Tone perception in Northern and Southern Vietnamese. *Journal of Phonetics* 37: 79–96.
- Chao, Yuen Ren. 1930. A system of tone letters. *Le Maître Phonétique* 30: 24–27.
- Chao, Yuen Ren. 1968. *A grammar of spoken Chinese*. Berkeley & Los Angeles: University of California Press.
- Coupe, Alexander R. 2003. *A phonetic and phonological description of Ao: a Tibeto-Burman language of Nagaland, north-east India*. Canberra: Pacific Linguistics.
- Coupe, Alexander R. 2007. *A grammar of Mongsen Ao*. Berlin & New York: Mouton de Gruyter.
- Coupe, Alexander R. 2011. On core case marking patterns in two Tibeto-Burman languages of Nagaland. *Linguistics of the Tibeto-Burman Area* Vol 34.1: 21–47.
- Coupe, Alexander R. 2012. Referent tracking strategies in Mongsen Ao and Chang discourse. Paper read at the 22nd Conference of the South-East Asian Linguistics Society (SEALS22). Agay, France, 30 May–2 June. <http://jseals.org/seals22/seals22.htm>
- Coupe, Alexander R. 2013. Tense, but in the mood: diachronic perspectives on the representation of time in Ao. *Language and Linguistics* Vol. 14.6: 1105–1138.
- Christaller, Rev. Johan G. 1875. *Grammar of the Asante and Fante language, also known as Tshi [Chwee, Twi], based on the Akuapem dialect with reference to other (Akan and Fante) dialects*. Basel: Basel Evangelical Missionary Society.
- Cruz, Emiliana & Anthony C. Woodbury. Finding a way into a family of tone languages: The story and methods of the Chatino Language Documentation Project. *Language Documentation and Conservation* 8: 490–524
- Evans, Jonathan. 2001. Contact-induced tonogenesis in Southern Qiang. *Language and Linguistics*, Vol. 2.2: 63–110.
- Gedney, William A. 1972. A checklist for determining tones in Tai dialects. In Smith, Estelle M. (ed.) *Studies in linguistics in honour of George L. Trager*, 423–37. The Hague: Mouton.
- Goldsmith, John A. 1976. Autosegmental phonology. Doctoral dissertation, MIT.
- Goldsmith, John A. 1990. *Autosegmental and metrical phonology*. Oxford: Blackwell.
- Haudricourt, André-George. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* 242: 69–82.
- Haudricourt, André-George. 1972. Two and three-way splitting of tonal systems in some Far Eastern languages (translated by Christopher Court). In Harris, Jimmy G. & Richard B. Noss (eds.) *A Conference on Tai Phonetics and Phonology*, 58–86. Bangkok: Mahidol University.
- Hildebrandt, Kristine. 2007. Tone in Bodish languages: typological and sociolinguistic

- contributions. In Miestamo, Matti and Wälchli, Bernhard (eds.), *New challenges in typology: broadening the horizons and redefining the foundations*. Trends in Linguistics, Studies and Monographs 189, 77–100. Berlin and New York: Mouton de Gruyter.
- Howie, John M. 1974. On the domain of tone in Mandarin. *Phonetica* 30:129–148.
- Hyman, Larry. 2007. Elicitation as experimental phonology: Thlantlang Lai tonology. In Maria-Josep Solé, Pam Beddor & Manjari Ohala (eds.), *Experimental Approaches to Phonology in Honor of John J. Ohala*, 7–24. Oxford University Press.
- Ladefoged, Peter. 2003. *Phonetic data analysis: an introduction to fieldwork and instrumental techniques*. Malden MA & Oxford: Blackwell.
- LaPolla, Randy J., with Chenglong Huang. 2003. *A grammar of Qiang; with annotated texts and glossary*. Berlin & New York: Mouton de Gruyter.
- Lehiste, Isle. 1970. *Suprasegmentals*. Cambridge, MA: MIT Press.
- Maddieson, Ian. 1997. Phonetic universals. In Hardcastle, William J. & John Laver (eds.) *The handbook of phonetic sciences*, 619–639. Oxford & Malden, MA: Blackwell.
- Matisoff, James A. 2003. *Handbook of Sino-Tibetan reconstruction: system and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California.
- Mazaudon, Martine. 1974. Notes on tone in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 1.1: 27–54.
- Mazaudon, Martine 1977. Tibeto-Burman tonogenetics. *Linguistics of the Tibeto-Burman Area* 3.2: 1–123.
- Mazaudon, Martine. 2005. On tone in Tamang and neighbouring languages: synchrony and diachrony. In Shigeki Kaji (ed.) Proceedings of the Symposium Cross-Linguistic Studies of Tonal Phenomena. Historical Development, Tone-Syntax Interface and Descriptive Studies [held at ILCAA, 14–16 Dec. 2004], 79–96 Tokyo: ILCAA, Tokyo University of Foreign Studies.
- Mazaudon, Martine, 2012. Paths to tone in the Tamang branch of Tibeto-Burman (Nepal). In Vogelaar, Gunther de & Guido Seiler (eds.) *The dialect laboratory: dialects as a testing ground for theories of language change*, 139–177. Amsterdam & New York: John Benjamins.
- Mazaudon, Martine. Studying emergent tone-systems in Nepal: pitch, phonation and word-tone in Tamang. *Language Documentation and Conservation* 8: 549–574
- Michaud, Alexis. 2011. The tones of numerals and numeral-plus-classifier phrases: on structural similarities between Naxi, Na and Laze. *Linguistics of the Tibeto-Burman Area*, Vol. 34.1: 1–26.
- Michaud, Alexis & He Xueguang. 2007. Reassociated tones and coalescent syllables in Naxi (Tibeto-Burman). *Journal of the International Phonetic Association* Vol. 37.3: 237–255.
- Michaud, Alexis & Barbara Kühnert. 2006. La courbe de F₀ des sonantes initiales de syllabe joue-t-elle un rôle prosodique? Etude-pilote de données d’anglais britannique. Proceedings of XXVIe Journées d’Etude de la Parole, 121–124. Dinard.
- Morey, Stephen. Studying tones in North East India: Tai, Singpho and Tangsa. *Language Documentation and Conservation* 8: 599–633
- Pike, Kenneth L. 1948. *Tone languages: a technique for determining the number and type of pitch contrasts in a language, with studies in tonemic substitution and fusion*. Ann Arbor: University of Michigan Press.

- Remijsen, Bert. The study of tone in languages with a quantity contrast. *Language Documentation and Conservation* 8: 634–651
- Rice, Keren. On beginning the study of the tone system of a Dene (Athabaskan) language: looking back. *Language Documentation and Conservation* 8: 652–668
- Rose, Phil. 1992. Bidirectional interaction between tone and syllable coda: acoustic evidence from Chinese. In Jeff Pittam (ed.) *Proceedings of the fourth Australian international conference on speech science and technology*, 292–297. Canberra: Australian Speech Science and Technology Association.
- Snider, Keith. On establishing underlying tonal contrast. *Language Documentation and Conservation* 8: 669–699
- Sprigg, Richard Keith (1966). Lepcha and Balti Tibetan, tonal or non-tonal languages? *Asia Major* (New Series) 12: 185–201.
- Styler, Will. 2013. *Using PRAAT for linguistic research*, v.1.3.6. <http://savethevowels.org/praat/>. Accessed 2013-10-23.
- Sun, Jackson T.-S. 2003. Variegated tonal developments in Tibetan. In Bradley, David, Randy J. LaPolla, Boyd Michailkovsky and Graham Thurgood (eds.) *Language variation: papers on variate and change in the Sinosphere and in the Indosphere in honour of James A. Matisoff*, 35–51. Canberra: Pacific Linguistics.
- Thurgood, Graham. 1981. *Notes on the origins of Burmese creaky tone*, *Monumenta Serindica* 9. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Thurgood, Graham. 2002. Vietnamese and tonogenesis. *Diachronica* 19.2: 333–363.
- Watters, David E. 2003. Kham. In Thurgood, Graham & Randy J. LaPolla (eds.) *The Sino-Tibetan languages*, 683–704. Oxford: Routledge.
- Weidert, Alfons. 1987. *Tibeto-Burman tonology*. Amsterdam/Philadelphia: John Benjamins.
- Yu, Kristine M. The experimental state of mind in elicitation: illustrations from tonal fieldwork. *Language Documentation and Conservation* 8: 700–739

Alexander R. Coupe
arcoupe@ntu.edu.sg