

Reassociated tones and coalescent syllables in Naxi (Tibeto-Burman)

Alexis Michaud, Xueguang He

▶ **To cite this version:**

Alexis Michaud, Xueguang He. Reassociated tones and coalescent syllables in Naxi (Tibeto-Burman). Journal of the International Phonetic Association, 2007, 37 (3), pp.237-255. hal-00144485

HAL Id: hal-00144485
https://shs.hal.science/hal-00144485

Submitted on 3 May 2007

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Preprint version of article published in Journal of the International Phonetic Association *in 2007*

Text and original data available online from http://hal.archives-ouvertes.fr/hal-00144485

Reassociated tones and coalescent syllables in Naxi (Tibeto-Burman)

Alexis Michaud

Centre National de la Recherche Scientifique, laboratoire Langues et Civilisations à Tradition Orale

alexis.michaud@vjf.cnrs.fr

He Xueguang

Naxi Culture Education Centre, Lijiang

Abstract

The Western dialect of Naxi has four lexical tones: High, Mid, Low and Rising; the latter is rare in the lexicon. Rising contours on monosyllables are frequent in connected speech, however, as a result of a process of syllable reduction: reduction of a morpheme carrying the High tone results in re-association of its tone to the syllable that precedes it in the sentence, creating a rising contour. An experiment (with one speaker and five listeners) establishes that there is not only one rising contour that originates in tonal reassociation, as reported in earlier descriptions, but two: Low-to-High and Mid-to-High—as could be expected by analogy with phenomena observed in Niger-Congo languages and elsewhere. A second set of experiments (same speaker; six listeners) investigates the reduction of Mid- and Low-tone syllables: they reduce to [ə̄] and [ə̀], respectively, and coalesce with the preceding syllable (in Naxi, syllabic structure is simply *consonant + glide + vowel*). Unlike High-tone syllable reduction, this process stops short of complete tonal de-linking. These experiments aim to provide a complete picture of syllable reduction patterns in Naxi. It is argued that the notions of floating tones and tonal reassociation can be usefully applied to the Naxi data.

1. Introduction

The Western dialect of Naxi (纳西, a.k.a. Na 纳, Moxie 麽些, Moso 摩梭), a Tibeto-Burman language spoken in China, could appear as a textbook example of a level-tone language: each syllable carries a High (˝), Mid (¯), or Low (ˋ) lexical tone (He & Jiang 1985: 10). A rising lexical tone is also reported on a small set of native monosyllables as well as on numerous (monosyllabic) loanwords, resulting in minimal quadruplets such as the following: /lá/ 'to strike', /lā/ 'tiger', /là/ 'hand', /lǎ/ 'candle'. Rising contours are also created in connected speech by a process of syllable reduction: some monosyllabic morphemes carrying High tone can undergo a form of reduction whereby their segments are deleted and their High tone associates to the preceding syllable. If the latter has a lexical Low or Mid tone, it acquires a rising contour; if it has a High lexical tone, this tone remains unchanged.

This is reminiscent of phenomena found in other languages—typically in languages of the Niger-Congo family, as described by Goldsmith (1976), but also in other families, e.g., among creole languages, in the variety of Saramaccan studied by Good (2002)—which are now commonly modelled in terms of association lines between a tonal tier and a syllabic tier, as in (1a-b) below: when syllable σ_2 undergoes reduction (full deletion of its segmental contents, i.e. disappearance of σ_2 from the syllabic tier), its tone, T_2, reassociates to syllable σ_1.

(1a) σ_1 σ_2 (1b) σ_1



Earlier accounts of Naxi (He & Jiang 1985: 15, 54, He Zhiwu 1987: 61) only acknowledge one rising contour. For instance, He Jiren and Jiang Zhuyi (1985: 15) state that example (2a) can simplify to (2b):

(2a) [dɯ˧ **ɲi˧** **tɑ˥** dʑy˧ se˩]

one day only *existential* *perfective aspect*

'There is only one day left.'

(2b) [dɯ˧ **ɲi˩˧** dʑy˧ se˩]

one day *(tone becomes rising)* *existential* *perfective aspect*

(Same glosses; the only difference with (2a) is that /tɑ˥/, 'only', undergoes reduction. The syllables involved in this process are emphasised typographically for the sake of clarity.)

This notation reflects the belief that the four lexical tones of Naxi are High ˥, Mid ˧, Low-falling ˩ and Low-rising ˩˧, and that the elliptical construction (2b) results in the creation of a Low-rising contour identical to the lexical Low-rising tone. By analogy with phenomena observed in Niger-Congo languages and elsewhere, the first author of the present article hypothesised (i) that the 'Low-falling' tone of Naxi was simply a Low tone, and (ii) that two contours resulting from syllable reduction were to be distinguished, the rising contours observed in continuous speech corresponding either to Low plus High (̀ ́), or to Mid plus High (̄ ́). Thus, (2b) would be transcribed as: /dūɯ̄.ŋī ́.gȳ.sè/. (Note in passing that He's [ɲ] is analysed phonemically as /ŋ/, and his [dʑ] as /g/; for details, see Michailovsky & Michaud 2006.) Likewise, in example (3), from my field data, what is underlyingly a Low-tone morpheme /kʰà/ occurs with a rising fundamental frequency contour: in (3b), /kʰà/ becomes /kʰǎ́/, i.e. with a Low-to-High contour. Example (3b) is taken from a dialogue translated from the Chinese; the explicit phrasing, (3a), was later elicited from the same language consultant.

(3a) ŋɤ̀ ʈʂʰūɯ̄ **kʰà ́** ʈʂʰū lù tʰɣ̄ mɤ̀.

1st sg. DEICT. **moment + only *(reduced to H)*** here reach affirmative final particle

'I've only just arrived here. (= I'm a newcomer here.)'

The explicit phrasing is the following:

(3b) ŋɤ̀ ʈʂʰūɪ **kʰà tá** ʈʂʰū lù tʰɣ̄ mɤ̀.

'I've only just arrived here. (= I'm a newcomer here.)' (Same glosses as (3a); **/tá/**, 'only', is not reduced.)

The syllables involved in this process are typographically emphasised for the sake of clarity.

In (3a) the segments of adverb /tá/ are deleted, and its tone is reassociated to the preceding syllable, /kʰà/ 'moment'; the stylistic nuance is, predictably, that (3b) is more colloquial than (3a). This is frequent in connected speech, as illustrated in figure 1, which shows a spectrogram and fundamental frequency (hereafter F_0) tracing of an excerpt from a narrative, (4):

(4) ɪ́ ‖ tō m͡bà kā **ʂù́** lūɪ

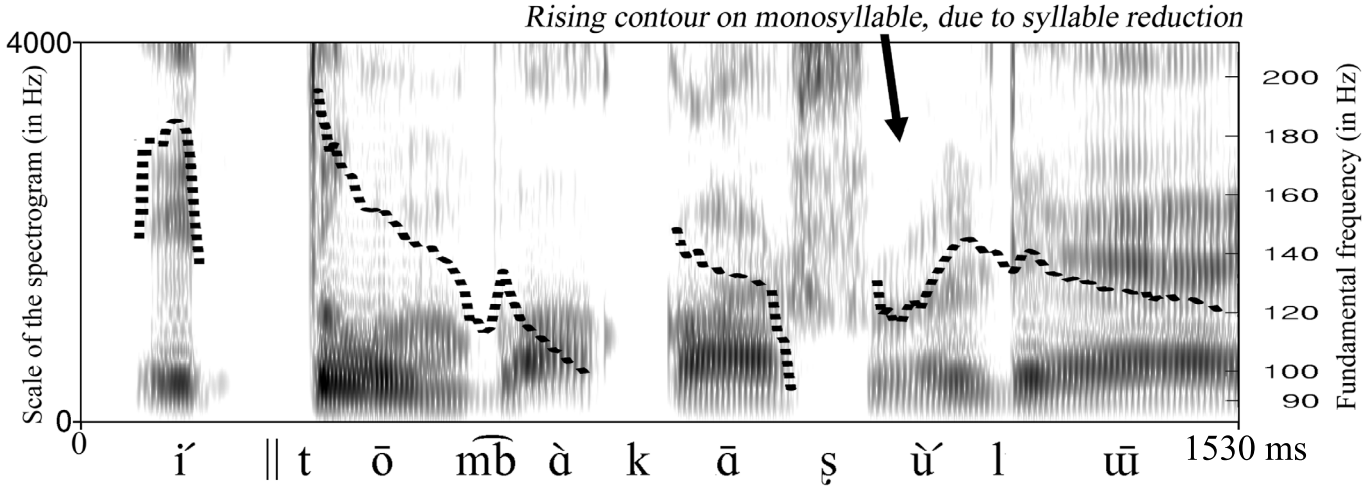
discourse partic. priest forward fetch (lexical form: **/ʂù/**) + added High tone come

... and [they] invited the priest to come along...

[INSERT FIGURE 1 HERE]

In (4), the rising contour on the verb 'to fetch' points to an added High tone; this High tone is due to the reduction of the grammatical word /pú/. As a verb, /pú/ means 'to bring, to offer'; in its grammaticalised uses, it has been described as a 'gerundiser' by Pinson (1998: 11). A synchronic overview of High-tone syllables that can undergo reduction to a High tone, which reattaches to the preceding syllable, is provided in Michaud (2006a).

Section 1 describes experiments on High-tone syllable reduction. It is argued that the notions of floating tones and tonal reassociation can be usefully applied to the Naxi data. Section 2 aims at presenting the broader picture of syllable reduction in Naxi by investigating the reduction of Mid- and Low-tone syllables to a coalescent [ə̄] and [ə̀], respectively. This process stops short of complete segmental deletion and 'tonal de-linking'.


Rising contour on monosyllable, due to syllable reduction
4000
0
Scale of the spectrogram (in Hz)
200
180
160
140
120
100
90
Fundamental frequency (in Hz)
0
1530 ms
í || t ō m͡b à k ā ş ù́ l ɯ̄

General overview of the research method

The present analyses are essentially based on two dialects of Naxi, AS (for /ā ʂə̀/) and FK (for /fý kʰō/); see Michailovsky and Michaud (2006) and Michaud (2006b) for details. Unless otherwise stated, the utterances cited come from the AS dialect, which belongs to Western Naxi (on the division between Western and Eastern Naxi: He & Jiang 1985:104-116). First-hand data from other dialects are also adduced when they yield relevant evidence.

Production data were elicited from one main language consultant for each dialect: a list of 1,000 words; 800 sentences translated from Chinese; and narratives. These data were sifted for examples of syllable deletion and tonal changes; rephrasings were tested with the language consultant; and production and perception experiments involving more subjects (reported in detail below) were set up to confirm and complement the observations. To obtain a precise measurement of F_0—and indications on voice quality, not reported here—, an electroglottographic signal was recorded throughout recording sessions.

2. Experiments on High-tone syllable reduction, establishing a difference between the /MH/ contour and the /LH/ contour

As the /LH/ and /MH/ contours are not distinguished in previous descriptions of Naxi (e.g. He & Jiang 1985: 15, 54, He Zhiwu 1987: 61), which only report one rising contour, it appeared necessary to provide evidence for the claim made here that they are distinct.

All the production data reported in this paper were elicited from the same language consultant, a native speaker of the AS dialect, aged 56 at the time of the recordings. He will be referred to hereafter as M4, the label assigned in the Naxi database collected over several field trips; this label allows for unambiguous reference. In the course of work sessions, M4

developed an awareness of the tonal change in utterances such as (2a-b, 3a-b) above. He indicated that there was no homophony between /LH/ and /MH/ sequences. His intuition alone is not sufficient to settle the issue, however. The experimental setup used for confirmation was the following: seven minimal pairs of M- and L-tone predicates (verbs and adjectives), such as: /ndzɯ̄/ 'eat', /ndzɯ̀/ 'sit', were recorded in randomised order in carrier sentences where they received a floating High tone from the following morpheme. These M- and L-tone predicates were interspersed with 13 other predicates so that the M-*vs.*-L pairs would not be conspicuous. The list of recorded items is provided in the Appendix, together with an internet link to the original audio and electroglottographic data. Two different carrier sentences were used, (5) and (6), with a view to obtaining evidence on the hypothesised contrast in two different contexts.

(5) t^hɯ̄ mɤ̄ *TARGET SYLL.* ˊ**sé** ‖ sè bē bē bɯ̄ lē ?

3rd sg. NEG. (predicate) COND./TOP. how do go interrogative particle

In the case of a verb: 'If (s)he doesn't *(TARGET PREDICATE)*, what are we going to do ?' In the case of an adjective: 'If it is not *(TARGET PREDICATE)*, what are we going to do?'

In (6), the target syllable follows a major group juncture:

(6) t^hɯ́ ‖ *TARGET SYLL.* ˊ**sé** ‖ sè bē bē bɯ̄ lē ?

then, if so (predicate) COND./TOP. how do go interrogative particle

'But then, concerning *(TARGET PREDICATE)*, what are we going to do?'

The conditional/topic /ˊ sé/ carries a floating High tone—tentatively placed to the left of the syllable in the transcription used here—which attaches to the preceding syllable, i.e. to the target syllable of the experiment; for details on this process, see Michaud (2006a). In this sentence, the COND./TOP. serves as topicaliser, and the predicate behaves as a noun; it can be glossed as follows, taking the verb 'eat' as an example: 'But then, what about the eating?', i.e. 'How are we going to manage the meals?'

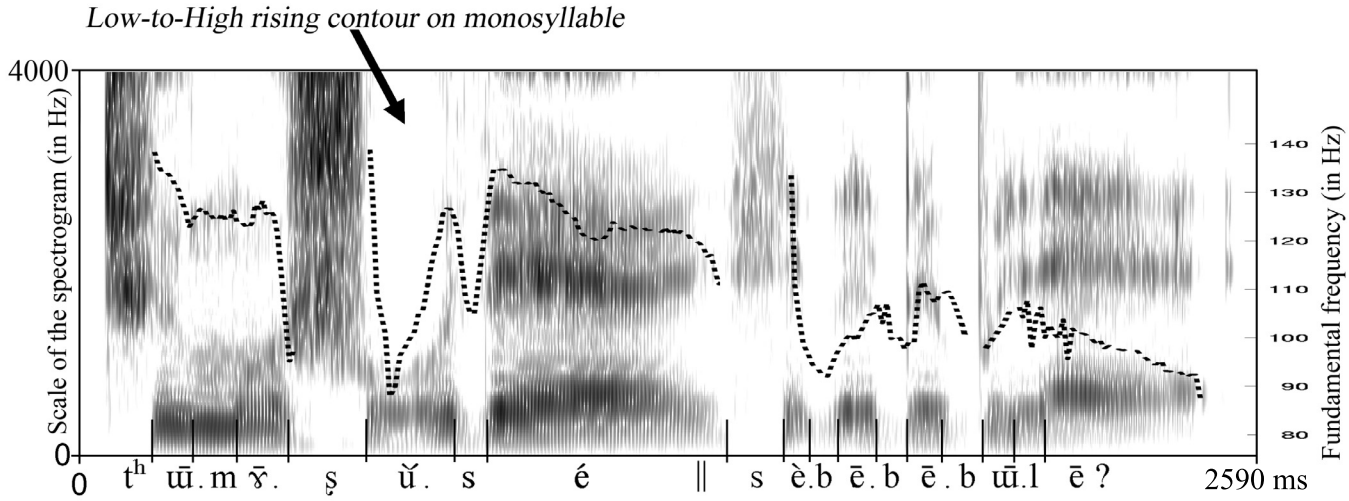
The conditions of recording (and playback, see 2.2 below) were optimal: high-fidelity equipment (TASCAM DA-P1 Digital Audio Tape, and EG2-PC glottograph, documented in Rothenberg 1992) and a quiet environment.

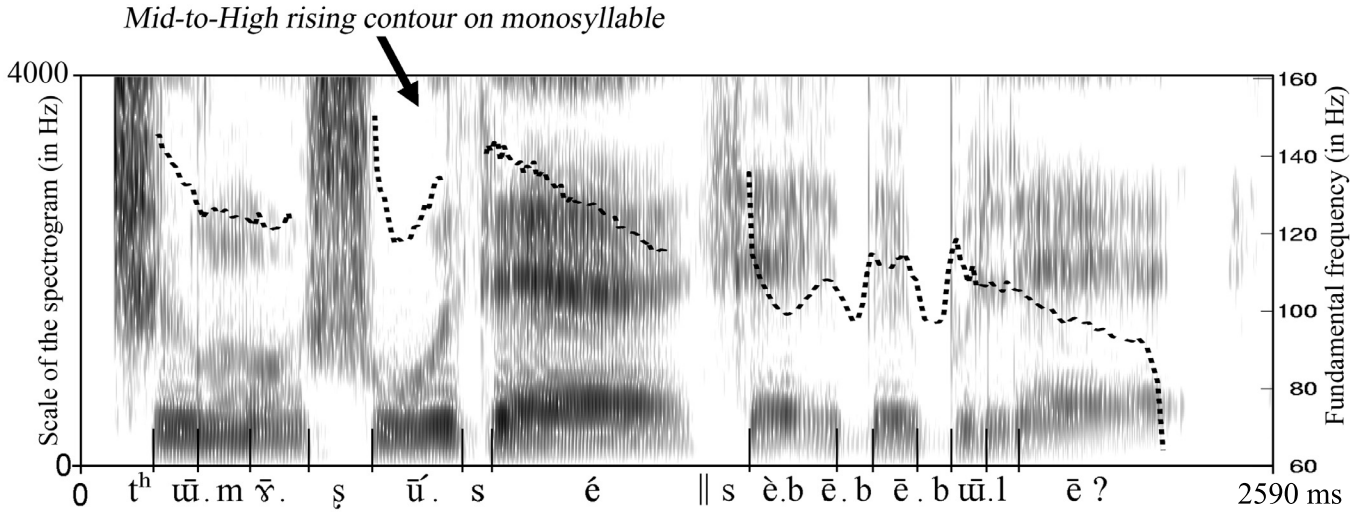
2.1. Production data

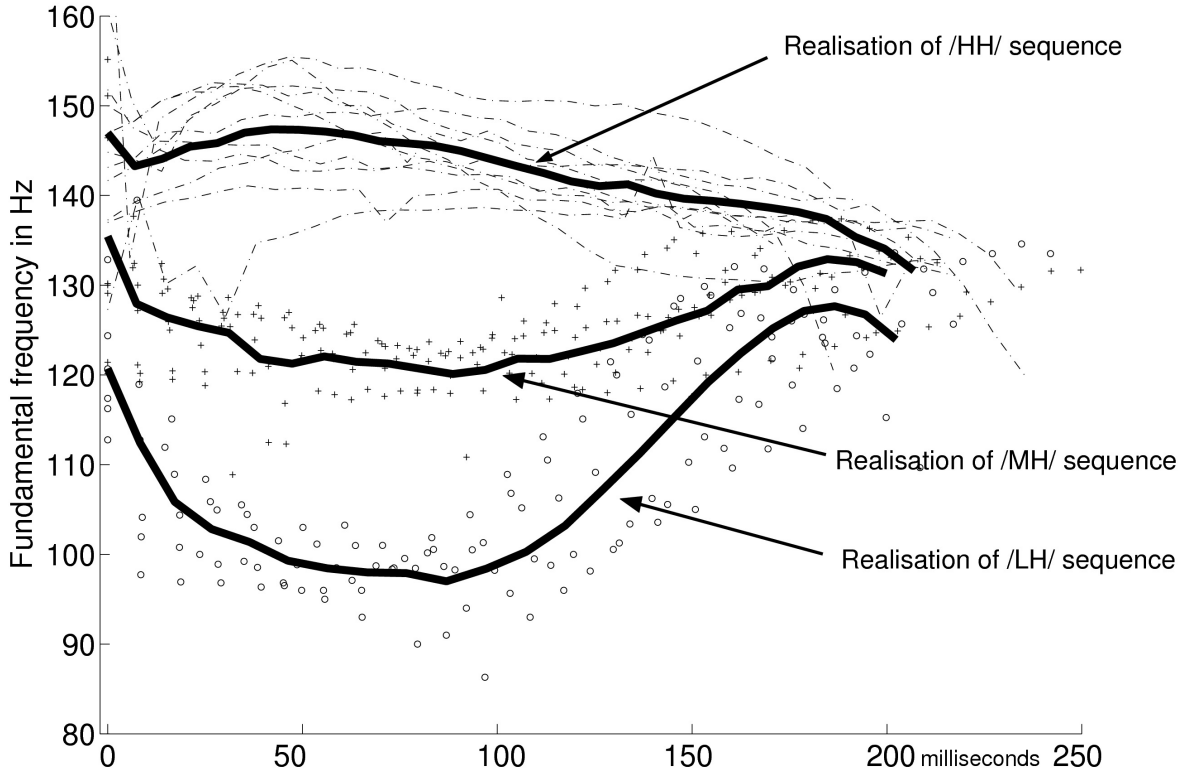
Figures 2 and 3 show verbs /ʂù/ 'to look for' and /ʂū/ 'to receive, to get', respectively, in carrier sentence (5)[1]. [INSERT FIGURES 2 AND 3 HERE] They show the presence of a rising F_0 contour on the syllables at issue. The vowel that bears this rising contour has greater duration than the other non-group-final vowels in the sentence; this may be put down to physiological factors: raising F_0 requires more time than lowering it (Ohala 1978: 30-31, Ohala & Ewan 1973, Sundberg 1979). The degree of lengthening is not considerable, however. The vowel /u/ is much shorter than the vowel of the group-final particle /sé/ and the sentence-final particle /lé/, which are lengthened to indicate intonational junctures: vowel length, which is not phonemic in Naxi, is used extensively to signal intonational phrasing (Michaud 2005:41-84). The F_0 range in which the rising contour is realised appears to differ according to the lexical tone of the syllable: it is lower for /ʂù/ (figure 2; see also figure 1, which contains the same verb) than for /ʂū/ (figure 3).

Averaged F_0 curves for the syllables recorded within carrier sentence (5) are plotted in figure 4, which presents both the raw data and curves averaged under the technical computing environment MATLAB following a procedure described in Michaud (2004). Due

[1] The /s/ of the conditional/topicaliser / ˊ sé/ in figure 2 is voiced throughout. This is frequently observed (Michaud 2005:365-367), due to the fact that function words are prosodically weaker than lexical words in Naxi (as in many other languages). This effect is not systematic: the same morpheme is less clearly voiced in figure 3. The same account holds for the realisation of the initial consonant of /sè/ in figure 6.


Low-to-High rising contour on monosyllable
4000
0
Scale of the spectrogram (in Hz)
Fundamental frequency (in Hz)
140
130
120
110
100
90
80
0
t^h ūɪ̄. m ɤ̄. ʂ ŭ. s é || s èb ēb ē b ūl ē ?
2590 ms


Mid-to-High rising contour on monosyllable
4000
0
Scale of the spectrogram (in Hz)
160
140
120
100
80
60
Fundamental frequency (in Hz)
0
t^h ūī . m ɤ̄ . ʂ ū́ . s é || s è.b ē . b ē . b ūī . l ē ?
2590 ms


Realisation of /HH/ sequence
Realisation of /MH/ sequence
Realisation of /LH/ sequence
Fundamental frequency in Hz
160
150
140
130
120
110
100
90
80
0
50
100
150
200
250
milliseconds

to some repetitions, the total number of items is 25. [INSERT FIGURE 4 HERE] From the figure, it is clear that a High tonal target is present, as predicted by the autosegmental tonal model. The /LH/ and /MH/ sequences are rising, without exception, whereas the /HH/ sequences are realised as a gently decreasing curve, similar to that of syllables with a simple H tones. (No experiment was set up to verify that the phonetic output of /H/ and /HH/ is actually identical.) The raw curves are not smooth due to the fact that they were obtained, not by autocorrelation, but by detection of each of the peaks indicative of glottis-closure-instants, on the derivative of the electroglottographic signal; this yields a precise estimation of the length of each glottal cycle, but no condition is placed on the continuity of the values from one cycle to the next (cf. Henrich et al. 2004 and references therein). The software that we developed is available online, with some documentation: see Appendix.

The curves corresponding to /LH/ and /MH/ are clearly different, though final values are close. The final High target is perturbed in the expected direction by assimilation to the preceding tone, L or M. The results for the items in carrier sentence (6) are similar.

2.2. Perception data

Perception data were elicited to verify whether the difference between syllables with /LH/ and /MH/ contours is perceived by listeners. Five native speakers participated in the listening test. One of them was M4, the language consultant who had produced the utterances, and with whom the experimental setup had been discussed; the other four had no awareness of the purpose of the test. Forty-two sentences—twenty-one containing /LH/ contours, twenty-one /MH/ contours, inside frame 5—were played once in random order through loudspeakers; the listeners were instructed to provide a translation into Chinese for each. The intended syllable was recognised in 204 cases out of 210. This warrants the conclusion that the /LH/ vs. /MH/ distinction exists.

3. Experiments on Mid- and Low-tone syllable reduction

This section investigates Mid- and Low-tone syllable reduction. With one illuminating exception (presented in section 3.2.2), these reductions respectively result in [ə̄] and [ə̀], not in de-linking of the tone: deletion of the segments is not complete.

Naxi presents some cases of creation of new lexical items from the coalescence of two syllables: the first with Mid or Low tone, the second with High tone. This is phonetically unsurprising: syllable-initial vowels have a potential for coalescence with the preceding vowel, as they have a soft onset: they begin by a semi-vowel in isolation, and link with the preceding syllable rhyme in connected speech. For example, 'mung bean jelly', a local dish, is called /hɯ̄.á/ in AS and other dialects outside the town of Lijiang, whereas the word has simplified to a monosyllable in the dialect of the old town of Lijiang 丽江 (a.k.a. Dayanzhen dialect 大研镇): /hǎ/, carrying a rising contour tone. There is apparently a single lexical contour tone, which we transcribe as /LH/, thus yielding /hǎ/ and not /hā́/; this fact was not verified experimentally. The resulting vowel—in this case, /a/—is identical to a simple vowel, i.e. the simplification to a monosyllable does not create new vowel phonemes.

No parallel creation of falling contours is observed: in cases of coalescence, /ML/ sequences simplify to /L/, /HM/ sequences simplify to /H/; no examples of lexical items coming from H + L were observed. For example, 'pickles' is /kjā.à/ in AS and other dialects outside Lijiang, /kjà/ in Lijiang. Likewise, the verb /bɯ̄/ 'go' in association with the perfect aspect particle /sè/ can simplify to /bɤ̀/; the tone of the resulting form is simply a Low tone, not a contour tone. The /HM/ pattern is illustrated by 'money': the conservative form (AS) is /ki ˊ.jɤ̄/, the coalescent form /kjɤ́/ (old town of Lijiang), without any trace of the M tone. Comparison of tables 1 and 2 brings out the discrepancy between lexical and phrasal syllable-reduction patterns. [INSERT TABLES 1 AND 2 HERE]

Further evidence for an asymmetry in the tone system comes from the comparison of two dialects: in AS, the final particle cluster /mɣ̩ sɯ̄/ (which indicates *conjecture*) simplified to /ˊsɯ̄/; in the dialect of /ndɑ̄ lè/ (丽江，金山乡，贵峰大来行政下村 /大来二村), following the inversion of all /H＋M/ sequences on lexical items to /M＋H/, this same particle cluster became /mɣ̩̄ sɯ́/, then simplified to /sɯ́/, the Mid-tone syllable disappearing without trace.

It seemed interesting to gather further synchronic evidence on this issue. The results below show that, in contrast to High-tone syllables, Mid- and Low-tone syllables only reduce to [ə̄] and [ə̀], not to a Mid and Low tone, respectively—with one exception, presented in section 3.2.2.

3.1. Reduction of Low-tone syllables results in a coalescent [ə̀]

L-tone syllable reduction will be presented before M-tone reduction because it is more frequent. Low-tone syllables were observed to reduce to [ə̀].

3.1.1. Reduction of Low-tone grammatical words to [ə̀]

The particle /t^{h}è/, which conveys 'static aspect' (Pinson 1998:103), frequently simplifies to [ə̀], a process already reported by Fu Maoji (1984:314). Fu's fieldwork dates back to the 1930s and 1940s; over half a century later, this morpheme shows no signs of simplifying further and becoming a floating Low tone.

My data show this pattern to extend to other grammatical words; by order of frequency: the copula /wù/, the perfective aspect particle /sè/, the ongoing aspect particle /nɯ̀/, and the final particle /là/, which conveys invitation: 'Let's…'. Figure 5 shows spectrograms and F_0 tracings of two realisations of the same passage of a narrative (by the same speaker, M4), /ŋgi¯.ŋgi¯. nɯ̀.mē/, without syllable reduction in the first case (7a), and with reduction in the second (7b). The folk tale from which this sequence is excerpted was

told twice from memory at a four-day interval; the phrasing of this passage is similar in both cases, but in the second recording (7b), the ongoing aspect particle /nʊ̀/ simplifies to [ə̀] and coalesces phonetically with the preceding syllable.

(7a) … li¯ lā há (…)	ŋgi¯ ŋgi¯	nʊ̀	mē…
'good old grandfather Li'	walk (reduplicated)	ongoing aspect	particle

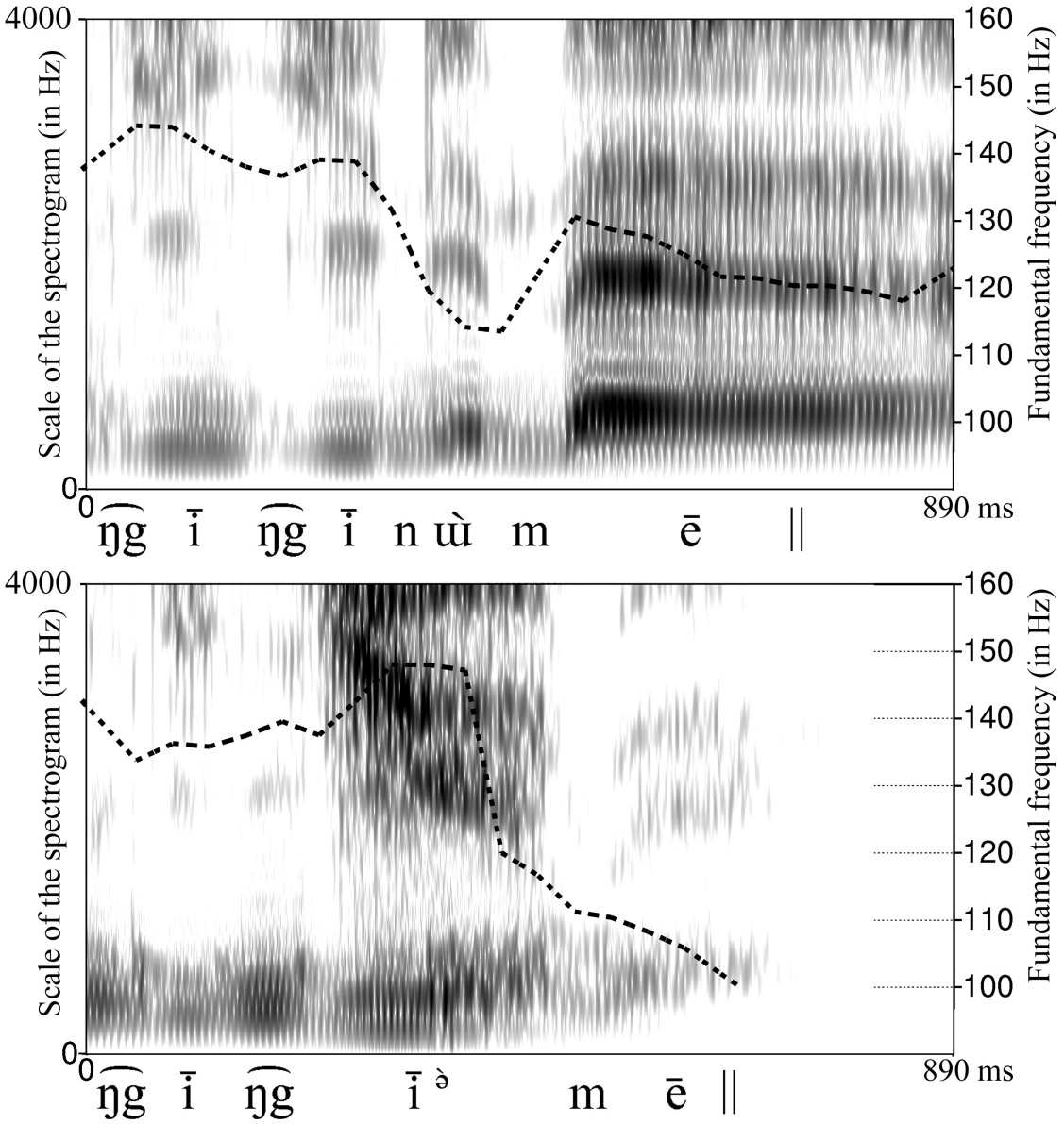
'…good old grandfather Li was taking a stroll…' (from a narrative)

(7b) ...	ŋgi¯ ŋgi¯	ə̀	mē…

(same glosses as above; particle /nʊ̀/ reduces to /ə̀/)

[INSERT FIGURE 5 HERE]

As is bound to be the case with examples taken from continuous speech, the two realisations that figure 5 aims to contrast do not provide an ideal illustration of the phenomenon of reduction: in (7b), the strong friction noise at the beginning of vowel [i] (which palatalises the initial /ŋ͡g/ to [ɲɟ]) and the nonmodal phonation towards the end of the vowel tend to blur the movement of the formants in the course of the vowel, which is expected to provide evidence of the presence of a final schwa-target. The change from a level F_0 curve on /ŋgi¯/ in (7a) to a falling contour in (7b) is likewise difficult to assess from figure 5, as the decrease in F_0 over the vowel /i/ in (7b) could in theory be put down to a gradual decrease towards an intonation-group-final low target on the final particle, i.e. a mere interpolation between the level pitch targets of successive syllables, rather than the effect of a coalescent [ə̀]. (As the tale from which these passages are excerpted were recorded with a microphone only, the F_0 curves in figure 5 were not calculated from electroglottography as in the other figures, but from the audio, using the software PRAAT: Boersma 2001.) An experiment that aims to bring out more clearly the contrast between full forms and reduced forms is reported in section 3.4; in a corpus of narratives, intermediate cases are in fact observed, e.g. /sè/ (which conveys perfect aspect) can reduce to [ɦə̀], with


4000
0
Scale of the spectrogram (in Hz)
160
150
140
130
120
110
100
Fundamental frequency (in Hz)
0
890 ms
ŋg̀ ī ŋg̀ ī n ɯ̀ m ē ||
4000
0
Scale of the spectrogram (in Hz)
160
150
140
130
120
110
100
Fundamental frequency (in Hz)
0
890 ms
ŋg̀ ī ŋg̀ ī ə m ē ||

centralised vowel and weakened fricative, and further to [ə]. This stands in contrast to High-tone syllable reduction, which in the present state of the language involves complete deletion of the segmental content of a syllable: in sentence pairs (2a-b, 3a-b), no phonetically intermediate forms span the categorical gap between, e.g., /ʈʂʰū.kʰà.tá/ (this-moment-only) and /ʈʂʰū.kʰá/ (same meaning, with reduction of final syllable).

3.1.2. Reduction of an open set of Low-tone items

In addition to grammatical words, an open set of syllables can reduce to [ə]. 'Really, truly' is pronounced /əsúəbē/ in the AS dialect (where /bē/ plays the role of an adverbialising particle). The language consultants perceive the form as trisyllabic and not quadrisyllabic, not dictating the [ə] as a separate syllable: [ə.súə.bē]. Evidence from other dialects reveals the identity of the underlying syllable: quadrisyllabic /ə.sú.gù.bē/ is standard in FK, and is in fact still intelligible in AS and the other Western dialects. The second syllable of a reduplicated verb was also observed to undergo occasional reduction:

(8) …	dū̃	sỳ	mɤ̃	wà	kʰō ə		bē…
	one	thing, sort-	NEG.	copula	pretend (RED.; *simplex*: /kʰò/)		adverbialising particle

(from a narrative, AS dialect)

'…pretending that nothing at all had happened…'

In passage (8), the second syllable of the reduplicated word /kʰō kʰò/ simplified to [ə]. The non-reduced form was dictated during the transcription session, the two informants consulted being unaware of the reduction.

3.2. Mid-tone syllables

Reduction of M-tone syllables in continuous speech is similar to L-tone syllable reduction in that it does not result in de-linking of the lexical tone. The only counterexample observed is presented in section 3.2.2.

3.2.1. Noncategorical reduction of an open set of Mid-tone syllables to [ə̄]

An open set of Mid-tone syllables can undergo reduction to [ə̄], e.g. /sɯ̄/ 'yet, as yet':

(9)	hā	ɔ́	mi ˊ	sɯ̄	lē?
	food	interrogative particle	cooked, well done	yet, as yet	interrogative particle

'Is the meal ready?'

Sentence (9) can become [hā ɔ́ mi ˊ ə̄ lē]. The phenomenon is not restricted to grammatical words: it was observed in the case of syllable /dɯ̄/ 'one (1)', in the phrase (10b) [dɯ̄ p^hé ə̄ p^hé], which constitutes a reduction of the full form (10a):

(10a) …	dɯ̄	p^hé	dɯ̄	p^hé …
	one	slice	one	slice

'slice by slice' (from a narrative, AS dialect)

3.2.2. Categorical reduction of /ī/ before final particles results in Mid-tone reassociation

There is one example where reduction of a Mid-tone syllable does result in M-tone reassociation: the reduction of the particle /ī/ when it precedes utterance-final particles. The resulting tone sequences are /HM/, lengthened M, and /LM/ after H, M and L tones, respectively. In Naxi, the particle /ī/ punctuates colloquial speech. In the AS dialect, it is almost systematically present, in the form of a floating Mid tone, before some final particles such as /mɣ̍ sɯ̄/ (conjecture), /mɑ̀ˊ/ (certainty, strong affirmation), /tsɯ́ nɯ́/ (hearsay, indirect knowledge), /tsō wɑ̀ pá/ (conjecture, call for confirmation), /nɯ̄ wā/ (surprise), /lá/ (interrogation, call for confirmation), /mɤ̀ lá/ (pressing call for confirmation), /mē/ (emphasis), /mɤ̀/ (affirmation). Most language consultants are unaware of the underlying syllable; its identity was arrived at by substitution tests.

It appears noteworthy that M-tone floating is confined to this very restricted context, and that the vowel involved is /i/. Following the reviewers' recommendation, discussion of these facts is deferred until a later publication.

3.3. Perception data on the recognition of reduced Low- and Mid-tone syllables

A test was performed with six listeners to evaluate the recoverability of the syllable having undergone reduction. In a first task, 14 utterances containing Low- and Mid-tone syllable reduction were played through loudspeakers, and the hearers were asked to repeat each utterance carefully. Their performance was written down by the author. In a second task, the identity of the syllable was verified by playing again the stimuli, proposing that an underlying syllable was present, and suggesting several candidates.

It appeared that the language consultants' awareness of the Mid- and Low-tone syllables having undergone syllable reduction varies from item to item, as is also the case for High-tone syllables having undergone reduction, described in Michaud 2006a. In the cases where a coalescent, tone-bearing schwa now systematically accompanies a certain word—typically final particles—, the identity of the underlying syllable could not be indicated by the majority of speakers. Indeed, no extra syllable was felt to be present: when asked to repeat carefully, they brought out the [ə] very clearly, and emphasised the melodic movement, but did not split the words into two syllables. Conversely, in the cases of free variation—e.g. (8) above, where the reduplicated word /k^{h}ō k^{h}ò/ becomes /k^{h}ō ə̀/—, language consultants perceive the underlying syllable vividly, so much so that they are not aware that only a reduced form is pronounced, and dictate the full form during transcription sessions. Lastly, in some cases, a stylistic difference has crept in between full form and reduced form, and paraphrases have to be elicited before the underlying syllable can be arrived at and finally confirmed by the speakers.

3.4. Measurement of the downward tilt in F_0 resulting from L-tone syllable reduction

The experiment reported here aims to quantify the changes in F_0 that result from L-tone syllable reduction. Speaker M4—recall that he is the speaker who produced the data for all the experiments—recorded 27 monosyllabic predicates (verbs or adjectives) within the following carrier sentences:

(11)	tʰū̱	*TARGET SYLL.*	sè	mē.
	3rd sg.	*(TARGET SYLL.)*	perfect aspect particle	final particle
(12)	tʰū̱	*TARGET SYLL.*	ə̀	mē.

(as above; with perfect aspect particle in its reduced form)

If the target syllable is a verb, both (11) and (12) mean that the person referred to by the pronoun has performed the action referred to by the verb: '(S)he has __ed.' If the target syllable is an adjective, these sentences mean '(S)he has become __'.

The corpus is the same as for utterances (5) and (6) above. It was intended to select H-M-L series such as: /ndzú/ 'fell (a tree)', /ndzū̱/ 'eat', /ndzù/ 'sit', but such a minimal set could not be found in every case. The predicates were arranged in random order. The speaker was aware that the focus of the experiment was on the reduction of the aspect particle /sè/. He was instructed to produce each sentence once at a slow pace, as if repeating for someone who was hard of hearing (thus eliciting carrier sentence 11), then repeat it more casually (eliciting carrier sentence 12). A translation of the predicate into the local dialect of Chinese was used as a prompt.

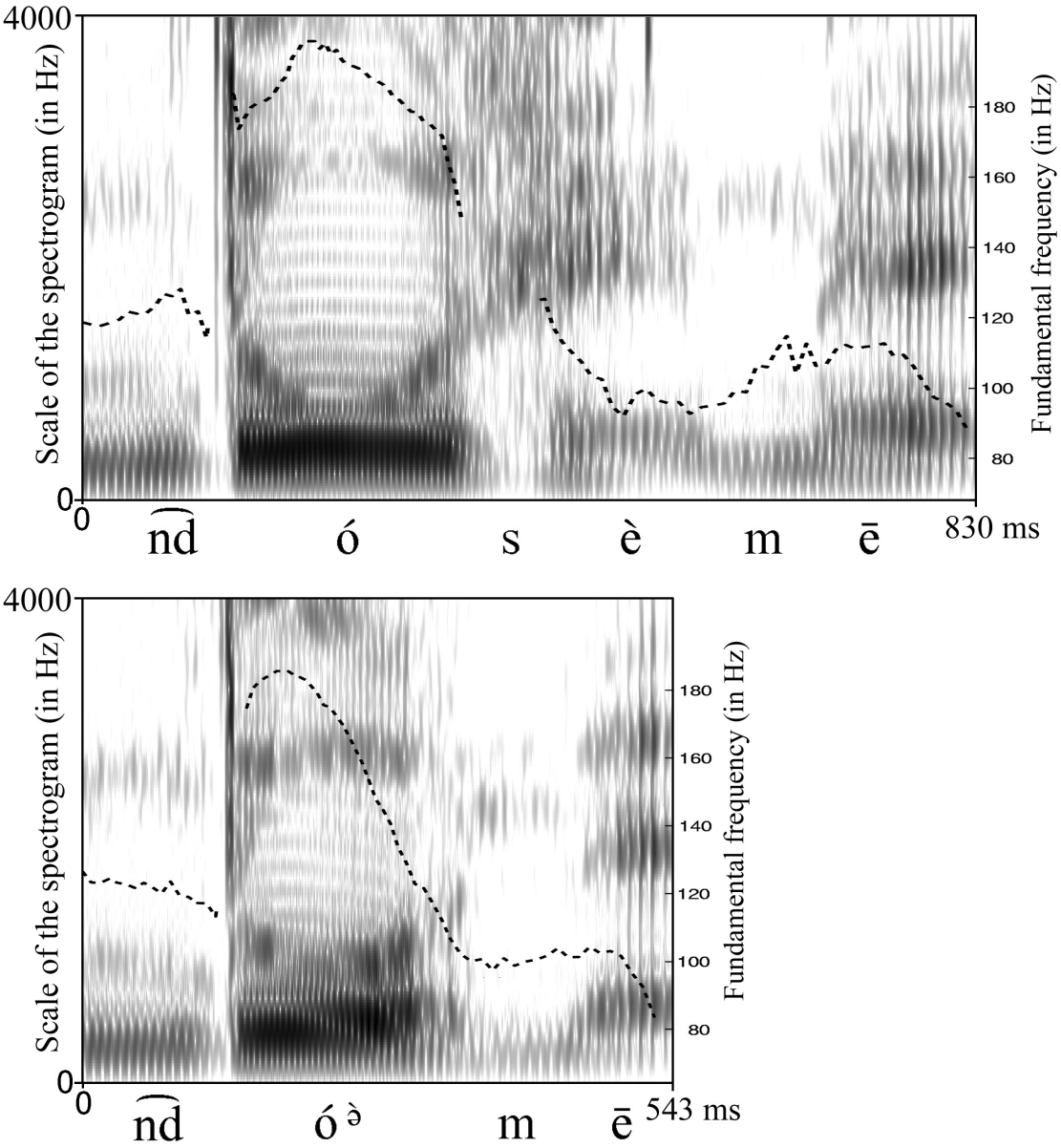
The effect of reduction of the following syllable on the target syllable is illustrated by figure 6, which shows the realisation of the verb /ndó/ 'to climb' in carrier sentences (11) and (12). [INSERT FIGURE 6 HERE] In both cases, F_0 decreases in the course of the vowel /o/; however, the decrease starts earlier and is greater when the syllable is in frame (12), i.e.

when the following syllable is reduced. As for formant frequencies, when /sè/ is fully realised (top of figure 6), the only formant movement observed is a rise in F2, which is an effect of the following alveolar consonant (a similar movement is observable at the beginning of the vowel, after the alveolar initial /n͡d/), whereas in the case when /sè/ has undergone reduction and coalescence with /ndó/ (bottom of figure 6), F1 increases from about 380 Hz to about 520 Hz in the course of the vowel, and F2 from about 850 Hz to about 1480 Hz, approaching the theoretical locus of schwa, namely F1≈ 500, F2≈ 1,500, F3≈ 2,500. For refined measurements, theoretical formant values of schwa may be calculated for a given speaker; a study of tongue movement for reduced vowels in English concludes that schwa has a target position, which corresponds to the mean of the positions for all the full vowels of the system (Browman & Goldstein 1992, Barry 1992).

Figure 7 provides average F_0 curves for the H tone, also plotting the other two tones as they appear in (11), taken as reference. The standard deviation is shown only for the two realisations of the High tones, which the figure aims to contrast. [INSERT FIGURES 7, 8 and 9 HERE] From figure 7, it is clear that tone H becomes falling when followed by a reduced L-tone syllable. By contrast, the results for tones M and L (in figures 8 and 9) do not show any clear change in F_0; this small-scale experiment alone is not a sufficient basis to settle the issue whether the phonetic output of such /LL/ and /ML/ sequences are identical with /L/ and /M/, respectively. It appears reasonable to think that the underlying additional L tone does have some effect on F_0 but that it is less salient than the effect on vowel quality.

3.5. Production data on Mid-tone syllable reduction

The production task designed to investigate the effects of Mid-tone syllable reduction was based on the following carrier sentence:


4000
Scale of the spectrogram (in Hz)
0
0
nd͡
ó
s
è
m
ē
830 ms
180
160
140
120
100
80
Fundamental frequency (in Hz)
4000
Scale of the spectrogram (in Hz)
0
0
nd͡
ó ə̀
m
ē
543 ms
180
160
140
120
100
80
Fundamental frequency (in Hz)

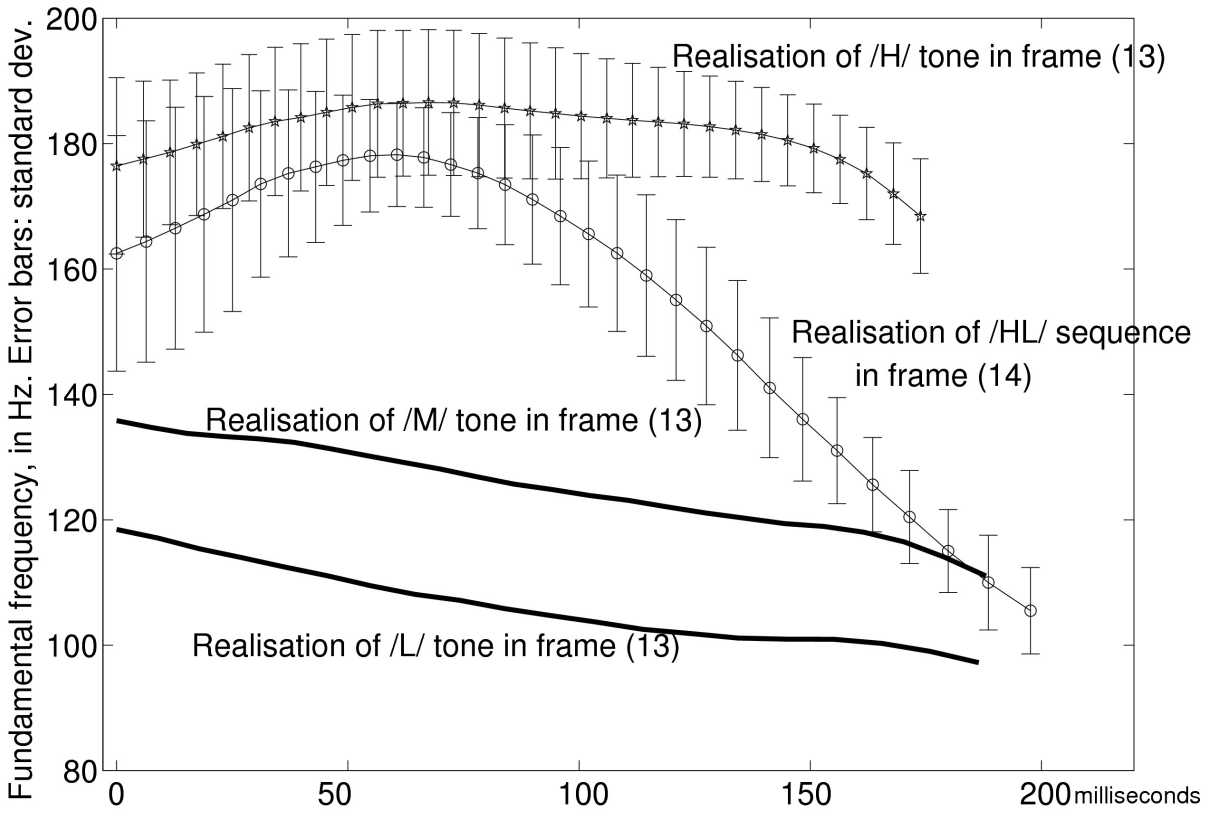
(13) ʈʂʰɯ̄́ sé ‖ tʰɯ̄ gjà *(slot A)* *(slot B)*

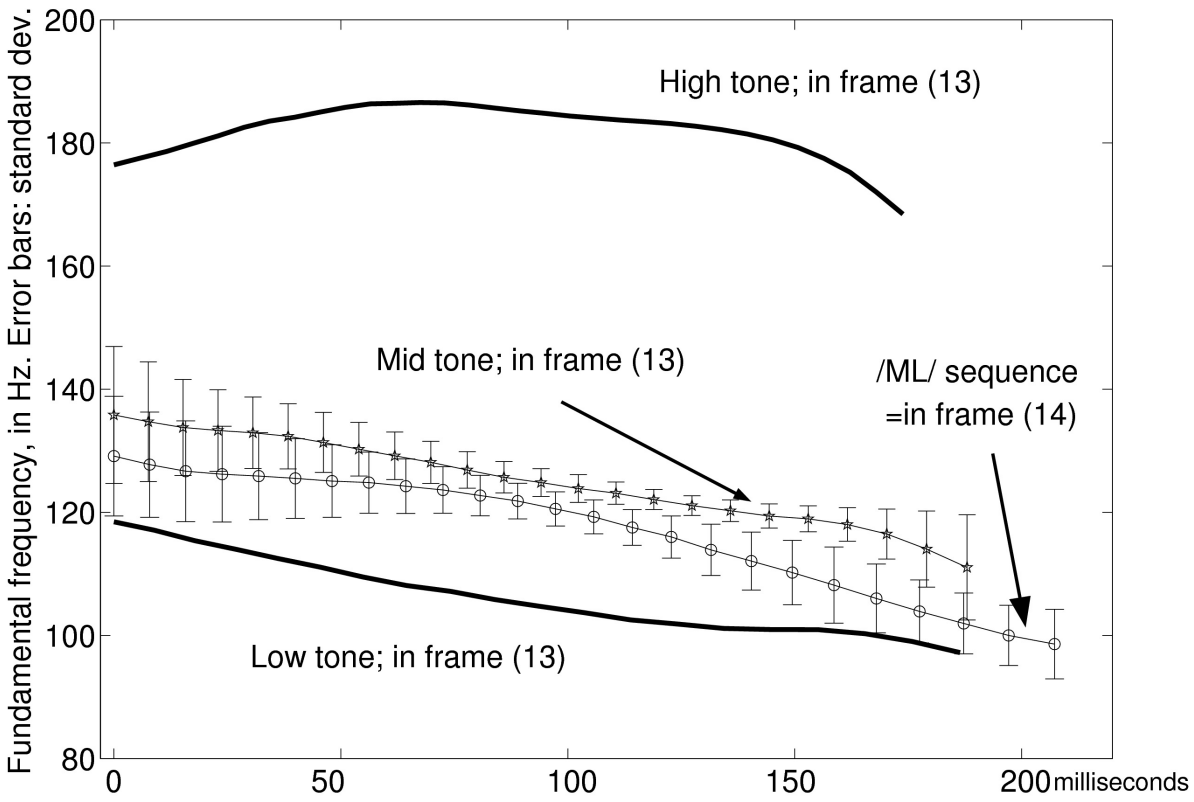
DEICT. COND./TOP. DEICT. very (adjective) (final particle)

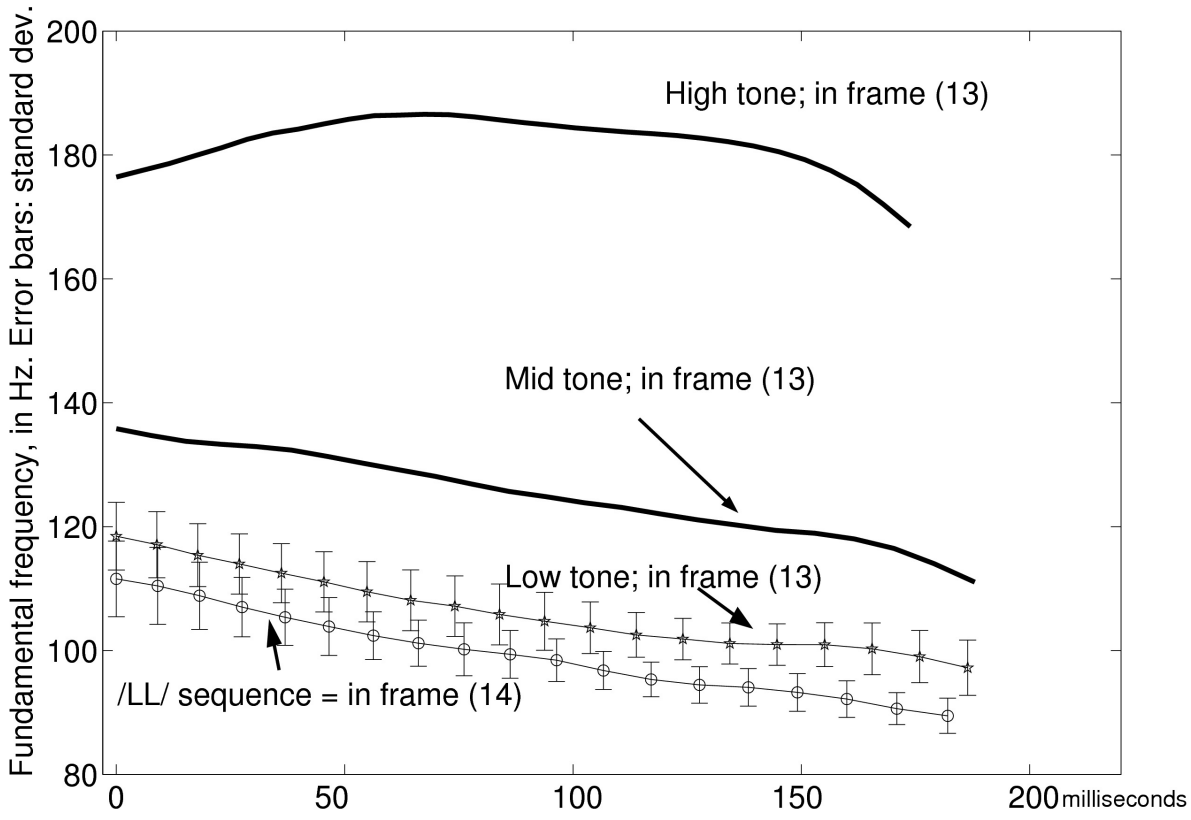
'This one is more *(slot A)* than that one, *(slot B)*'

Slot A was filled by an adjective, and slot B by a final particle. Twenty-five adjectives were selected: see the Appendix for the list of items and an internet link to the original recordings. Sixteen particles were used, eight of which carry an accompanying floating Mid tone whereas the other seven do not have a floating tone. All the combinations were recorded (384 tokens in all; each token was said twice), in three separate recording sessions. The adjectives were arranged in randomised order; the 25 were elicited in succession (their Chinese equivalent being used as a prompt) with the same final particle, which was indicated to the speaker in Naxi before the beginning of the series. The speaker (again, M4) was not aware that the experiment focused on syllable reduction.

Figure 10 presents F_0 curves averaged over the first 50 syllables in the series whose final particle is accompanied by a floating M tone. [INSERT FIGURES 10 AND 11 HERE] The results of F_0 measurements show that the /HM/ and /LM/ contour tones have a straightforward phonetic implementation. They also suggest that the realisation of a /MM/ tone sequence is identical to that of a simple M tone, in the same way as /HH/ sequences were shown to be identical to /H/. The slight difference in length can hardly be significant, in view of the variation in length across items; the overall tendency for the /LM/ sequences to be longer can be explained in light of physiological factors, raising F_0 requiring more time than lowering it, as was already mentioned in section 2.1. Figure 10 can be compared to figure 11, which shows the same adjectives, also in carrier frame (15), with one of the final particles that is not accompanied by a floating M tone. In figure 11, the Low tone is phonetically falling, contrasting with its rise in figure 10. The High tone is falling in both


Fundamental frequency, in Hz. Error bars: standard dev.
200
180
160
140
120
100
80
0
50
100
150
200
milliseconds
Realisation of /H/ tone in frame (13)
Realisation of /HL/ sequence
in frame (14)
Realisation of /M/ tone in frame (13)
Realisation of /L/ tone in frame (13)


Fundamental frequency, in Hz. Error bars: standard dev.
200
180
160
140
120
100
80
0
50
100
150
200
milliseconds
High tone; in frame (13)
Mid tone; in frame (13)
/ML/ sequence
=in frame (14)
Low tone; in frame (13)


Fundamental frequency, in Hz. Error bars: standard dev.
200
180
160
140
120
100
80
0
50
100
150
200
milliseconds
High tone; in frame (13)
Mid tone; in frame (13)
Low tone; in frame (13)
/LL/ sequence = in frame (14)

cases, but the final downward tilt in figure 11 differs clearly from the steady fall in figure 10.

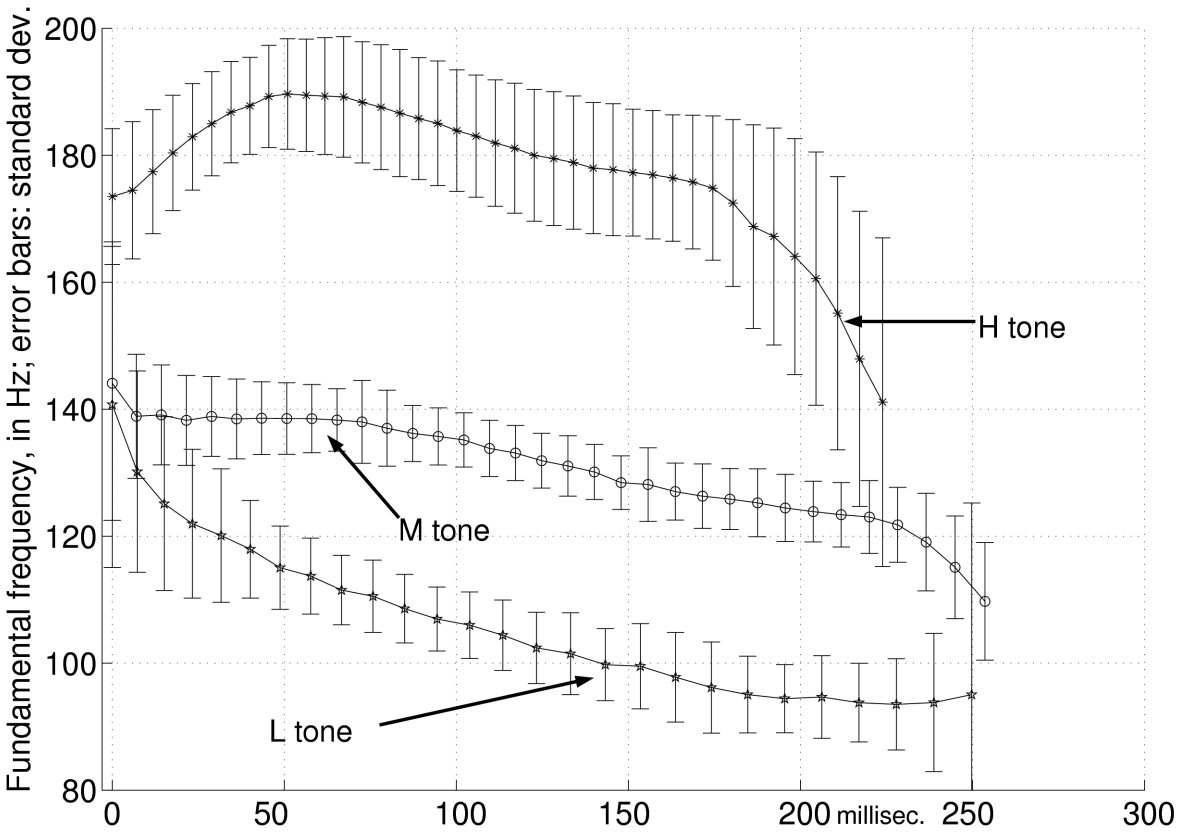
An issue which remains open: the difference among falling contours

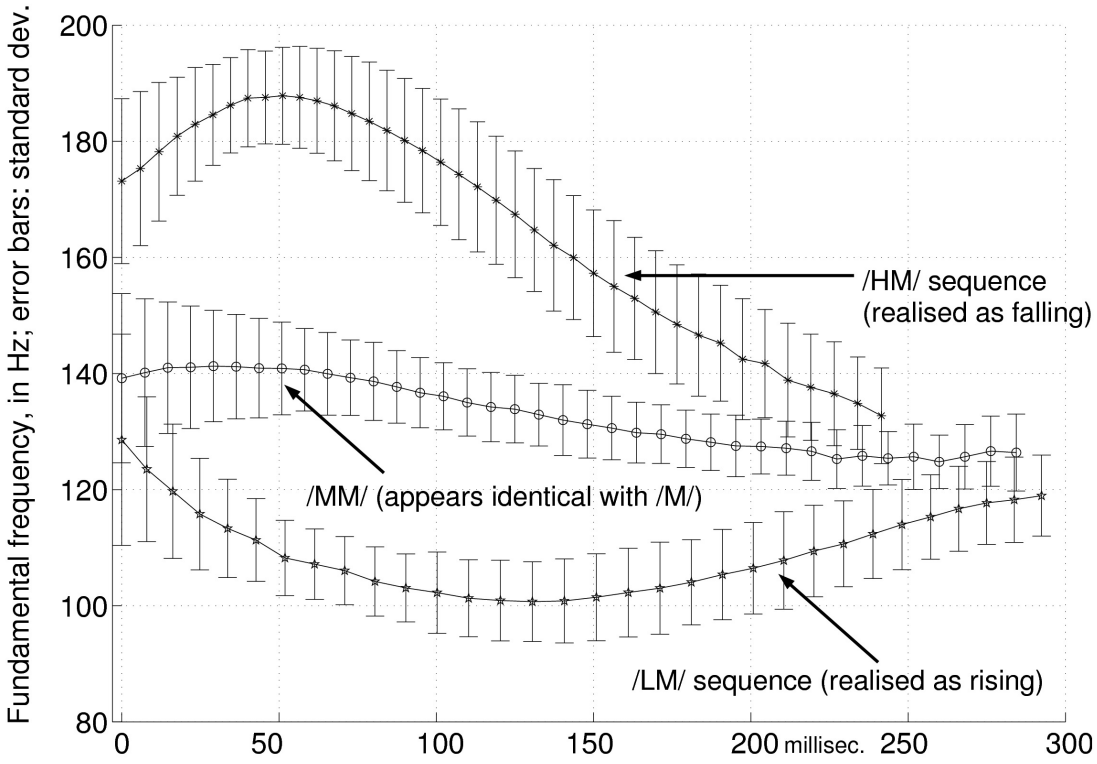
Experiment 1 gave some experimental evidence for the difference between /LH/ and /MH/ tone contours. One may wonder whether /HL/ tone contours are likewise categorically different from /HM/ contours. Comparison of the production data in figure 7 and figure 10 suggests so. Straightforward experimental evidence could not be obtained because no plausible 'minimal pairs' of utterances could be found to contrast these contours. The AS and NL language consultants were asked whether the /lá/ + /sè/ ('strike' plus perfect aspect) and /lá/ + /jɤ̄/ ('strike' plus particle indicating direct observation) sequences 'sound exactly the same or not' after reduction of the second syllable; their intuition is that the contour is not the same, and impressionistic listening likewise suggests that there is a difference, though it may not be realised and perceived in all cases, its functional load being very light.

4. Conclusions

4.1. Conclusions concerning tonal contours in Naxi

The experiments can be taken to settle the first issue raised in the introduction, that of rising contours in Naxi: the /LH/ contour and the /MH/ contour are actually distinct. Sequences made up of a Mid or Low lexical tone followed by a High tone (either delinked or floating) are realised as different dynamic F_0 shapes which transparently show the tautosyllabic realisation of the High target after the Low or Mid tone. There is no separate realisation of the floating High after a lexical High.


Fundamental frequency, in Hz; error bars: standard dev.
200
180
160
140
120
100
80
0
50
100
150
200
millisec.
250
300
H tone
M tone
L tone


Fundamental frequency, in Hz; error bars: standard dev.
200
180
160
140
120
100
80
0
50
100
150
200
millisec.
250
300
/HM/ sequence
(realised as falling)
/MM/ (appears identical with /M/)
/LM/ sequence (realised as rising)

Mid- and Low-tone morphemes can reduce to [ə̄] and [ə̀], respectively, and coalesce with the preceding syllable. Said differently, sequences made up of a High, Mid or Low tone followed by a Mid or Low can also be realised on a single syllable, as before, but this syllable acquires a schwa offglide. Thus, in Naxi, syllable-reduction patterns are asymmetrical: High-tone syllable reduction leads to tonal delinking and reassociation, whereas Mid- or Low-tone syllable reduction stops short of complete tonal delinking. Possible explanations of these processes, in light of universal phonetic tendencies and of language-specific factors, are being actively pursued.

4.2. On the typological validity of the notion of floating tone

As mentioned in the Introduction, phenomena of tone reassociation are now commonly modelled in terms of association lines between a tonal tier and a syllabic tier. This analysis applies, beyond the languages for which it was originally devised (essentially Niger-Congo languages), to the lexical tones of many languages. The present results go to show that the notions of floating tone and tonal reassociation are also applicable to Naxi. The present research aims to contribute data for the comparison of Far Eastern and Subsaharan tone systems—a line of research which is only in its incipient stage now.

Appendix: audio and electroglottographic data, and software for electroglottographic analysis

1. Audio and electroglottographic data

Part of the recordings on which this research is based are available online from the address below:

http://halshs.archives-ouvertes.fr/hal-00144485

The audio and electroglottographic (EGG) signals were simultaneously recorded into stereo files (audio in one channel, EGG in the other); as this format may not be convenient for all users, each stereo file was split into two: one audio file (name ending in _AUD.wav), one EGG file (name ending in _EGG.wav). There is no EGG recording for examples 7a and 7b.

The file names are as follows: 'Naxi_tones_' + number of sentence in main text + AUD for audio, EGG for electroglottographic files. For instance, Naxi_tones_4_AUD.wav is the audio file for example 4.

The list of items recorded in carrier sentences (5), (6), (11) and (12) is as follows (in the order of recording):

number of item	phonemic transcription	English translation	Chinese translation
1	ndə̀ˇ	to call out loud	喊
2	ʈʂʰū	to read aloud	读
3	ndý	to graze (cattle)	放牧
4	ndzù	to sit	坐
5	ʈʂʰú	to insert, to pierce	插
6	ʂù	to look for	寻找
7	pʰý	to vomit	吐
8	ndū	to come	来
9	ndə́ˇ	to measure	量
10	ù	to grind	磨
11	mé	to teach	教
12	ndzū	to eat	吃
13	kʰá	to shoot (an arrow…)	射
14	ndə̄ˇ	to hit (the mark)	中（射中）
15	ndó	to fall	跌倒
16	ndù	to plow	犁

17	p^hū	to open	开
18	ts^hý	to compensate for	赔偿
19	ʂū	to get back, to collect back	收(收回来)
20	p^hù	to scamper off	逃跑
21	ndzú	to fell	(用斧子) 砍
22	$tʂ^h$ù	rapid; of knife: sharp	快
23	ū	good, skilled	好(技术好)
24	bē	to do	做，干
25	hū	to go *(past tense)*	去
26	t^há	apt, good	可以
27	ts^hù	to arrive	到达

Doublets were removed from the sound files. Item 17 is missing from the recordings in frame 5, items 24 to 26 from the recordings in frame 6, and items 8, 24 and 25 from the recordings in frames 11 and 12.

The list of items recorded in carrier sentence (13) is as follows (in the order of recording):

number of item	phonemic transcription	English translation	Chinese translation
1	lá	thick	厚
2	ndzā	dark	阴
3	pà	broad	宽
4	gɣ̄	good (e.g. good-hearted)	好
5	pɣ̀	dry	干
6	mí	ripe, cooked	熟
7	sī	poor	穷
8	mbì	thin, watery	稀
9	ʂ^{w}ɚ̀	full	满
10	də̄	tender, soft	嫩

11	ʂə̀	long	长
12	mú	old	老
13	mbū	bright	亮
14	ʈʂʰù	rapid; of knife: sharp	快
15	ʂʊ́	new	新
16	ɳɖʐʊ́	beautiful	美
17	ʂʊ̀	yellow	黄
18	bȳ	coarse	粗
19	ỳ	light (not heavy)	轻
20	tsʰā	hot	热
21	ndzà	annoying; *homophone:* colourful	讨厌（同音词：花）
22	hó	deep	深
23	ndò	stupid	愚蠢
24	mbē	shallow	浅
25	ŋgjɤ̀	difficult	困难

Due to restrictions of the vocabulary list available, this set is not phonemically or tonally balanced (10 L tones, 8 M tones, 7 H tones).

In order to illustrate the difference between final discourse particles with and without a floating M tone, two recordings are provided: (i) one with final particle /mē/—conveying *obviousness*—, which carries a floating Mid tone that attaches to the syllable to its left; (ii) one with final particle /sjɤ̄/—used when the speaker was an eye witness to what (s)he reports—, which does not have such a floating tone. One single repetition is provided for each item.

2. Software for analysis of the electroglottographic signal

The software developed in collaboration with other researchers for the analysis of the electroglottographic signal (under the computing environment MATLAB) is available from http://voiceresearch.free.fr/egg/

Acknowledgments

Many thanks to David Bradley, François Dell, John Esling, Guillaume Jacques, Larry Hyman, Randy LaPolla, Martine Mazaudon, Boyd Michailovsky, David Mortensen, Annie Rialland, Jackson T.-S. Sun, Jacqueline Vaissière, and to Philip Rose and an anonymous reviewer for very helpful comments and suggestions. Many thanks to researchers in China for their friendly cooperation: Guo Dalie 郭大烈, Huang Xing 黄行, Kong Jiangping 孔江平, Xu Lin† 徐琳, and to the Naxi language consultants and friends, in particular He Xixian 和希贤 and Mu Ruhua 木汝华. The authors alone are responsible for remaining shortcomings.

Financial support was given by the *Fondation de France* and by *Université Paris 3*. Professional recording equipment was lent by the *Laboratoire de Phonétique et Phonologie* (CNRS-LPP) and the *Laboratoire Langues et Civilisations à Tradition Orale* (CNRS-LACITO).

References

BARRY, W. (1992). Comments on Browman and Goldstein: ‘ ‘Targetless’ schwa: an articulatory analysis’. In Docherty, G. J. & Ladd, R. (editors), *Papers in laboratory phonology II: Gesture, segment, prosody*, 65–67. Cambridge: Cambridge University Press.

BOERSMA, P. (2001). Praat, a system for doing phonetics by computer. *Glot International* **5**, 341–345.

BROWMAN, C. P. & GOLDSTEIN, L. (1992). ‘Targetless’ schwa: an articulatory analysis. In Docherty, G. J. & Ladd, R. (editors), *Papers in laboratory phonology II: Gesture, segment, prosody*, 26–56. Cambridge: Cambridge University Press.

FU MAOJI 傅懋勣(1984). 纳西语图画文字 ″白蝙蝠取经记″ 研究 *Naxi-zu Tuhua-wenzi ‘Bai bianfu qu jing ji’ yanjiu [A study of a Naxi pictographic manuscript, ‘White Bat's Search for Sacred Books’],* vol. 2. Tokyo: Computational Analyses of Asian and African Languages: Monograph Series 6.

GOLDSMITH, J. (1976). Autosegmental phonology. Cambridge, MA: M.I.T., published by Garland Publishing, New York, 1980.

GOOD, J. (2002). Tonal morphology in a creole: High-tone raising in Saramaccan serial verb constructions. In Booij, G. & van Marle, J. (editors), *Yearbook of Morphology 2002*, 105–134. Dordrecht: Kluwer.

HE JIREN 和即仁 & JIANG ZHUYI 姜竹仪 (1985). 纳西语简志 *Naxiyu Jianzhi (A Brief Monography on the Naxi Language)*. Beijing: 民族出版社 Minzu Chubanshe.

HE ZHIWU 和志武 (1987). 纳西语基础语法 *Naxiyu Jichu Yufa (A Basic Grammar of Naxi)*. Kunming: Yunnan Minority Publishing House (云南民族出版社).

HENRICH, N., D'ALESSANDRO, C., CASTELLENGO, M. & DOVAL, B. (2004). On the use of the derivative of electroglottographic signals for characterization of non-pathological voice phonation. *Journal of the Acoustical Society of America* **115**, 1321–1332.

MICHAILOVSKY, B. & MICHAUD, A. (2006). Syllabic inventory of a Western Naxi dialect, and correspondence with Joseph F. Rock's transcriptions. *Cahiers de linguistique - Asie Orientale* **35**, 3–21.

MICHAUD, A. (2004). Final consonants and glottalization: new perspectives from Hanoi Vietnamese. *Phonetica* **61**, 119–146.

MICHAUD, A. (2005). Prosodie de langues à tons (naxi et vietnamien), prosodie de l'anglais : éclairages croisés. Ph. D. dissertation, Paris: Université Paris 3–Sorbonne Nouvelle. Available online at: http://tel.archives-ouvertes.fr/tel-00130149

MICHAUD, A. (2006a). Tonal reassociation and rising tonal contours in Naxi. *Linguistics of the Tibeto-Burman Area* **29**(1), 61–94.

MICHAUD, A. (2006b). Three extreme cases of neutralisation: nasality, retroflexion and lip-rounding in Naxi. *Cahiers de linguistique - Asie Orientale* **35**, 23–55.

OHALA, J. (1978). The Production of Tone. In Fromkin, V. A. (editor), *Tone: a linguistic survey*, 5–39. New York/San Francisco/London: Academic Press.

OHALA, J. & EWAN, W. G. (1973). Speed of pitch change. *84th Meeting of the Acoustical Society of America*: abstract published in the *Journal of the Acoustical Society of America* **53**(1):345.

PINSON, T.M. (1998). *Naxi-Chinese-English Glossary, with English and Chinese Indexes*. Dallas: The Summer Institute of Linguistics.

ROTHENBERG, M. (1992). A multichannel electroglottograph. *Journal of Voice* **6**, 36-43.

SUNDBERG, J. (1979). Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics* **7**, 71-79.