Replicating in Naxi (Tibeto-Burman) an experiment designed for Yorùbá: An approach to ‘prominence-sensitive prosody’ vs. ‘calculated prosody’
Alexis Michaud

To cite this version:
Alexis Michaud. Replicating in Naxi (Tibeto-Burman) an experiment designed for Yorùbá: An approach to ‘prominence-sensitive prosody’ vs. ‘calculated prosody’. 2006, pp.819-822. hal-00130341

HAL Id: hal-00130341
https://hal.archives-ouvertes.fr/hal-00130341
Submitted on 11 Feb 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.
Replicating in Naxi (Tibeto-Burman) an Experiment Designed for Yorùbá: An Approach To ‘Prominence-Sensitive Prosody’ vs. ‘Calculated Prosody’

Alexis Michaud

Laboratoire Phonétique et Phonologie (UMR 7018) CNRS/ Paris 3 Sorbonne Nouvelle
alexis.michaud@univ-paris3.fr

Abstract

The starting-point of this study is the hypothesis (suggested by an overview of typologically varied languages) that it may be useful to characterise prosodic systems in terms of the degree to which they rely on the calculation of tone sequences. Each language could be placed at a certain position along a typological continuum between two types of prosodic organisation: (i) ‘calculated prosody’, in languages such as Yorùbá (in which tone serves complex morphophonological functions), whose prosodic structure hinges on the calculation of a tone sequence, by categorical processes such as the association of lexical tones and/or boundary tones, reassociation/‘tone floating’, and downstep; and (ii) ‘prominence-sensitive prosody’, typologically more common, found in languages such as Chinese, which have fewer elements of categorical tonal calculation, and in which intonation appears to reflect phrasing and informational structure in a largely noncategorical way. In an effort to test (and refine) this hypothesis, an experiment used for Yorùbá [8, 9] is adapted to Naxi, a Sino-Tibetan language which, like Yorùbá, has three lexical tones (High, Mid and Low), but which is hypothesised to be closer to the ‘prominence-sensitive prosody’ type, whereas Yorùbá would be closer to the ‘calculated prosody’ type. This pilot study on sentences in which all syllables bear the same tone does bring out differences between the two languages in terms of phenomena of phrasing and of prominence.

1. A typological proposal: ‘prominence-sensitive prosody’ vs. ‘calculated prosody’

1.1. Some definitions

As a preliminary to the proposed typological overview, some indications on the terms used may be useful. Prosody is here understood in its broadest sense: it includes lexically distinctive suprasegmentals, intonation, the expression of attitudes and emotions, and performance factors: rhythm and speech rate [15]. Lexically distinctive suprasegmentals may be, depending on the language, lexical stress, lexical tone, lexical pitch accent... Intonation is an abstract structure, not to be confused with the parameters whereby it manifests itself (in particular, it is a different notion from F0, which is a physical parameter, and pitch, which is a percept). Intonation can be conveniently divided into two components: (i) intonational phrasing, sometimes called syntactic intonation because of its strong (though not bijective) link to syntax, (ii) pragmatic intonation, determined by the information structure of the utterance. The term intonational group is used here loosely as a cover term for the various levels of intonational phrasing.

1.2. An overview of typologically varied tone languages

Typologically, there appears to be considerably more diversity among prosodic systems than labels such as ‘tone languages’, ‘pitch accent languages’ and ‘stress languages’ suggest [7, page 1]. ‘Tone languages’ do not make up a unified typological pole. An overview suggests that some prosodic systems, more than others, go by a calculation of tone sequences. The prosody of Mambila (Niger-Congo, Bantoid) seems to be highly constrained by its four level lexical tones. Ngamambó (Bantoid) [1] also appears highly constrained, as a five-term paradigm is open after a High-tone syllable: High, Mid, Low, downstepped Low (Ḷ), or flat Low (Ḷ̣̣̣̣̣). On the other hand, in Hausa, an Afro-Asiatic language which only has two tones, speakers apparently play to a larger extent on intonation to convey degrees of informational prominence (A. Rialland, p.c.). It thus seems that the degree of constraint on the margin of freedom left to the speaker depends on the number of level tones and of categorical processes of tone modification.

1.3. Formulation proposed

The proposal put forward here is that prosodic systems could be placed along a typological continuum between (i) prosodic systems that hinge on the calculation of a tone sequence (by categorical processes of association of lexical tones and/or boundary tones, reassociation/’tone floating’, and downstep), and (ii) prosodic systems in which phonetic realisation is sensitive to the prominence of individual items within the utterance, i.e. where intonation reflects informational structure (with such typical phenomena as contrastive emphasis). ‘Prominence-sensitive prosody’ can be illustrated by languages such as Mandarin Chinese. Despite having four lexical tones, Mandarin allows local intonational modifications which carry a communicative load comparable to the one they have in a language such as English (as has been known since [3]): the phonetic realisation of tones depends on the degree of prominence of individual syllables, and on intonational phrasing. Rather than unambiguously categorical processes such as downstep and tonal reassociation, Mandarin Chinese shows many noncategorical modifications of lexical tones, and few categorical modifications. In the study of Mandarin tones, much attention has been directed to the sandhi process whereby a sequence of two tone-3 syllables yields a tone 2+tone 3 sequence; on the other hand, the many cases of “minor sandhi” that have been proposed for Mandarin (e.g. by Chao Yuen-ren) are not clearly categorical, so much so that “It is not clear whether it is desirable or even possible to segregate tonal coarticulation [i.e. a noncategorical modification] from tone sandhi proper [i.e. a categorical modification]” [4, page 25].
It seems that ‘calculated-prosody’ systems would lend themselves especially well to the automatic synthesis of F0 curves, on the basis of tone sequences alone, whereas for ‘prominence-sensitive’ prosodic systems, other parameters have to be taken into account: for instance, a system of synthesis proposed for Mandarin assigns a coefficient of strength to each individual syllable [14].

Section 2 aims to provide some experimental evidence to help refine these tentative typological reflections.

2. A pilot study comparing two widely different prosodic systems (Naxi and Yorùbá) using a similar experimental setup

2.1. Overall presentation of the experiment

The pilot study reported below consists in adapting to Naxi, a Sino-Tibetan language, an experiment originally designed to investigate Yorùbá, a language of the Niger-Congo family [8, 9]. Both languages have High, Mid and Low lexical tones; unlike in Naxi, however, tone plays complex morphophonological functions in Yorùbá (see, e.g., the discussion by [2] of the subject marker as a floating High tone); and the Mid tone of Yorùbá can be considered as the absence of tone [6 and references therein], whereas the Mid tone of Naxi is a tone in its own right. In some respects, the prosodic system of Naxi (investigated first-hand [10, 11]) appears closer to that of neighbouring tone languages (such as Mandarin Chinese [3, 4, 14]) than to that of Yorùbá. In the terms proposed in section 1, Yorùbá is closer to the ‘calculated prosody’ type, whereas Naxi is closer to the ‘prominence-sensitive prosody’ type (though it possesses some elements of tonal calculation, namely the addition of a High tone for intensification, and some cases of ‘floating’ High tone [11]). Using a similar method for both languages shows the cross-speaker differences in declination over the speakers differences will not be discussed in detail; graph 1 as representative of the behaviour of the four speakers (cross-averaged (two on fig. 1, four on fig. 2), it did not appear useful to indicate standard deviation. These figures can be proposed as representative of the behaviour of the four speakers (cross-speakers differences will not be discussed in detail; graph 1 shows the cross-speaker differences in declination over the utterance).

On figure 1, the third syllable, which corresponds to an intonational juncture, is markedly longer than the seven others; this third syllable and the last syllable (the 8th) both have a more strongly decreasing F0 than the others (the decrease is on the order of 3 musical tones), and, contrary to all other syllables, have a decreasing Oq, indicating a tensing of the vocal folds, as opposed to the relaxing of the vocal folds at the end of the other six syllables. On figure 2, the longest syllable is, again, the one located at an intonational juncture (syllable 2). The curves in figure 2 do not show a strict layering as in figure 1, where the mean F0 of successive syllables decreased from the first to the last: in figure 2, syllables 4 and 7 have a higher mean F0 than the syllable that precedes them. The extent of the F0 rise also differs across syllables.

These piecemeal observations, by themselves, do not have demonstrative value; they nonetheless illustrate the general fact that, in Naxi, grammatical words are typically weaker (which does not appear to be equally true in Yorùbá), e.g. /mʊː/, ‘and’, the fourth syllable, is realised with flat rather than rising F0, in the continuity of the preceding syllable (its starting-point by and large corresponds with the F0 endpoint of the preceding syllable). The adverb ‘very’, /gíá/ (4a)
syllable), likewise a grammatical word, is also weak—weaker than the adjective /hɔ/ (syll. 7), which is realised slightly higher, and with a more clearly rising F0.

These facts have a bearing on the representation of F0 curves. In the study of Yorùbá that inspired the present experiment, F0 curves over individual syllables are simplified to two data points, “a point towards the beginning of the downward shift in F0 register, or a decrease in F0 in the course of the group is typically lengthened, with final lowering (a declination from one syllable to the next), and the last syllable.”

The junctures indicated below by the symbols [], and [[]] (realised by all five speakers): (i) /kodont o só sè lè n'yí kò lè /pè jè rì /pè ài (2) /pè ìlì | pè tì jì ní | kə ðu sú mè mò̀ n'dááùi/ (3) /pè ì / kə ò yè kò lè | tè bì mò̀ yè (4) /pè ài lè yì lè nàMD nà pè ñà gíà lá /hò/. (The presence or absence of the optional breaks [] apparently depends on tempo—a variable which was not controlled by explicit instructions given to the speakers.)

To sum up briefly the observations: in Naxi, all the syllables within an intonational group are realised at a similar height (modulo lexical tones, intrinsic F0, and a slight declination from one syllable to the next), and the last syllable of the group is typically lengthened, with final lowering (a downward shift in F0 register, or a decrease in F0 in the course of the syllable, or again a combination of both). The syllable which follows the intonational juncture (i.e. the first syllable of the next group) typically starts from a somewhat raised height (partial resetting of the declination line). For reasons of space, individual strategies in the realisation of intonational junctures cannot be set out in detail; to take one example, in the M9’s careful, deliberate style of delivery, group-final (and utterance-final) syllables stand out very clearly by a length about double that of the other syllables, whereas final lengthening is minimal.

In view of these observations, normalising the data for duration does not appear adequate for Naxi. Moreover, it does not appear enlightening to approximate declination over the utterance by comparing the average F0 of the first and last syllables (as was done by [8, 9] for Yorùbá), because this procedure overlooks the division of the utterance into intonational groups. The most apparent downward trend in Naxi appears to be due to group- and utterance-final lowering, rather than to (continuous) declination. With a view to comparison with the Yorùbá data, declination was nonetheless estimated as a ratio of the average F0 of the first and last syllable. The results are shown on graph 1: values on the order of -1.6 musical tone over the utterance for tone H and tone M, -1.75 musical tone for tone L. (The instruction that the speakers should take a deep breath between sentences may have contributed to this relatively steep declination.) F0 decrease in the course of the last syllable is more frequent than downward shift in F0 register. Graph 1 brings out the variability observed across speakers: in Naxi, even under controlled, ’labouratory’ conditions (uttering sentences that have been prepared beforehand), local modifications are of such amplitude that the declination line cannot be recovered from the observed F0 curve in a straightforward way.

3. General discussion and conclusion

Though very slender indeed (so much so that no statistical treatment was attempted), and open to some criticisms, the experimental results set out in section 2 reflect the considerable structural difference between the prosody of Naxi and of Yorùbá. The following differences emerge:

(i) Final lowering is observed in Yorùbá [8, page 17] as well as in Naxi; in Naxi, however, it is usually accompanied by lengthening.

(ii) The marking of intonational junctures (by lengthening and a final dip in F0) is very salient in Naxi prosody.

(iii) Grammatical words, as opposed to lexical words, are typically realised (all other things being equal) with slightly lower F0 in Naxi, whereas it seems to be a general observation that grammatical words are NOT realised as ‘weaker’ or with lower F0 than lexical words in Yorùbá (and in other languages of the Niger-Congo family; A. Rialland, p.c.).

(iv) Monosyllables elicited in a carrier phrase in Naxi are much longer (average over 600 syllables: 260 ms [10]) than would be expected for Yorùbá; I consider this lengthening as evidence of the prominence-sensitive realisation of Naxi words (a syllable placed inside a carrier sentence is in focal/prominent position), as opposed to the more calculated realisation of tone in Yorùbá, whereby F0 realisations do not usually depart considerably from what can be predicted on the basis of the tone sequence of the utterance.

Unlike in Yorùbá, where it appears possible, in read speech, to elicit sequences of like tones that come out as regular, gradually lowered F0 curves, it appears that in Naxi, under a similar experimental setup, at least two factors disrupt (to a certain extent) the regularity of the tone sequences: (i) junctures between intonational groups are clearly marked by lengthening and F0 decrease; and (ii) the informational importance (in the broad sense) of individual words influences their prosodic realisation. (These factors are even more salient in continuous ‘spontaneous’ speech.)

Concerning perspectives for future research, the typological distinction between ‘calculated prosody’ and ‘prominence-sensitive prosody’ could arguably be extended to languages that do not possess lexical tones: a language such as Wolof, which signals information structure by morphosyntactic means (a verb form may be ‘subject-emphatic’, ‘object-emphatic’ or ‘verb-emphatic’), appears to have a prosodic system very unlike ‘prominence-sensitive’ systems, despite not having any lexical tones [12].
4. Hyperlink
Some recordings (audio and EGG) are available at:

5. Acknowledgments
Many thanks to Larry Hyman and Annie Rialland for useful discussions, and to Jacqueline Vaissière, Cédric Patin and two anonymous reviewers for suggesting improvements. Any remaining errors are naturally my own responsibility.

6. References


