Voice quality and voice quality change in the Bai language of Yunnan Province

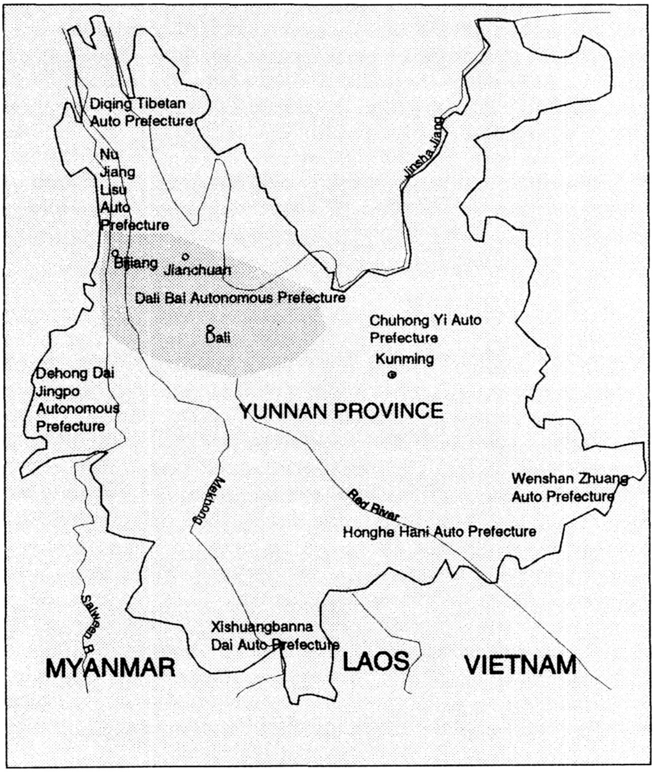
Jerold A. Edmondson
University of Texas at Arlington

Li Shaoni
Central University of Nationalities, Beijing

1. INTRODUCTION. The Bai are one of China's 55 minority nationalities and one of more than twenty in Yunnan Province.[1] Their total population amounts to 1.13 million divided into three vernacular areas, which are found in the extreme west of the province bordering on Burma. The first of these is centered on Dali 大理, where there live 500,000 Bai. The second area is found at Jianchuan 劍川 and the third is at Bijiang 碧江. Jianchuan is representative of the Central Vernacular with a population of 300,000 and Bijiang of the Northern Vernacular with 100,000. They are today called in Chinese *Bai* or *Baizu* 白族, but in the past they have been referred to as the 民家 *Minjia* or 'the minority' in opposition to the Han Chinese. Although the name *Minjia* is mentioned in Scott's *Gazetteer of Upper Burma and the Shan States* (1900), no one is today aware of their presence in Myanmar.

Classification of the Bai language is controversial. The majority opinion states that Bai belongs to the Tibeto-Burman branch of the Sino-Tibetan language family, cf. Matisoff (1991), Fu (1991), Bradley (1993), and Ma et al. (1994:17), though convictions differ about which subgroup it belongs to. There are also conjectures contra a TB affiliation altogether. Paul K. Benedict (p.c.) and recently Sergei Starostin (1994) have suggested that it might, in fact, be a daughter language of Ancient Chinese branching off from it before 200 BC. As Starostin (1994:1) says, "...it turns out that there are virtually no Tibeto-Burman elements in Bai other than those having their counterparts in Chinese, too." Crucial to his argument about the Chinese lineage of Bai are a small number of lexical items. These include: *ne*^{6} 'two'; *ɕi*^{6} 'four'; *ŋua*^{6} 'outside'; *sua*^{6} 'year'; *phia*^{6} 'lung'. By having these items in Tone 6, he reasons, Bai preserves a correspondence to the Chinese *yinru* category (i.e. syllables of the shape C_iVC_f, where C_i was *voiceless and C_f was a stop). Ancient Chinese possessed a final stop consonant (in his reconstruction) for these words that later (second century BC) changed to a fricative and which subsequently softened and merged with *-j* (4th c. AD). Contemporary Sinitic forms demonstrate no residues of the original consonant coda for these words.

[1] Support for some of this research was provided by a research grant from the Committee for Scholarly Communication with the People's Republic of China and the National Endowment for the Humanities to the first author for the period January-June 1990. The speaker analyzed here is Professor Li Shaoni of Central University of Nationalities, who first brought the problems of the Bai language to the attention of the first author. The first author is responsible for the analysis and writeup here.



MAP 1. Yunnan Province and the Bai Areas

In this paper we will not take a stand on the affiliation question, but note only that the voice quality properties we will describe below may in part be phonologically displaced vestiges of ancient final consonants and perhaps the study of their properties may shed light on this issue. These singular features may suggest for Bai a divergent history arising from an early separation either from Chinese or from TB. And they also suggest that Bai, far from being a bastardized mixture of Chinese and TB, may in fact retain the truest traits of the parental segment inventory in prosodic disguise.

On the descriptive side, notable research contributions to the study of Bai have been made by: Xu and Zhao (1964, 1984), François Dell (1981), Grace Wiersma (1990) and Starostin (1994). Also of relevance are the essays on voice quality in related languages by: Hu and Dai (1964), Ma Xueliang et al. (1981), Maddieson and Ladefoged (1985), and Chen Kang (1988). None of these has treated the richness of voice quality contrasts in Bai.

In our earlier work we have noted systematic voice quality contrasts in both Dali and Jianchuan Bai, cf. Edmondson and Li (1988, 1991). The Dali voice quality contrasts are not as remarkable as those found in Jianchuan. As we will outline below, Jianchuan Bai possesses not only a distinctive set of pitch contrasts, but also four different kinds of voice quality contrasts: (a) *modal voice*, (b) *breathy voice*, (c) *tense voice*, and (d) *harsh voice*. As an aid in this description, we have employed the notion of *setting* as in Laver (1980), according to which various muscle groups are adjusted or *set* to produce different voice-qualities. Jianchuan Bai has an inventory of such *settings* richer to our knowledge than for any other language of China, and equal in complexity (but different in organization) to that found in Chong, a Mon-Khmer language found on the Thai-Cambodian border (Huffman 1985, L-Thongkum 1991, and Edmondson forthcoming).

In Bai the settings can be divided physiologically into:
(1) *glottal stricture settings*—which result in three types of glottal vibration we will associate with: (a) *modal voice*, (b) *breathy voice*, and (c) *harsh voice*;
(2) *global settings*—which affect tensing of the speech apparatus from the diaphragm to the lips yielding two types of voice quality, (a) *tense* and (b) *lax* (Hu and Dai 1964);
(3) *supralaryngeal settings* —which have produced resonance voice-qualities we will associate with the features (a) *nasal* and (b) *oral*.

There are also four pitch trajectories in Bai, contrasting *high level*, *mid level*, *mid to high rising* and *mid to low falling*. Xu and Zhao (1984) distinguish three types of mid-low falling pitch to which they assign values 42, 21, and 31. We did not find differences this large in Jianchuan Bai, cf. Figures 3 and 5.

Indeed, for Jianchuan it is likely that the pitch contours for the mid to low falling tones are not the contrastive feature of the syllable.

In fact, only some but not all of these pitch and voice quality features are combined in actual Bai vocabulary. In particular [±nasal] can combine with both tense and lax voice, with breathy voice, and with harsh voice as well as with three of the four pitch trajectories, thus (3 values of pitch [high, mid, falling] X 2 values [±nasal] X 2 values [±tense]) + 2 [+harsh, ±nasal] +1 [rising pitch] = 15. Breathy voice quality and harsh voice quality combine only with the mid to low falling pitch trajectory. These combinations are illustrated for a minimal set in Figure 1 using the syllable /tɕi/.

Pitch Setting	Glottalic stricture setting	Glottalic + Overall setting	Glottalic + Resonance	Glottalic + Overall + Resonance
[high level] [mid rising]	[Modal] jil [tɕi55] 'much' jif [tɕi35] 'restless'	[Modal+ Tense] jib [tɕi5̟5̟] 'to mail'	[Modal + Nasal] jih [tɕĩ55] 'gold'	[Modal + Tense + Nasal] jik [tɕĩ5̟5̟] 'to perceive'
[mid level]	[Modal] ji [tɕi33] 'to pull'	[Modal + Tense] jiw [tɕi3̟3̟] 'leech'	[Modal + Nasal] jim [tɕĩ33] 'apricot'	[Modal + Tense + Nasal] jis [tɕĩ3̟3̟] 'naughty'
[mid falling]	[Harsh] jid [tɕi2͈1͈] 'flag'		[Harsh + Nasal] jit [tɕĩ2͈1͈] 'bracelet'	
[mid falling]	[Breathy] jiq [tɕi3̤1] 'earth'	[Breathy + Tense] jig [tɕi3̤1̟] 'to chase'	[Breathy + Nasal] jip [tɕĩ3̤1] 'alkali'	[Breathy + Tense + Nasal] jiy [tɕĩ3̤1̟] 'arrow'

Figure 1: Pitch and voice quality combinations in Jianchuan Bai

In order to stress the close connection of voice quality and pitch, we have indicated the voice-qualities with diacritics placed under the numerals of pitch contour, instead of writing them under the vowel as is the usual practice. Specifically, we have used two dots under a numeral to signify breathy voice, e.g. 3̤1, a plus sign for tense voice, e.g. 3̟3̟, and a line with two slashes through it, e.g. 2͈1͈ for harsh voice. Laver (1990) employs the plus sign for tenseness in this fashion; Chinese publications usually use a line below a vowel. The symbol for harsh voice is of our own invention.

We are still uncertain about the system of correspondences of Bai tones with traditional tone categories in TB or Chinese. The problem is complicated, as one must tease out originally inherited forms from possibly multiple levels of later borrowings. Nevertheless, we do give the correspondences found in Xu and Zhao (1984) and Starostin (1994):

Tone Cat (XZ)	*MC (Starostin)*	*Our description: pitch and voice quality*
1	yinshang	mid level modal voice
2	yangru	mid to low falling breathy tense voice
3	qu	mid to low falling breathy voice
4	yinping	high level modal voice
5		mid to high rising modal voice
6	yinru	mid level tense voice
7	yangping	mid to low falling harsh voice
8		high level tense voice

2. Pitch and voice quality. Pitch and voice quality in Jianchuan Bai are not independent of one another. Tenseness causes a higher pitch over most of the syllable, but it is particularly evident at syllable onset. It also results in an overall shortening of the duration of a syllable, presumably an effect from preemptory cessation of vibration.

A series of audio recordings was made of Li Shaoni's speech using a Sony TCM 5000 Professional Quality tape recorder, ATUS ATR-20 unidirectional microphone, and high quality audio tape. We later examined briefly two additional speakers from Jianchuan County, but do not include examples of their speech here. Since the Bai language demonstrates *voice quality sandhi* or change of voice quality in context, it was decided to examine syllables in isolation and not in a frame, for fear that the frame would influence the test item. Five to ten repetitions of each of the 15 constrastive syllables were recorded. The recordings were analyzed and the fundamental frequency extracted using the CECIL (JAARS International, Inc.) A-to-D hardware box and software running on a DOS laptop computer. At a later time, a locally developed compositing program was used to compute pitch means and to remove idiosyncratic features from five repetitions of a given syllable. The comparisons of syllables with the features [±tense] and [±nasal] are illustrated for the pitch trajectories in Figures 2-5 below.

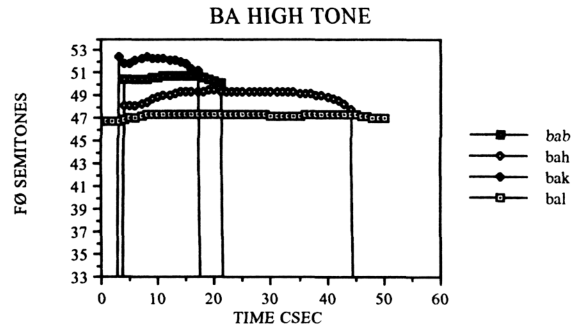


Figure 2: FØ of high level tones: *bal* [pa55] 'father', *bab* [pa5̣5̣] 'extortionist', *bah* [pã55] 'liter', and *bak* [pã5̣5̣] 'clam'

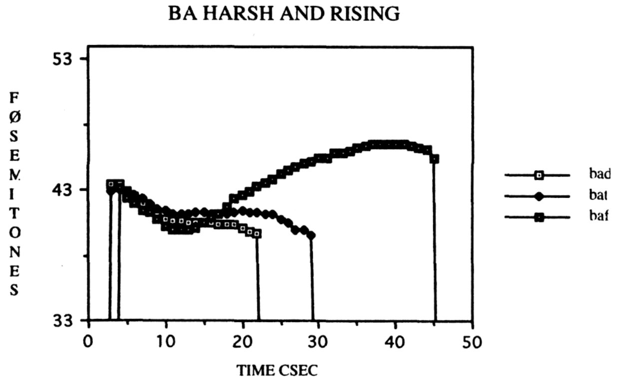


Figure 3: FØ of mid-low falling (harsh) and rising tones: *bad* [pa2̤1̤] no meaning, *bat* [pã2̤1̤] 'basin', and *baf* [pa35] 'eight'.

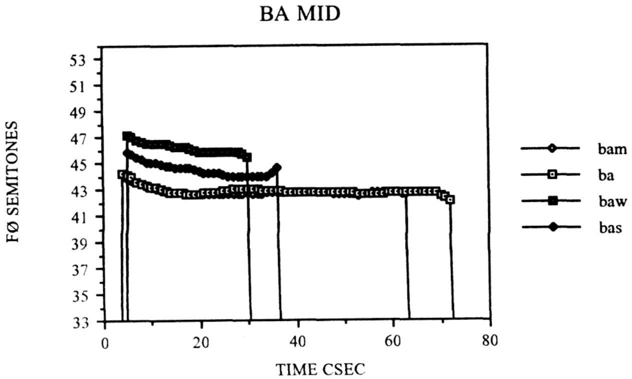


Figure 4: FØ of mid level tones: *ba* [pa33] 'foam', *bam* [pa3̣3̣] 'overrun, to', bas [pã33] 'help, to', and *baw* [pã3̣3̣] no meaning.

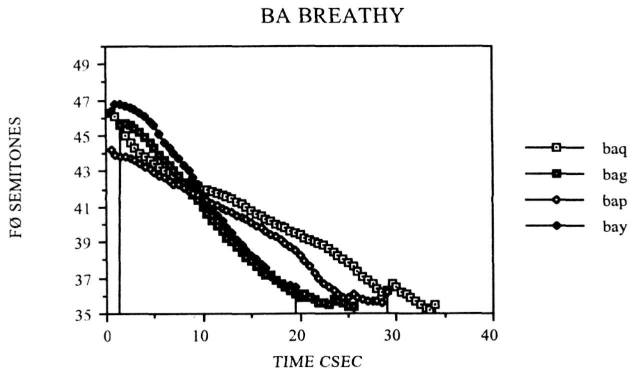


Figure 5: FØ of mid-low falling tones *baq* [pa3̤1] 'stir, to; fuss, to', *bag* [pa3̤ 1̣] 'milk', *bap* [pã3̤1] 'button', and *bay* [pã3̤1̣] 'trip someone with a rope, to'

We have indicated the tense syllables with filled symbols and the lax ones with open symbols in order to highlight the pitch-raising feature at the onset of tense syllables. Examining Figures 2, 4, and 5 one can see that there are consistent differences between syllables with the feature [+tense] and those with the feature [-tense]. The first obvious difference is that tense syllables demonstrate a fundamental frequency that is higher than the non-tense syllables. This difference is especially evident at the onset of voicing. Moreover, the duration of syllables of the tense set is always shorter than that of the non-tense syllables. These properties are found in all syllables with tense voice quality, including those with breathy and tense voice shown in Figure 5. Although there is not a strong auditory impression of tenseness in those tones that we characterized as breathy tense, the raised fundamental frequency and the shortened syllable length was always present.

In order to confirm these results in another way, the volume velocity of each syllable was measured. This procedure involved the Rothenburg Mask. The Rothenberg Mask is a plastic mask that fits over the nose and mouth and is pressed against the face by means of a handle protruding from the front. A sponge rubber base insures a tight fit without air leakage. The mask is fitted with two air pressure transducers or pneumotachographs—one for the oral airflow and one for the nasal airflow—connected to a filter-amplifier hardware unit. The hardware unit is, in turn, interfaced with the CECIL eight-bit analog-to-digital converter connected to a DOS computer, as shown in Figure 6.

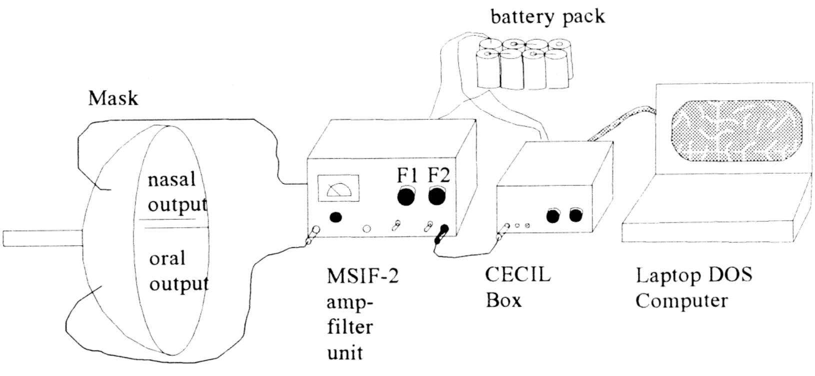


Figure 6: Equipment configuration for the inverse filtering experiments

3. RESULTS OF THE VOLUME VELOCITY MEASUREMENTS. As it is uncertain whether the setting [nasal] would have any influence on airflow, we undertook an experiment to discover a possible dependency. In order to test this hypothesis, we collected airflow data on the fifteen Bai syllables with the initials [p ph tɕ m s ts] as well as the syllable [tʏ̩].[2] The mean values for the airflow were determined by calculating the area under the curve divided by unit time. This information was entered in columns in MS-Excel®, column 1 for [-tense], column 2 for [+tense], column 3 for [+nasal, -tense] and column 4 for [+nasal, +tense]. The four sets were then analyzed pairwise for difference of means using the t-test. The following results were obtained:

Table 1: Results of significance test for tenseness and nasality

[-tns] vs. [+tns]	HIGHLY SIGNIFICANT	$.0005 < p \leq .005$
[-nsl -tns] vs. [+nsl -tns]	NOT SIGNIFICANT	$.1 < p \leq .375$
[+nsl -tns] vs. [+nsl +tns]	HIGHLY SIGNIFICANT	$.0005 < p \leq .005$
[-nsl +tns] vs. [+nsl +tns]	NOT SIGNIFICANT	$.1 < p \leq .375$

The influence of tenseness can be quite dramatic as these results show, whereas the nasal feature does not significantly change airflow or pitch values. In the level tones the pitch is raised so dramatically by tenseness that this feature cannot be ignored in transcription. If, for example, *bal* 'father' were assigned the value of 55, then its tensed counterpart *bab* 'extortionist' would have a value significantly higher than 5, say 6.5. At present there does not appear to be any adequate means of transcribing this affect other than in prose.

4. INVERSE FILTERING AND OTHER EXPERIMENTS ON VOICE QUALITY IN BAI AND SOME OTHER TIBETO-BURMAN LANGUAGES. In our previous work we relied on acoustic results to document voice quality differences. In particular we employed a procedure like that used by Bickley (1982), who showed that modal and breathy voice quality had different *spectral tilts*.

4.1. Spectral Tilt. Spectral Tilt or Harmonic Differential is a property one can extract from a waveform that compares the amount of intensity in the first and second harmonic of the acoustic signal. These are properties of the glottal wave and not of the filtering effect of the vocal tract. This measure revealed itself to be extraordinarily sensitive to voice quality differences and, in

2 The syllable [tʏ̩] involves the "bilabial vowel" [ʏ̩], which is the "motor boat" bilabial trill sometimes ending in [-u], although the vocalic ending is not obligatory. There are examples of real lexical items in all voice qualities and pitches.

fact, we found (Li and Ai 1990) differences of Harmonic Differential for all four voice qualities of Bai.[3] Three of these are seen in Figure 7 below.

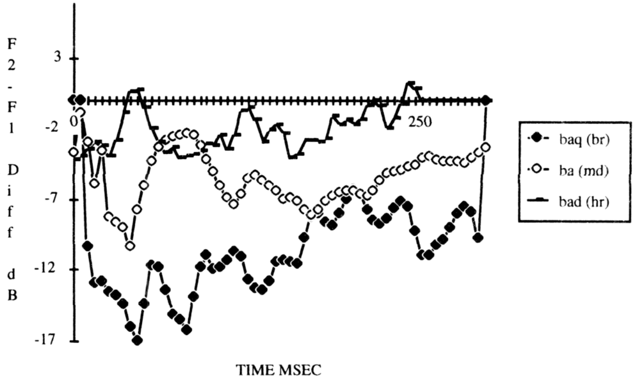


Figure 7: Harmonic differential for three kinds of voice quality in Jianchuan Bai of Yunnan Province

Figure 7 shows that the ΔI track for harsh voice quality crosses the ΔI track for modal voice only briefly, and the breathy voice quality track is even more distant on the lax side of modal voice. This result agrees with the other experiments conducted on the voice-qualities of Bai that suggest that harsh voice shows some affinities with tense voice.

We also carried out a set of informal experiments of *maximum phonation time* in which Li Shaoni held a syllable until his air supply was exhausted. These all showed that airflow (duration) varied as a function of voice quality. In particular, breathy voice syllables had the shortest duration lasting no more than a second; tense voice syllables lasted longer than the corresponding lax one, reflecting the greater impedance of the glottal folds to airflow. Harsh voice occupied a position between the other two.

---

3 Our experiment was somewhat different than Bickley (1982) in that we calculated the differential at every 15 msec point across the syllable. Those ΔI values for a given voice quality were found not to cross the values of any other voice quality, a feature we called *stratified*.

These measure were helpful in clarifying our understanding of voice quality but they all rely on secondary properties of the mechanisms of glottalic stricture. In order to measure voice quality using a technique that gets at the source of the voice quality differences in finer detail we began to use the technique of inverse filtering.

4.2. Inverse Filtering.

4.2.1. Methods. According to the Source-Filter Theory of Speech Production (cf. Fant 1961, Fant and Sonesson 1962, Lindqvist 1965), the air pressure wave exiting the mouth is said to be the coproduct of *glottal source wave*, *resonance filtering* of a tube-like vocal tract, and *radiation* out into the atmosphere. By using the Rothenberg Mask and the tunable filters in the MSIF-2 Hardware Unit in the inverse filtering mode (cf. Figure 6), we can cancel the influence of the vocal tract configuration and recover the glottal source wave. At the same time, the mask allows us to capture all the airflow discharge into the atmosphere; consequently there is no radiation effect to be considered. In order to remove these vocal tract components from the airpressure wave, the filters for the first and second formant frequency (F_1 and F_2) are adjusted as the subject produces a modal voice syllable until the visually most completely filtered profile is obtained, i.e., the resulting wave form is supposed to resemble the idealized model portrayed in Figure 8. One begins with the values of F_1 and F_2 obtained from a sound spectrogram, but precise tuning is still necessary as formant values are subject to considerable bandwide and frequency variation, and the slightest deviation from the resonance poles (formant frequency) leaves some of the resonance unremoved. A second problem is that even after the filters are set for the optimal F_1 and F_2, there is the added complication that the vowel formants in the test utterances might change over the course of the syllable; or even that there was some interaction between the formants and the voice quality so that the settings might be correct for modal voice quality but incorrect for others. Despite this potential for error, the results we obtained suggest that our simplifying assumptions—the language has pure or unchanging vowels and the voice quality-vowel quality interaction is minimal—were largely justified. Once the optimal formant values had been determined, we maintained these settings for all other voice quality types, so that comparison to the modal voice settings would be possible. In this sense modal voice was taken as the standard of comparison.

The net result of the capturing mask and properly tuned Hardware Unit interfaced as in Figure 6 is to display the vibratory mode of the original air pressure wave on the computer screen. The mask allows us to measure both the steady-state egressive airflow exiting the mouth (called the DC component) as well as the varying air pressure wave of glottal vibration (called the AC component). An ideal glottal air pressure wave for modal voice has a configuration illustrated in Figure 8 (Lindqvist 1965; Laver, 1980).

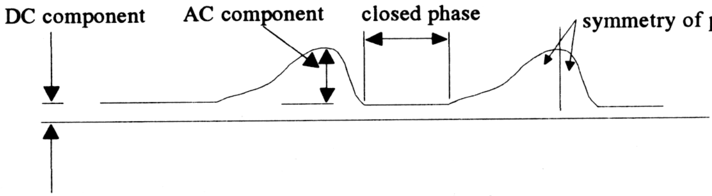


Figure 8: Hypothetical glottal pulses

As can be seen in this drawing, the glottal waveform has a dynamic opening phase, a dynamic closing phase, and a static closed phase. It is noteworthy that opening phase in modal voice is relatively slower than the more rapid closing phase. It is also clear that the pulse height off the baseline is composed of two parts: (1) a varying AC part, and (2) a steady-state DC part.

In order to show how effective the inverse filtering technique can be to enlighten instrumentally the auditory and articulatory descriptions of a tense/lax contrast in another TB language, we first illustrate the glottogram wave profiles in a language with this contrast. Figure 9 is a comparative plot of tense and lax voice syllables in Zaiwa (Atsi), a Tibeto-Burman language of Dehong Dai-Jingpo Autonomous Prefecture, Yunnan Province and Kachin State, Myanmar.

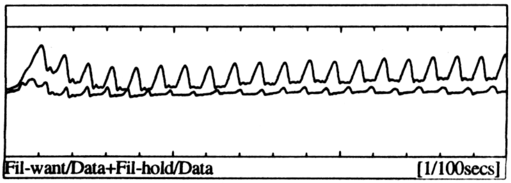
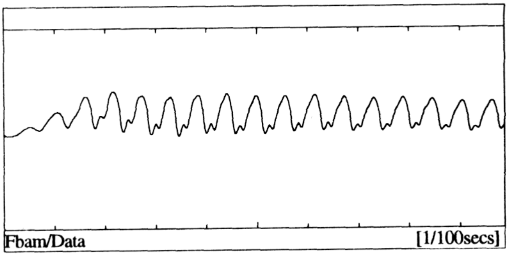


Figure 9: Zaiwa tense vs. lax voice quality contrast in *ta*^{55} 'to want, request' vs. *ta*$^{\underline{55}}$ 'to hold the hands of a walking infant'

Figure 9 shows two 150 msec long glottograms of the contrasting syllables, *ta*^{55} 'to want, request' vs. *ta*$^{\underline{55}}$ 'to hold the hands of a walking infant' with the beginnings of the two waves aligned and with the same filter settings for F_1 and F_2, so that a direct comparison of the onset of the syllables is possible. One sees directly the dramatic difference in DC and AC components as well as the different profiles of the glottal pressure wave. The upper track *ta*^{55} 'to want', has a much greater steady-state or DC airflow of a kind produced by laxer settings of the vocal folds in comparison with the tense voice quality of *ta*^{55} 'to hold the hands of a walking infant'. Note especially the height above the baseline, which is the level at which the wave enters from the left. Relatively high airflow corresponds to low active and passive tension of the vocal fold tissue during lax phonation. In the tense counterpart the wave is much closer to the baseline and also the wave has a much smaller amplitude in comparison to the lax phonation. This behavior suggests both lower values for the steady state as well as dynamic airflow, a property consistent with great passive tension. A second feature in Zaiwa is the relatively longer closed phase in the tense syllable. So, for example, a single period of *ta*^{55} 'want' lasts about 7 msec of which 4 msec are open and 3 msec closed. In *ta*^{55} 'to hold a baby's hands' the ratio of open to closed is reversed. The closed portion amounts to about 4 msec, whereas the open part lasts 3 msec. The Zaiwa tense vs. lax contrast seen here is not identical in nature to that found in Bai, but we do believe that it is representative of many other Tibeto-Burman languages.

4.2.2. Results. We now turn to glottograms for Bai syllables.



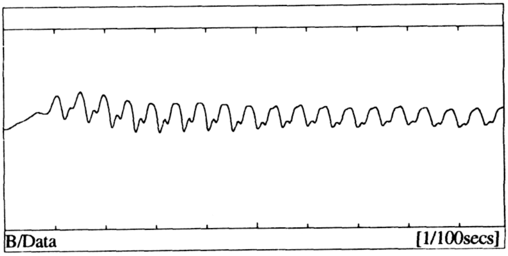


Figure 10: Glottograms (a) comparing *bam* 'overrun' and (b) *bas* 'help', mid level tones with modal nasal and tense voice quality.

In Figure 10a, the glottogram of a modal voice quality shows an asymmetric wave shape with a slower rise, a more rapid fall, and a moderate amount of airflow. There are also some unfiltered components during the closed phase that are especially prominent at the beginning of the syllable but become less obvious toward the end. Figure 10b shows the tense voice syllable [pa^{33}]. It is quite evident that the two waves in Figure 10 differ in amplitude. The tense voice quality wave is much closer to the baseline. These effects are due presumably to the greater vocal fold tension in syllables with tense voice quality, just as in the case of Zaiwa. It is also evident that the airflow is greatest at syllable onset. In fact, the two waves have nearly the same profile until the damping influence of vocal fold tension makes itself felt.

In harsh voice (Figure 11) the most striking feature of the glottal wave is the double-peaked structure in many of the periods.

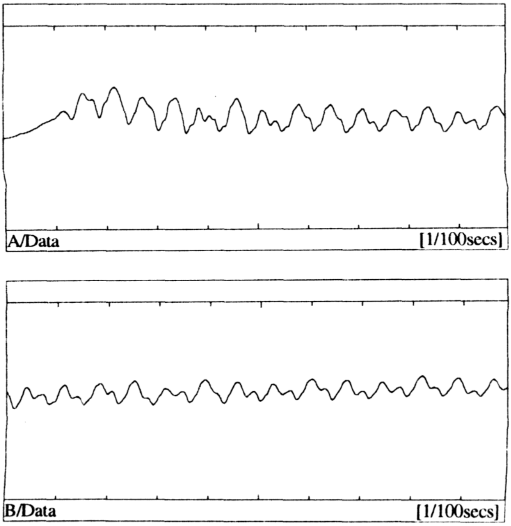


Figure 11: Glottogram for *bat* 'basin', a mid-low falling tone with harsh nasal voice

Note that the dual profile does not begin until about the sixth or eighth period after onset. On close inspection it can be seen that the second peak has a smaller amplitude and a slightly shorter wavelength, as evidenced by its moving a small increment to the left in each successive period. This circumstance suggests that two bodies are vibrating at slightly different frequencies. The stronger wave (from the true vocal folds, presumably) has a frequency as measured directly off the glottogram of about 128Hz and the weaker wave (perhaps, the false or ventricular folds) has a frequency of about 140Hz. This voice quality is accompanied by a throbbing series of pulses produced by the interference of the sound waves of these two slightly differing periods of vibration (beat frequency). That it begins after a few milliseconds after onset is apparently a delay caused by the ventricular folds requiring some time to descend. It also appears that this mode of phonation is rather inefficient, since progressively more air is required toward the end of the syllable. The curve is also located a considerable distance off the baseline, indicating relatively large amounts of airflow. Bai harsh voice shares many features that Rose reports for Zhenhai Chinese, cf. Rose (1989, 1990).

Breathy voice quality shows by far the highest level of steady-state airflow, with a very regular symmetrical pattern of alternating airflow.

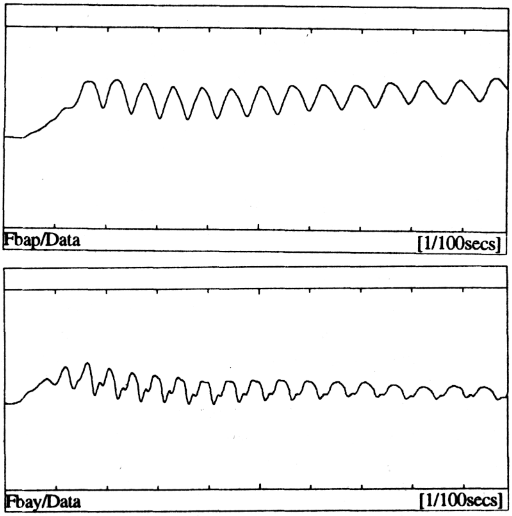


Figure 12: 12a Glottogram for *bap* 'button', a mid-low falling tone with breathy voice and 12b *bay* 'trip someone with a rope, to' with a mid-low falling with tense breathy voice.

In Figure 12a there is no evidence of unfiltered residues of resonance, nor a second peak in the wave. The frequency is about 116Hz and falling. There is also rising DC airflow over the course of the syllable as the wavelength of the AC component becomes longer (decreasing frequency). Just as in the case of tense voice in Figure 11, there is here heightened airflow at syllable onset. At the same time Figure 12 differs from Figures 10 and 11 by being further off the baseline, having a greater airflow throughout its course.

4.2.3. Discussion. The figures above show very evident glottographic differences that correspond well to what we believe to be the articulatory nature of the four voice qualities, and therefore constitute good evidence for these features. These figures also support the claim that breathy and tense qualities can combine in a single syllable. One may well ask if they can suggest how such a complex and compounded system might have developed. While much is still uncertain, one scenario for this prosodic compounding on a single syllable might be that breathiness is a result of displacing the contrast originally found in the initial consonant of the syllable, as is so widely attested in Mon Khmer and Kadai languages, whereas tenseness might be the leftover of a final stop consonant contrast, as has been suggested for some TB languages (see, e.g. Hu and Dai (1964)). Separate locations in the proto-language, i.e. syllable-initial and syllable-coda sources for breathy and tense voice quality respectively, would also provide a simple model of the history of Bai voice qualities; there would not need to be elaborate traffic rules about how prosodies were added from the same or adjacent sources in some environments and not in others. The feature [nasal] can have come from original final nasal consonants, of course. That would leave open only the source of the [harsh] voice quality.

Harsh voice might have evolved somewhat as the *hỏi* and *ngã* voice-qualities of modern northern Vietnamese. Thompson (1987:40) describes these as follows:

"hỏi is the 'dipping' tone;

[ngã] is accompanied by the rasping voice quality occasioned by tense glottal stricture. In careful speech such syllables are sometimes interrupted completely by a glottal stop (or rapid series of glottal stops)."

In both cases the feature [glottal stricture] is changing over the course of the syllable. The muscular effort of tensing and laxing the true glottal folds may have in time been extended to involve the ventricular folds located just superior to them, resulting in harsh voice. We would note that Dali Bai vocabulary corresponding to items with harsh voice in Jianchuan Bai have lower pitch than in Jianchuan but the mode of vibration of the one speaker we have studied suggests intervocalic constriction or closure not unlike that described for Vietnamese *hỏi* and *ngã* syllables. See Figure 13 below:

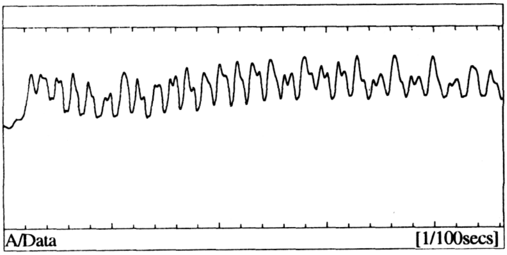


Figure 13: Dali Bai *bat* 'basin' with some ventricular voice

Moreover, the source for the Vietnamese *hỏi* and *ngã* tones in the theory of Haudricourt (1954) is a fricative coda. According to Starostin's theory mentioned at the beginning, Bai would have had a fricative coda for a part of its history in those syllables that eventually evolved into syllables with harsh voice quality.

5. TONE AND VOICE QUALITY SANDHI. Jianchuan Bai not only has multiple ways in which the vocal folds can oscillate, it also possesses changes of prosodies in context. The change of pitch trajectory in context has been of paramount interest to linguists in recent years, as it can inform us about the underlying form and the general processes of tonal change. Yip (1988) among others has characterized tone sandhi changes as belonging to two types: *edge effects* and *unit replacement*. Edge effects are changes in which the right or left edge of the trajectory is altered from the effect of surrounding pitch trajectories, whereas unit replacement involves the effacement of the pitch on a syllable and its replacement by that of a neighboring syllable. For example, Chinese generally shows edge effects as in the famous case of tone 3 → tone 2/___tone 3. So, *mai*^{323} *ma*^{323} 'buy horse' becomes *mai*^{35} *ma*^{323} which is homophonous with *mai*^{35} *ma*^{323} 'bury horse'. While several analyses of this sandhi process may be possible, one would be that the right edge of the first syllable HLH [323] becomes LH [35] while the left edge (underlined portion) is deleted, [<u>H</u>LH] -> [LH]. An example of unit replacement might be the case of Lhasa Tibetan, cf. Duanmu San (1993), in which 'bride' *na*^{53} + *ma*^{12} becomes *na*^{55}*ma*^{53} [HL-<u>LH</u>] → [H-L]. Duanmu argues that the tonality of the second syllable is effaced entirely and the tonality of the first syllable distributes over the larger disyllabic domain.

In the Jianchuan Bai language there are only a few changes in context described by Xu and Zhao (1984:7) as well as Li (1987:244). What is of interest is not only that there are changes of pitch trajectory, but that there are also changes of voice quality in Li Shaoni's speech. Basically, three environments cause tone changes. First, whenever the interrogative particle *mo*^{33} is added to the end of a verb to signal doubt, uncertainty, or interest in the matter, there is a change in the prosodies of the preceding syllable. All tones other than 55 become 35 and the first syllable exchanges its harsh, breathy, or tense quality for modal voice quality. Secondly, *jĩ*^{21} 'person' in compounds such as *ko*^{31} *jĩ*^{55}*po*^{55} 'wealthy' changes to *jĩ*^{55} and modal voice.

There is much about these changes in context that is unclear. However, it does seem clear that the changes occur irrespective of the pitch of surrounding syllables. It is likely that these mutations are not edge effects from the influence of coarticulation of polysyllables. Instead, they occur without regard for the pitch of the environment; or, in some cases, they depend on an item occurring in a particular construction, e.g. *xx*$^{(3)5}$*mo*^{33}, no matter what *xx* is. Moreover, the Jianchuan data also make clear how closely the voice quality is associated with the pitch. Whenever the pitch trajectory is altered, the voice quality is also replaced with modal voice (presumably the voice quality of the replacement pitch contour). Such data strongly support our choice of associating voice quality with pitch shapes.

In this paper we have stressed the complex and compounded nature of voice quality contrasts in Jianchuan Bai. At this time we cannot be sure whether these features are inherited or innovative. In any case, they should reawaken an interest in searching for these kinds of contrastive features in other TB languages.

REFERENCES

Bickley, Corine. 1982. "Acoustic analysis and perception of breathy vowels." *Working papers, Speech Communication Group, MIT* 1.71-81.

Bradley, David. 1993. "The subgroupings of Proto-Tibeto-Burman." Paper presented at the 26th ICSTLL. Osaka, Japan.

Chen Kang. 1988. "The tense tones of the Yi language." *MZYW* 1.18-26.

Dai Qingxia. 1990. "Guanyu wo guo zangmian yuzu xitong fenlei wenti." [On the problem of the systematic classification of the Tibeto-Burman languages of China] In *Zangmian yuzu yuyan yanjiu*. Kunming: Yunnan Minzu Chubanshe.

Dell, François. 1981. *La Langue Bai: Phonologie et Lexique*. Paris: Éditions de l'École des Hautes Études en Sciences Sociales.

Duanmu San, 1992. "An autosegmental analysis of tone in four Tibetan languages." *Linguistics of the Tibeto-Burman Area* 15.1.65-92.

Edmondson, Jerold A. forthcoming. "Voice quality in Chong." To appear in *Mon-Khmer Studies* 27.

Edmondson, Jerold A. and Li Shaoni. 1988. "Voice quality settings and pitch in the Bai language of Yunnan Province." (in Chinese) as Ai Jierui he Li Shaoni. 1989. Yunnan Jianchuan Baiyu yinzhi he yindiao anliang *Zhongyang Minzu Xueyuan Xuebao 70-74.*

_____ . 1991. "Voice quality and inverse filtering in the Bai language of Yunnan Province." In Dai Qingxia, ed., Proceedings of the International Conference on Yi-Burmese Languages, Xichang, Sichuan, Aug 1991.

_____ . 1960. *The Acoustic Theory of Speech Production*. The Hague: Mouton.

Fant, Gunnar. 1961. "A new anti-resonance circuit for inverse filtering." *Speech Transmission Laboratory Quarterly Progress and Status Report*-4/1961.1-6.

Fant, Gunnar and B. Sonesson. 1962. "Indirect studies of glottal cycles by synchronous inverse filtering and photoelectric glottography." *Speech Transmission Laboratory Quarterly Progress and Status Report*-4/ 1962.1-3.

Fu Maoji. 1991. *Zangmianyu Yuyin he Cihui*. Beijing: Zhongguo Shehui Kexue Chubanshe.

Hu Tan & Dai Qingxia. 1964. "Haniyu yuanyin de ssongjin." *ZGYW* 1.76-87. (Tenseness and laxness in Hani vowels.)

Huffman, Franklin. 1985. "The phonology of Chong." In Ratanakul, Thomas, and Premsrirat (eds.) *Southeast Asian Linguistic Studies Presented to André-G. Haudricourt*. Bangkok: Mahidol University.

Laver, John. 1980. *The Phonetic Description of Voice Quality.* Cambridge University Press.

Li Shaoni. 1992. "Lun Baiyu de shengmen hun he jicayin." [On the combination of glottal squeezing and friction in the Bai language.] *MZYW* 4.68-72.

Lindqvist, J. 1965. "Studies of the voice source by means of inverse filtering." *Speech Transmission Laboratory Quarterly Progress and Status Report-*2/1965.8-13.

L-Thongkum, Theraphan. 1992. "Chong register." In Davison, ed. *Shorto Festschrift.* London: School of African and Oriental Studies.

Matisoff, James A. 1991. "Sino-Tibetan linguistics: Present state and future prospects." *Annual Review of Anthropology* 20.469-504.

Ma Xueliang. 1951. *Sani Yiyu Yanjiu.* Beijing. (A study of the Sani Yi language.)

_____ et al. 1981. *Yuyanxue Gailun.* [An Introduction to Linguistics] Wuhan: Wuhan Gongxueyuan.

Ma Xueliang, Hu Tan, Dai Qingxia, Huang Bufan, and Fu Ailan. 1994. *Recent Contributions to Tibeto-Burman Studies.* Beijing: The Central University of Nationalities Press.

Maddieson, Ian & Peter Ladefoged. 1985. "*Tense* and *Lax* in four minority languages of China." *Journal of Phonetics* 13.433-54.

_____ & Susan Hess. 1986. "*Tense* and *Lax* revisited: more on phonation type and pitch in minority languges in China." *UCLA Working Papers in Phonetics* 63.103-9.

Scott, James George. 1990. *Gazetteer of Upper Burma and the Shan States.* Rangoon: Printed by the Superintendent, Government Printing. 5 Vols.

Rose, Phil J. 1989. "Yang tone phonation types in Zhenhai." *Cahiers de Linguistique Asie Orientale* 18.229-45.

_____ . 1990. "Acoustics and phonology of complex tone sandhi." *Phonetica* 47.1-35.

Starostin, Sergei. 1994. "The historical position of Bai." Paper presented at the 27th International Conference on Sino-Tibetan Languages and Linguistics, Paris, Oct. 11-15, 1994.

Wiersma, Grace Claire. 1990. *A Study of the Bai (Minjia) Language Along Historical Lines.* Ph.D. Dissertation, Universtiy of California, Berkeley.

Xu Lin & Zhao Yansun. 1964. "Baiyu Gaikuang." *ZGYW* 4.321-5. (Description of the Bai language.)

_____ . 1984. *Baiyu Jianzhi.* Beijing: Minorities Publishing House. (Sketch of the Bai language.)

Yip, Moira. 1988. "Contour tones." *Phonological Yearbook.* Cambridge: Cambridge University Press.