



HAL
open science

Voice onset contrast in Naxi

Thomas M. Pinson, Zhao Qinglian

► **To cite this version:**

Thomas M. Pinson, Zhao Qinglian. Voice onset contrast in Naxi. 34th International Conference on Sino-Tibetan Languages and Linguistics, Yunnan Institute for Nationalities, Kunming and Nankai University, Tianjin, Oct 2001, Kunming, China. halshs-02123969

HAL Id: halshs-02123969

<https://shs.hal.science/halshs-02123969>

Submitted on 9 May 2019

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.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

Voice Onset Contrast in Naxi*

Thomas M. Pinson and Zhao Qinglian

SIL International East Asia Group and Ethnic Language Guiding Committee of Yunnan

0. INTRODUCTION

It has been claimed that Naxi has a four-way voice onset time (VOT) contrast within the stops and affricates. For example, within the set of bilabials Naxi is said phonemically to have the four segments: /p^h/, /p/, /b/, and /mb/. The voiceless aspirated, voiceless, and voiced¹ distinctions in Naxi are unquestioned. The voiced and pre-nasalized voiced distinction is what this study examines. In this paper we present evidence indicating that the distinction between the voiced set (e.g. /b/) and the pre-nasalized voiced set (e.g. /mb/) does exist for some speakers of Naxi but does not exist for other speakers.

Naxi belongs to the Sino-Tibetan language family, the Tibeto-Burman language group, and the Lolo language branch. Ramsey (1987:250, 265) points out that there is some debate by Western scholars over whether Naxi should be in this group. The Naxi language is widely recognized to be Tibeto-Burman of the Burmic sub-branch, but that is where agreement ends. Where some Chinese scholars have listed Naxi in the Loloish branch, some Western scholars (e.g. DeLancy 1990) place Naxi by itself extra-Lolo-Burmese.

The Naxi language is divided into two major dialects: the Western, spoken in the Lijiang area, and the Eastern, spoken outside of the Yangtze loop. This paper examines the major Western dialect located around Lijiang's Old Town (*Dayanzhen*).

Many sources have claimed that Naxi has a four-way VOT contrast within the stops and affricates: Rock (1963), He and Jiang (1985), He Zhiwu (1987), and most recently Yang, Liu and Pu (1998). But several of our own observations led us to doubt the existence of the four-way VOT contrast.

First, in the orthography developed by Naxi scholars there is no accommodation of a four-way contrast. Naxi Pinyin has voiceless-aspirated segments <p>, voiceless and voiced <bb>. If there really were a four-way contrast, why would the fourth category not be included in the orthography? The answer is found in He and Jiang (1985) and He Zhiwu (1987). These sources state that Naxi scholars took the Lijiang Old Town dialect as the standard for developing the orthography, but speakers of the Lijiang Old Town dialect only make a three-way VOT contrast.

* We would like to express our appreciation to Stuart Milliken for his helpful input into the design of this experiment. We also would like to thank Keith Slater for his editorial feedback when this write-up was still rough. Finally, we are indebted to Norman Geary for his advice and help with the statistics and for his invaluable editorial input. All errors are our own responsibility.

¹ Here, by "voiced" we mean both purely voiced and pre-nasalized voiced. Elsewhere, we mean only the purely voiced set.

The second reason we doubted the four-way contrast was that in our own study we had only found a three-way division. Informally, we checked with speakers from several rural areas and found that they made no contrast between the putative pre-nasalized voiced set and the purely voiced set.

Yang, Liu and Pu (1998) did some acoustical studies on the stops and affricates of Naxi. They recorded speakers from many different areas, examining the results with their own computer program, “Visual Speech For Windows.” Their results showed that speakers from the Lijiang Old Town had voiceless-aspirated, voiceless, and pre-nasalized voiced stops and affricates, but did not have purely voiced stops and affricates. Their study also showed that speakers from outlying areas had the four-way contrast. In their write-up they did not indicate whether or not they had concealed from their test subjects what they were testing for. They may have run the test in such a way that the subjects were not aware of what they were testing, but their write-up gave no indication of this.

Recently, some Naxi scholars have been discussing the need to revise Naxi Pinyin to include the fourth category of stops and affricates. If the contrast exists, this revision is probably warranted. If it previously existed but is now disappearing, or never existed, there would be no need to revise Naxi Pinyin. We decided to run a small experiment either to support our doubts or to confirm what other scholars were saying.

Starting with the premise that in some areas there are still people who can both produce and perceive the four-way contrast, we developed two experiments: a production test and a perception test. We questioned several Naxi scholars to determine the best area(s) in Lijiang for testing. Two areas were chosen for their proximity to Lijiang Old Town and for the type of Naxi spoken there (i.e. for their reputed perpetuation of the distinction). These were *Lashi* Township in the next valley west of Lijiang Old Town and *Yanxi* Administrative Village of *Jinshan* Township south of Lijiang Old Town. For the production test we went to *Enzong* Village, *Meiquan* Administrative Village, *Lashi* Township. For the perception test we visited *Anzhong* Village in *Haidong* Administrative Village, *Lashi* Township and *Liu-yi* Village and *Liu-er* Village of *Yanxi* Administrative Village, *Jinshan* Township.

The subjects were not chosen randomly. Instead, we tried to restrict the subjects to those native to the test sites. We chose people from different areas of the test sites and from different families. We make no attempt to extrapolate the findings to the population as a whole. Determining the demographics would demand a larger scale survey. The goal of this experiment was to try to determine whether or not there were people who could still produce and perceive the putative contrast. Before actually running the experiments we ran a brief pilot test in *Baihua* Village.

1. PRODUCTION TEST

1.1. Production Test Design

Our expectation with the production test was that when a person who makes the VOT contrast articulates a putative pair, the VOT for the pre-nasalized voiced word would be measurably longer than for the purely voiced word. With this in mind we designed a test to elicit and record pairs of words.

1.1.1. Preparation and Administration

Using the published wordlist in He and Jiang (1985), which identified VOT contrasting pairs, we compiled a list of words with two minimal pairs showing the pre-nasalized voiced and voiced contrast (i.e. /ndɿ³³/ ‘to dig’ and /dɿ³³/ ‘turning-plow’; and /ndzɿ³³/ ‘leopard’ and /dzɿ³³/ ‘town, market’) and three pairs of words showing the voiceless-aspirated and voiceless contrast (i.e. /p^hə²¹/ ‘to untie’ and /pə²¹/ ‘to pull up [weeds]’; /t^hu²¹/ ‘to drink’ and /tu²¹/ ‘to heat [water]’; and /t^hɿ²¹/ ‘bucket’ and /tɿ²¹/ ‘to plant’). The list was comprised of the tokens (i.e. the voiced sounds), which were repeated ten times each, and the distracters (i.e. the voiceless sounds), which were repeated six times each. We then randomized this list of 76 words.

We chose the above ten words because they can be easily illustrated with photos. We collected photos for the ten words and tested them to see if they were sufficiently clear to elicit the correct responses. Following this, we made multiple copies of each photo and pasted the photos into a notebook in the random order we had previously determined. There was one photo per two-facing pages so that only one photo could be seen at any one time. In our test we had a separate notebook with copies of the ten photos. We used this to introduce the test. We explained that we were studying Naxi and would like to record the subject saying several words. We showed the subject the introductory notebook, asking the subject to identify the action or object represented in the photo. We elicited the ten test words without ever articulating them ourselves. We did not indicate what we were testing.

Once the subject had clarified the words represented by the photos, we recorded the subject saying the 76 words from the test notebook one at a time. We recorded the subjects using a Sony (TCD-D7) digital audio tape recorder, with a Shure (SM10A) unidirectional head-worn dynamic microphone.

Five Naxi speakers from *Enzong* Village, *Meiquan* Administrative Village, *Lashi* Township were recorded: two men aged 42 and 54, a woman aged 31, and two more men aged 44 and 37. The woman reported she was from *Enzong* Village. Later we found out that she was actually from *Baisha* Township and had married into *Enzong* Village. We recorded mostly men with the hope of getting speakers native to the test site, realizing that most women marry into a village from outside. The production test was not intended to discriminate men from women or old from young.

1.1.2. Analysis

Our null hypothesis for the production test was that there would be no difference in the duration of the VOT for the voiced and pre-nasalized voiced pairs. If there were a statistical difference in the duration of the VOT for any speaker, then we could be fairly certain this speaker did produce the four-way VOT contrast.

We used the SIL computer program, “Speech Analyzer,” to examine the raw waveforms and spectrograms for each of the recorded voiced utterances. We measured the VOT for each utterance, from the beginning of the utterance until the release of the stop or affricate. Using MS Excel, we calculated for each subject the mean VOT duration of each word (i.e. the average of the ten utterances for each word) and the standard deviation. We then calculated the t-test score for both paired-sets of utterances for each

subject, with a two-tailed test assuming two-sample unequal variance. We used a 1% level of significance.

1.2. Production Test Results and Discussion

We recorded four men and one woman. Of these five people only one man (i.e. the second man we recorded) consistently produced a contrast in the VOT of the pre-nasalized voiced and purely voiced stops and affricates.²

The t-test scores for the first man, 42 years old, were 0.3502 for /ndzɿ³³/ ‘leopard’ and /dzɿ³³/ ‘town’ and 0.2590 for /ndɿ³³/ ‘to dig’ and /dɿ³³/ ‘turning-plow.’ At a significance level of 0.01 (or even 0.05), the t-test scores from this man cannot reject the null hypothesis. Statistically, this man makes no difference in the VOT of pre-nasalized voiced and voiced stops and affricates.

The t-test scores for the second man, who was 54 years old, were well below the 0.01 significance level. His scores were 0.0012 for ‘leopard’ and ‘town’ and 0.000001 for ‘to dig’ and ‘turning-plow.’

The woman’s scores were 0.2125 for ‘leopard’ and ‘town’ and 0.6445 for ‘to dig’ and ‘turning-plow.’ As indicated above, this woman was actually from *Baisha* Township, but she originally said she was from *Enzong* Village, *Lashi* Township. We report her scores along with those of the other test subjects since we are not using the results to extrapolate to the population as a whole. She is an example of a young woman living in the countryside who does not make the four-way VOT contrast.

Our fourth test subject was a 44-year-old man. His t-test scores were 0.1664 for ‘leopard’ and ‘town’ and 0.6437 for ‘to dig’ and ‘turning-plow.’

The last person we recorded for the production test was a 37-year-old man. His t-test scores were 0.9557 for ‘leopard’ and ‘town’ and 0.1798 for ‘to dig’ and ‘turning-plow.’ Neither of his scores is low enough to reject the null hypothesis. He does not produce the four-way VOT contrast.

Without doing a larger scale survey and trying to draw a random sample of people we cannot extrapolate these findings to the population as a whole, but we do feel that they are somewhat indicative of the language situation. Only the 54-year-old man produced the VOT contrast between pre-nasalized voiced and purely voiced stops and affricates we tested. It is possible that if we had chosen different tokens, one of the other test subjects might have produced the VOT contrast, but based on this small experiment it looks like the younger generation of people are not maintaining the contrast.

2. PERCEPTION TEST

2.1. Perception Test Design

Our expectation for the perception test was that a person who can hear the difference between voiced and pre-nasalized voiced stops and affricates would be able to

² In this section we only report the t-test scores from each test subject. For the means and standard deviations of each test subject please see Appendix A.

distinguish an isolated word with one of these consonants from the corresponding word of the minimal pair.

2.1.1. Preparation and administration

For the perception test we prepared another notebook with photos. This notebook presented photos of all 76 randomized words in the same order we used for the production test. Each word was paired with the corresponding word of the minimal pair. For example, photos of /ndv³³/ ‘to dig’ and /dv³³/ ‘turning-plow’ were paired and photos of /p^hæ²¹/ ‘to untie’ and /pæ²¹/ ‘to pull up [weeds]’ were paired. Each pair of facing pages in the notebook displayed the photo of the test utterance and the photo of its “partner”. We randomly picked one photo from each pair to be pasted on the first page of the pair of pages; the other photo was pasted on the facing page.

We prepared one of the production test recordings for the perception test. The one we chose was a recording that consistently displayed the VOT contrast, the voice of the 54 year-old man.³ We used the shareware program “GoldWave” to edit the recording. Between utterances, we inserted approximately five seconds of silence, and we deleted any extraneous noise. From this edited recording we made a magnetic tape recording, which we used to administer the test.

We also prepared some score sheets. On the score sheets we recorded biographical information and the responses to each of the 76 words. There were four possible responses: “A” represented the first of the pair of photos; “B” represented the second of the pair; “C” represented an answer of “both”; and “D” represented an answer of “neither”. Using a prepared score sheet enabled us to mark the test subjects’ responses quickly and accurately.

When we introduced the test to a subject, we used the same introductory notebook that we had used for the production test. The method of eliciting the test word using those photos was the same for both tests. We did add a few unrelated photos to help explain how the subject was to answer. We explained that the subject would hear a word on the tape. The subject was then to point to one of the two photos to indicate what the word they just heard meant. If the subject felt that the meaning of the word was represented in both photos (i.e. homophones), they were to point to both photos. If the subject did not think that the recorded word was represented in either photo they were to indicate that to us. If the test subject wanted to hear an utterance twice, we played the utterance again. Again we concealed from the subjects what we were testing. We did not administer the test to subjects who had previously observed someone else taking the test, nor did we administer it to anyone who had taken or observed the production test.

We tested twenty-two people at the two locations. We tested eight men (ages 23, 28, 30, 32, 36, 61, 66 and 67) and three women (ages 27, 45 and 57) at *Lashi Township*.⁴

³ Our driver, who is from *Yanxi Village, Jinshan Township*, could easily discern the VOT contrast in the voiced sounds. He watched us record the production tests and told us that the man whose recording we eventually chose to use for the perception test was the only “standard” speaker we had recorded. By “standard” he meant this man could produce the contrast.

⁴ One man we tested at *Lashi Township* reported he was a native resident. Later we found out that he was a native resident of *Shigu Zhen* near the Yangtze River. We discuss his scores with the others.

At *Jinshan* Township we tested 11 people, but we didn't record correctly the responses of one man. Of the remaining ten people, seven were men (ages 30, 32, 35, 51, 55, 75, and 75) and three were women (ages 28, 50, and 50). All the people we tested were native residents of the test sites.

2.1.2. Analysis

Our null hypothesis for the perception test was that there was no perceptible difference between the voiced and pre-nasalized voiced pairs. In other words, test subjects would not be able to distinguish correctly between the voiced and pre-nasalized voiced pairs. They would then choose one or the other with equal probability 0.5, under the null hypothesis. Results from subjects who chose “both” or “neither” among their responses could not be analyzed according to this simple statistical model. For each subject we counted the number of “correct” responses out of twenty for each pair of words. A correct score of 15 or more (or 5 or less) was sufficient to reject the null hypothesis at the 5% significance level.⁵

2.2. Perception Test Results and Discussion

2.2.1. *Anzhong* Village Results

The first two subjects from *Anzhong* Village, *Lashi* Township were statistically unable to discern the difference between the voiced and pre-nasalized voiced pairs, but they responded to the distracters without error.⁶

Subjects 3 to 5 and 7 and 8 responded “both” for many of the words, and were clearly unable to distinguish between voiced and pre-nasalized voiced pairs.

The sixth *Anzhong* Village subject was a 61-year-old man. This man was able to discern the difference between ‘leopard’ and ‘town’ and between ‘dig’ and ‘plow’. He correctly picked ‘plow’ ten times, but only correctly picked ‘dig’ five times.⁷ He responded correctly to all of the distracters except for once incorrectly responding to ‘pull up [weeds]’.

The ninth to eleventh subjects were unable to distinguish the tokens consistently well.

In summary of the *Anzhong* Natural Village perception test, we tested eleven people; eight were men and three were women. Only one man (age 61) was able to discern the difference between ‘leopard’ and ‘town’ and between ‘plow’ and ‘to dig’. All three women (ages 57, 45 and 27) were unable to discern the difference between the tokens, though two of the women responded to the distracters without error.

⁵ For a table with the binomial probabilities for all possible correct responses, see Appendix B.

⁶ For a more complete presentation of the perception test data, see Appendix C for *Anzhong* Village data and Appendix D for *Yanxi* Village data.

⁷ We discuss a possible reason for not scoring better when responding to the word ‘dig’ below.

2.2.2. Yanxi Village Results

Five subjects at *Yanxi* Village, subjects 1, 4, 5, 8 and 9, were unable to distinguish between ‘leopard’ and ‘town’ or between ‘plow’ and ‘to dig’.

Our second *Yanxi* Village subject was a 35-year-old man. This man was the only person in *Yanxi* able to distinguish ‘leopard’ from ‘town’ and ‘plow’ from ‘to dig’. He responded to the distracters without error.

The third subject from *Yanxi* Village was a 75-year-old man. This man was able to discern the difference between ‘leopard’ and ‘town’, but was not able to discern ‘plow’ from ‘to dig’. He responded correctly to all of the distracters except ‘pull up [weeds]’ (one time).

Our sixth *Yanxi* Village subject was a 30-year-old man. He was able to distinguish ‘leopard’ from ‘town’, but was unable to distinguish ‘plow’ and ‘to dig’. He responded correctly to all of the distracters except ‘to drink’ (one time) and ‘bucket’ (one time).

Our seventh *Yanxi* Village subject discerned the difference between ‘leopard’ and ‘town’, but not ‘plow’ and ‘to dig’. He responded correctly to all of the distracters except ‘untie’ for which he chose “neither” one time.⁸

Our tenth *Yanxi* Village subject was a 50-year-old woman. She was able to distinguish ‘leopard’ from ‘town’, but not ‘plow’ from ‘to dig’. She responded to the distracters without error.

In summary, we tested seven men and three women in *Liu-yi* and *Liu-er* Villages of *Yanxi* Administrative Village, *Jinshan* Township. Four of the men and one of the women were able to perceive the difference between one of the pairs of tokens (‘leopard’ and ‘town’). Only one of these men could discern ‘plow’ from ‘to dig’.

2.2.3. Discussion

Of the twenty-one people tested in the two villages, five men (aged 61, 35, 75, 30 and 55) and one woman (aged 50) were able to distinguish the ‘leopard—town’ pair (/ndzɿ³³/ and /dzɿ³³/). Only two of the twenty-one subjects, two of the five men referred to just above (aged 61 and 35), were able to discern the difference between the ‘to dig—turning-plow’ pair (/ndv̥³³/ and /dv̥³³/).

There could be several reasons for this. One possible reason is that the distinction between the words or consonants is being lost. A second possible reason for this is the voicing—pre-nasalized voicing contrast is harder to discern on a magnetic tape for an alveolar stop followed by a syllabic labio-dental fricative (/ndv̥³³/ and /dv̥³³/) than for a retroflexed affricate followed by an apical vowel (/ndzɿ³³/ and /dzɿ³³/). We cannot offer an explanation for this but recognize it as a possibility. A third and related possible reason has to do with a mistake we made in administering the test. We failed to have a few non-test words as a lead into the perception test. The first word on the test was ‘to dig’

⁸ This was the only instance of a subject picking “neither”. This man was talking when this utterance was played. We played the utterance again for him, but he spoke again while the utterance was being played. It is our opinion that he never heard the word correctly.

/ndv³³/. Though several times we played the word more than once for subjects, only two people responded correctly to it. One of those two people, a 32-year-old man from *Anzhong* Village, only responded correctly to the word ‘to dig’ that one time. The other person, a 27-year-old woman from *Anzhong* Village, only responded correctly to the word ‘to dig’ three times. This of course only explains why people missed one occurrence of ‘to dig’, but was certainly a factor in the results we got.

3. CONCLUSION

Of the five people we recorded for the production test only one man (aged 54) produced the VOT contrast. Of the twenty-one people from two locations to whom we administered the perception test, only five men and one woman were able to discern the VOT difference between a voiced—pre-nasalized voiced pair.

These results are not surprising in view of the low functional load born by the voiced—pre-nasalized voiced contrast. In a list of more than 1000 Naxi words in He and Jiang (1985) there are only twenty-two minimal pairs displaying this contrast. Within these there are only nine pairs in which both members share the same part of speech (e.g. verb, noun, adjective, etc.). Within this set of nine, for many speakers the vowel or the tone is different for two pairs. Also in this set of nine, one pair could be considered minimal pairs in the area near Lijiang Old Town (i.e. for those people who still make the VOT contrast), but the vowels are different for speakers in a more remote area (i.e. *Tacheng* Township). One of the words in one pair is quite uncommon, and many young people do not know the word. Also, we have not been able to verify one word of yet another pair. Finally, at least three pairs have one word that usually is modified in some way. For example, the pair we used in our tests, ‘town’ and ‘leopard’, in common speech are easily distinguished by the locative suffix /lɔ²¹/ affixed to ‘town’. The word ‘town’ is rarely said without this suffix. Moreover, many people had trouble remembering the Naxi word for ‘leopard’, instead they quickly thought of the Chinese loan /pɑ⁵⁵/. This leaves only one pair from this list of nine, and one word of this pair is often reduplicated in common speech. Basically, for this short list from He and Jiang (1985), there is a very low functional load carried on the VOT contrast.

We have shown that there are some Naxi speakers who can still hear and produce the voiced—pre-nasalized voiced VOT contrast for some words. But for many speakers of Naxi the contrast is lost. The Naxi orthography (i.e. *Naxi Pinyin*) is one of the approved minority orthographies for trial implementation in China. It would seem to be an unnecessary price to pay to introduce a change to the orthography in order to represent a distinction that seems to be lost for many speakers. Currently, *Naxi Pinyin* is quite easy for Naxi people who are somewhat familiar with *Hanyu Pinyin* to learn. There would be a significant increase in learning difficulty if this largely historical distinction were represented in the orthography. It would be similar to many of the strange spellings we have in English, which represent archaic pronunciations.

We recognize that our study was very limited in scope and that a much larger survey would be needed to determine more widely the demographics of this loss of VOT contrast. It is possible that there are pockets of Naxi who are maintaining this subtle phonetic difference. But with the increase of public education comes the increased influence of Chinese, a language with only a two-way VOT contrast. And with the

improvement of communication lines comes increased inter-village interaction, adding further pressure to a highly marked phonetic distinction in previously isolated pockets. Finally, among our subjects it was largely the older male speakers, and (notably) none of the younger women, who were able to distinguish the VOT difference. Since language changes tend to follow the speech patterns of young women in a language community, it is therefore probable that Naxi is losing its four-way VOT contrast.

APPENDICES

Appendix A

Production test means, standard deviations and t-test scores for the voiced and pre-nasalized voiced pairs for each subject are presented below.

<i>Enzong</i> subject #1: 42-year-old man, with junior high education				
VOT	/ndv ³³ / 'to dig'	/dv ³³ / 'plow'	/ndzɿ ³³ / 'leopard'	/dzɿ ³³ / 'town'
Mean	571 ms	526 ms	436 ms	367 ms
Standard Dev.	87 ms	91 ms	173 ms	148
T-test score	0.2590		0.3502	

<i>Enzong</i> subject #2: 54-year-old man, with 2 nd year high school education				
VOT	/ndv ³³ / 'to dig'	/dv ³³ / 'plow'	/ndzɿ ³³ / 'leopard'	/dzɿ ³³ / 'town'
Mean	1097 ms	118 ms	1500 ms	1043 ms
Standard Dev.	208 ms	22 ms	277 ms	252 ms
T-test score	0.0000001		0.0012	

<i>Enzong</i> subject #3: 31-year-old woman, with elementary education				
VOT	/ndv ³³ / 'to dig'	/dv ³³ / 'plow'	/ndzɿ ³³ / 'leopard'	/dzɿ ³³ / 'town'
Mean	1007 ms	966 ms	660 ms	589 ms
Standard Dev.	229 ms	158 ms	140 ms	100 ms
T-test score	0.6445		0.2125	

<i>Enzong</i> subject #4: 44-year-old man, with elementary education				
VOT	/ndv ³³ / 'to dig'	/dv ³³ / 'plow'	/ndzɿ ³³ / 'leopard'	/dzɿ ³³ / 'town'
Mean	896 ms	869 ms	589 ms	689 ms
Standard Dev.	120 ms	133 ms	109 ms	188 ms
T-test score	0.6437		0.1664	

<i>Enzong</i> subject #5: 37-year-old man, with 4 th grade elementary education				
VOT	/ndv ³³ / 'to dig'	/dv ³³ / 'plow'	/ndzɿ ³³ / 'leopard'	/dzɿ ³³ / 'town'
Mean	946 ms	812 ms	648 ms	651 ms
Standard Dev.	172 ms	258 ms	132 ms	126 ms
T-test score	0.1798		0.9557	

Appendix B

The following table is the binomial distribution for the number of correct responses out of 20. Under the null hypothesis that a correct answer has probability 0.5, the sum of the probabilities of responding correctly only 0 to 5 times out of 20 is 0.0207. This is the same as the total probability under the null hypothesis of responding correctly 15 times or more (out of 20 times). The sum of all of these probabilities is 0.0414, which gives rise to a significance level less than 5%. Thus under the null hypothesis, there is less than a 5% chance that a respondent will answer correctly less than 6 times, or more than 14 times (out of 20). Thus when a given respondent actually answers correctly more than 14 times out of 20, this shows that for this respondent the null hypothesis is probably false (i.e. that a correct answer has a higher probability than 0.5). In other words, it means that the respondent can identify the correct answer more frequently than just by chance; which means, in turn, that the respondent can correctly distinguish between the pictures associated with the voiced and pre-nasalized voiced words, respectively.

Number correct	Probability
0	0.0000
1	0.0000
2	0.0002
3	0.0011
4	0.0046
5	0.0148
6	0.0370
7	0.0739
8	0.1201
9	0.1602
10	0.1762

Number correct	Probability
11	0.1602
12	0.1201
13	0.0739
14	0.0370
15	0.0148
16	0.0046
17	0.0011
18	0.0002
19	0.0000
20	0.0000
Total	1.000000

Appendix C

Following, we present the scores of the subjects from *Anzhong Village, Lashi Township* in the order we tested them:

<i>Anzhong</i> Subject #1: 23-year-old man, with junior high education				
Word	Correct	Incorrect	Both	Neither
/ndz _l ³³ / 'leopard'	4	6		
/dz _l ³³ / 'town'	3	7		
First pair total	7	13		
/ndv ³³ / 'to dig'	3	7		
/dv ³³ / 'turning-plow'	7	3		
Second pair total	10	10		

<i>Anzhong</i> Subject #2: 28-year-old man, with high school education – from <i>Shigu Zhen</i>				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	5	5		
/dz ₁ ³³ / ‘town’	8	2		
First pair total	13	7		
/ndv ³³ / ‘to dig’	4	5	1	
/dv ³³ / ‘turning-plow’	6	4		

<i>Anzhong</i> Subject #3: 32-year-old man, with junior high education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	2		8	
/dz ₁ ³³ / ‘town’		1	9	
/ndv ³³ / ‘to dig’	1	9		
/dv ³³ / ‘turning-plow’	8	1	1	

<i>Anzhong</i> Subject #4: 30-year-old man, with junior high education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’			10	
/dz ₁ ³³ / ‘town’			10	
/ndv ³³ / ‘to dig’		1	9	
/dv ³³ / ‘turning-plow’	1		9	

<i>Anzhong</i> Subject #5: 57-year-old woman, with 4 th grade elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	2	7	1	
/dz ₁ ³³ / ‘town’	2	8		
/ndv ³³ / ‘to dig’	4	5	1	
/dv ³³ / ‘turning-plow’	4	4	2	

<i>Anzhong</i> Subject #6: 61-year-old man, with elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	9	1		
/dz ₁ ³³ / ‘town’	9	1		
First pair total	18	2		
/ndv ³³ / ‘to dig’	5	5		
/dv ³³ / ‘turning-plow’	10			
Second pair total	15	5		

<i>Anzhong</i> Subject #7: 45-year-old woman, with 3 rd grade elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	4		6	
/dz ₁ ³³ / ‘town’		4	6	
/ndv ³³ / ‘to dig’	1	3	6	
/dv ³³ / ‘turning-plow’	4	1	5	

Anzhong Subject #8: 27-year-old woman, with junior high education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	1	1	8	
/dz ₁ ³³ / ‘town’	1	2	7	
/ndv ³³ / ‘to dig’	3	1	6	
/dv ³³ / ‘turning-plow’	3	2	5	

Anzhong Subject #9: 36-year-old man, with high school education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	7	3		
/dz ₁ ³³ / ‘town’	6	4		
First pair total	13	7		
/ndv ³³ / ‘to dig’	3	7		
/dv ³³ / ‘turning-plow’	8	2		
Second pair total	11	9		

Anzhong Subject #10: 67-year-old man, with high school education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	6	4		
/dz ₁ ³³ / ‘town’	5	5		
First pair total	11	9		
/ndv ³³ / ‘to dig’	2	8		
/dv ³³ / ‘turning-plow’	8	2		
Second pair total	10	10		

Anzhong Subject #11: 66-year-old man, with junior high education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	6	4		
/dz ₁ ³³ / ‘town’	7	3		
First pair total	13	7		
/ndv ³³ / ‘to dig’	4	6		
/dv ³³ / ‘turning-plow’	5	5		
Second pair total	9	11		

Appendix D

Following, we present the perception test scores of the subjects from Yanxi Administrative Village, Jinshan Township in the order we tested them:

Yanxi Subject #1: 75 year-old-man, with elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	6	3	1	
/dz ₁ ³³ / ‘town’	4	6		
/ndv ³³ / ‘to dig’		10		
/dv ³³ / ‘turning-plow’	9	1		
Second pair total	9	11		

Yanxi Subject #2: 35-year-old man, with elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	9	1		
/dz ₁ ³³ / ‘town’	6	4		
First pair total	15	5		
/ndv ³³ / ‘to dig’	5	5		
/dv ³³ / ‘turning-plow’	10			
Second pair total	15	5		

Yanxi Subject #3: 75-year-old man, with elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	8	2		
/dz ₁ ³³ / ‘town’	9	1		
First pair total	17	3		
/ndv ³³ / ‘to dig’	4	5	1	
/dv ³³ / ‘turning-plow’	9	1		

Yanxi Subject #4: 51-year-old man, with secondary-level technical school education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	3	7		
/dz ₁ ³³ / ‘town’	9	1		
First pair total	12	8		
/ndv ³³ / ‘to dig’	5	5		
/dv ³³ / ‘turning-plow’	9	1		
Second pair total	14	6		

Yanxi Subject #5: 50-year-old woman, with no education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	7	3		
/dz ₁ ³³ / ‘town’	4	6		
First pair total	11	9		
/ndv ³³ / ‘to dig’	5	5		
/dv ³³ / ‘turning-plow’	9	1		
Second pair total	14	6		

Yanxi Subject #6: 30-year-old man, with elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	8	2		
/dz ₁ ³³ / ‘town’	7	3		
First pair total	15	5		
/ndv ³³ / ‘to dig’	6	4		
/dv ³³ / ‘turning-plow’	4	6		
Second pair total	10	10		

Yanxi Subject #7: 55-year-old man, with 4 th grade elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	8	2		
/dz ₁ ³³ / ‘town’	9	1		
First pair total	17	3		
/ndv ³³ / ‘to dig’	2	8		
/dv ³³ / ‘turning-plow’	10			
Second pair total	12	8		

Yanxi Subject #8: 28-year-old woman, with junior high education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	6	4		
/dz ₁ ³³ / ‘town’	4	6		
First pair total	10	10		
/ndv ³³ / ‘to dig’	2	8		
/dv ³³ / ‘turning-plow’	7	3		
Second pair total	9	11		

Yanxi Subject #9: 32-year-old man, with 5 th grade elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	6	4		
/dz ₁ ³³ / ‘town’	1	9		
First pair total	7	13		
/ndv ³³ / ‘to dig’	3	7		
/dv ³³ / ‘turning-plow’	9	1		
Second pair total	12	8		

Yanxi Subject #10: 50-year-old woman, with 4 th grade elementary education				
Word	Correct	Incorrect	Both	Neither
/ndz ₁ ³³ / ‘leopard’	9	1		
/dz ₁ ³³ / ‘town’	6	4		
First pair total	15	5		
/ndv ³³ / ‘to dig’	4	6		
/dv ³³ / ‘turning-plow’	9	1		
Second pair total	13	7		

REFERENCES

- DELANCEY, Scott. 1990. “Sino-Tibetan languages.” In *The major languages of East and South-East Asia*, ed. Bernard Comrie 69-82. London: Routledge.
- HE Jiren and JIANG Zhuyi (和即仁、姜竹仪). 1985. 纳西语简志 [A Concise grammar of the Naxi language]. Beijing: Renmin Chubanshe.
- HE Zhiwu (和志武). 1987. 纳西语基础语法 [A Basic Naxi Grammar]. Kunming: Yunnan Minzu Chubanshe.

RAMSEY, Robert S. 1987. *The Languages of China*. New Jersey: Princeton University Press.

ROCK, Joseph F. 1963. A Na-Khi -- *English encyclopedic dictionary*. Serie Orientale Roma, no. 28. Rome: Istituto Italiano per il Medio ed Estremo Oriente.

YANG Jian, LIU Bing and PU Yuanyuan (杨鉴、刘兵、普圆媛). 1998. 纳西语辅音的声学分析 [Naxi Consonantal Acoustical Studies]. *Minzu Yuwen* 5:37-40.