PROSODIC PATTERNS AND PHONETIC REALIZATION OF KANAKANAVU LEXICAL STRESS ON DISYLLABIC WORDS

Sally Chen

National Taiwan University
sallyhfchen@ntu.edu.tw

ABSTRACT

To the best of our knowledge, this is by far the first phonetic study on Kanakanavu, an endangered Austronesian language spoken in southern Taiwan. A disyllabic word pair only differs in stress position was recorded from four native speakers. Tone type and corresponding acoustic correlates were analyzed. Results showed that lexical stress is realized with ample prosodic patterns when it is on penultimate position, and differences of these patterns were reflected on the corresponding acoustic correlates: Duration of stressed and unstressed vowels did not serve as the major cue. The stressed vowel was even shorter when the word bears stress on the penultimate syllable. In terms of F0, peak alignment pattern conformed to the previous findings on other languages in that temporal coordination for L+H* was demarcated later than H*. However, the maximal pitch values of H* and L+H* did not differ much from each other. Further research will be done based on these findings.

Keywords: Austronesian, endangered languages, lexical stress, vowel duration, peak alignment.

1. INTRODUCTION

Once classified as southern Tsou, in June 2014, along with Hla'alua, Kanakanavu was certificated as the fifteenth and the sixteenth aboriginal peoples in Taiwan. Contrast to this political fusion, linguistically, the two languages have been believed to be different from Tsou back in the Japanese Colonial Era (1895-1945).

Hla'alu and Kanakanavu are similar in several aspects: Originated from Kaohsiung City in southern Taiwan, they are now spoken by no more than 1,000 people, which, in accordance with to the standard of UNESCO, belong to the languages that are in serious danger of disappearing.

Study of the Tsou language started quite early. Hla'alua, on the other hand, has not been extensively investigated until recently. In terms of word stress, the default position for Tsou is believed to be on the penultimate syllable, and suffixation does not cause stress shift. For example, the word *'amo ("father")* will become *a'mo-u ("my father")* after the suffix is added [14]. However, the distribution of word stress in Hla'alua is more complicated: For words made up with two syllables, it is usually the penultimate syllable that bears the primary stress, *e.g., vi'aru ("cat")*, whereas for words composed of three or more syllables, primary stress may fall on the penultimate or the antepenultimate syllable, *e.g., sa'ria or 'saria ("house")* [12].

Unlike Tsou and Hla'alua, Kanakanavu has not yet been thoroughly studied. Sound inventory of this language is similar to that of Tsou and Hla'alua. Impressionistic observation on the recorded data showed that distribution of word stress in this language is diverse, more like the case of Hla'alua. In this paper, a Kanakanavu disyllabic word pair contrasts in stress position was investigated to find out possible prosodic patterns and corresponding phonetic realizations. Results obtained are crucial for further research on the distribution of lexical stress of this endangered language.

The assignment of prosodic patterns, or more specifically, tone types, would be following the tradition of Tone and Break Indices (see [4] and [5] for details). First introduced by Pierrehumbert, the three-way distinction of rising accents, H*, L+H*, and L*+H, is based on the overall shape of their F0 contours [13]. Among them, the uni-tonal H* and the bi-tonal L+H* have been more frequently investigated. The main difference between the two resides in the fact that the former reaches the peak from the midpoint of one's pitch range and the latter shows a rise from a low point [1]. In terms of functions, in English, H* is usually used upon introduction of new information, whereas L+H* is exploited when contrastive information is indicated. However, choice of tone type differs from language to language. Even speakers of different varieties of one language may also show different preference for tones. For example, those who spoke Southern Californian English switched from H* to L+H* with increased emphasis, but this was not observed among speakers of Minnesotan English [1].

In general, word elicitation task is analogous to the elicitation of new information. However, since different default patterns might be adopted for individual languages or different dialects of a language, no specific expectation was set for Kanakanavu word stress prior to analysis.

Perception of stress is believed to be multidimensional. The four common psychological dimensions, or acoustic correlates, are segmental length, loudness (intensity), fundamental frequency, and vowel quality, as in [7] and [15]. Among them, intensity is severely influenced by heterogeneous recording conditions. As none of the recordings of this study was conducted in a sound attenuated booth (rather, they were done in the tribe), intensity would not be included for analysis. Vowel quality was not included in the analysis, either, since it is well controlled: Both vowels of the disyllabic words were not reduced.

Previous studies on the phonetic realization of stress showed that compared with their unstressed counterparts, stressed syllables usually have longer duration, higher intensity and longer duration [15]. Upon the two remaining acoustic correlates included in this study, duration and pitch height, we would like to see whether Kanakanavu shows similar features for lexical stress as other languages do. The more recent concern of alignment of F0 events with respect to segments was also included in the following analyses.

2. METHOD

2.1. Informants

Four native speakers of Kanakanavu, balanced in gender, were recorded in their homes in Namasia District, Kaohsiung, Taiwan. The informants were all in their 60s or 70s.

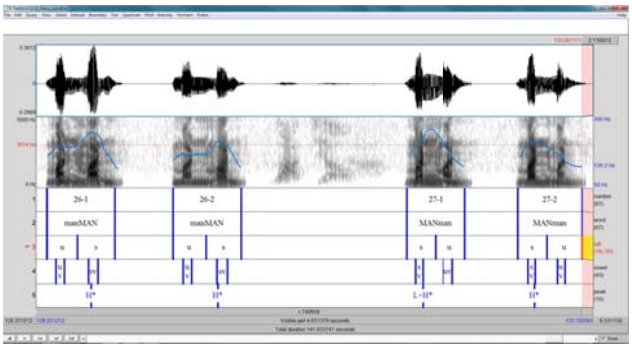
2.2. Material

A list of Kanakanavu words were recorded from each informant. In this paper, only different renditions of the production of the disyllabic word pair *man'man* (*"like"*) versus *'manman* (*"chili pepper"*) were analyzed. These two words are composed of exactly the same segments and only contrasts in stress position.

2.3. Data labelling and analyses

The Praat software [6] was used for data labelling and collection. For each token, five layers were labelled in the TextGrid file. As illustrated in Figure 1, the layers include number of the target word in the recording list, its orthography, the stressed and unstressed syllables/vowels of the word, and the tone type of the lexical stress. As aforementioned, the assignment of the tone type followed the tradition of Tone and Break Indices. Details can be found in [4] and [5].

Figure 1: Illustration of how target words were labelled in this study.



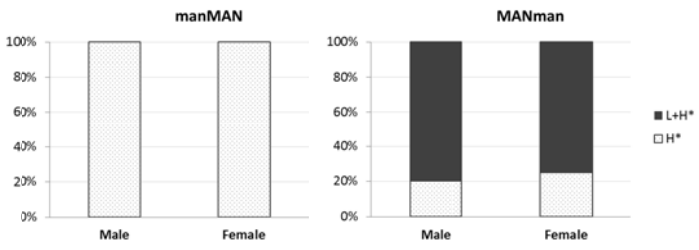
A Praat script was written to extract the phonological and phonetic labels in the TextGrid file. Tone types, and their corresponding durational and pitch information were collected and analysed.

3. RESULTS AND DISCUSSION

3.1. Tone types

Of the three most common tone types regarding rising contours [13], the uni-tonal H* (direct peak) and bi-tonal L+H* (gradual rise to peak) were used by the Kanakanavu informants in this disyllabic word pair. The four of them consistently pronounced *man'man* with an H*; whereas they preferred the gradual rise of L+H* for *'manman*, the word with the same segmental combination but accented in the penultimate position, as shown in Figure 2.

Figure 2: Tone type used in the production of the Kanakanavu *man'man*–*'manman* pair.



3.2. Phonetic realization

Vowel duration, peak alignment and maximal pitch value were included for analysis.

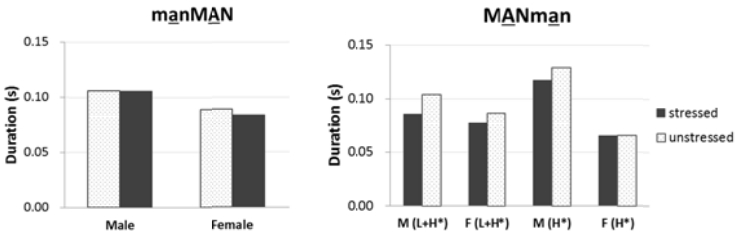
3.2.1. Vowel duration

As difference in the relative durations of vowels in a two-syllable word may change the perceived stress pattern for the word [7], lengthening or reducing of vowel duration has been believed to be a crucial perceptual cue for stressed or unstressed vowels.

Based on the previous findings of other languages, longer duration was expected for stressed vowels in Kanakanavu.

The results did not meet our expectation, though. As shown in Figure 3(b), for *'manman*, the word with stress on the penultimate syllable, the duration of the unstressed vowels was longer than, or equal to, that of their stressed counterparts, and this held true for both genders and tone types. One might speculate that the longer duration of the unstressed vowels could be resulted from the effect of final lengthening, as these target words were recorded in isolation and each of which was likely to be treated as an independent prosodic unit. This account was not supported, though. In Figure 3(a), for *man'man*, the word stressed in the final syllable, the stressed /a/ sound was not longer than the unstressed /a/, either. This contradicts our naïve intuition, as stressed vowels for isolated words in the final position should have been lengthened to an even greater degree due to the confounding effect of lexical stress and final lengthening.

Figure 3: Stressed and unstressed vowel duration of the Kanakanavu *man'man-'manman* pair.



(a) (b)

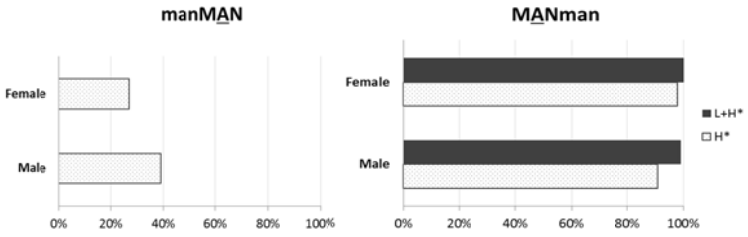
3.2.2. Peak alignment

Research on tonal alignment had not drawn much attention until the last decade. As any tonal target may vary along the "horizontal" (time) and the "vertical" (F0) directions, the phonetic properties regarding the two dimensions were called alignment, or scaling [2]. Moreover, the temporal coordination of fundamental frequency and "landmarks" in the segmental string is believed to be always aligned in a predictable way [3]. In terms of tonal alignment, one of the most extensively investigated languages is English, *e.g.*, [8], but recent research has been extended to various languages or different varieties of the same language all over the world. Factors in addition to tone type were also found to influence the F0 alignment. For instance, the durational account and the structural account have been proposed in the discussion of the effect of phonemic vowel length contrast on Dutch [8].

As previous findings in Dutch revealed that relative to the accented vowel, nuclear peaks (peaks in the final syllable) are aligned much earlier than prenuclear peaks (peaks in the penultimate syllable) [10], we hypothesized that the peak alignment of *mam'man* would be earlier than that of 'manman. Furthermore, final stresses realized as H* were expected to show earlier alignment than those realized as L+H*, since sufficient time was provided in preparation for full articulation.

Indeed, as illustrated in Figure 4, the word stressed in final syllable (which were consistently realized as H*) showed earlier alignment than the word stressed in the penultimate syllable. For *man'man* (Figure 4(a)), peaks were all aligned no later than 40% from the onset of the stressed vowel, and female informants produced even earlier peaks than males. In the case of *'manman* (Figure 4(b)), the word with a penultimate stress, H* was aligned much later (≥ 85%). Even so, the alignment was still slightly earlier than that of L+H* realization.

Figure 4: Peak alignment in the stressed syllable of the Kanakanavu *man'man-'manman* pair.



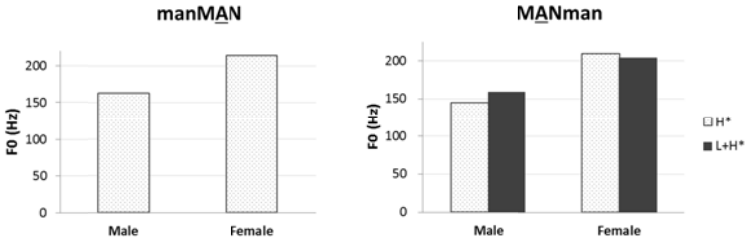
(a) (b)

3.2.3. Maximal pitch height

In addition to temporal alignment, the pitch maximum of F0 peaks is usually measured in studies with regards to phonetic realization of lexical stress. The maximal pitch of stresses realized as L+H* are expected to achieve for a higher value since longer preparation time is provided for full articulation.

As shown in Figure 5(a), the maximal pitch values of H* realization of both genders served as the baseline for the fundamental pitch difference between males and females. The fact that the average taken from the two stress types of *'manman* (Figure 5(b)) was comparable to the baseline indicates that the samples we collected were from a normal distribution. One can also find that the maximal pitch value of stresses realized as L+H* was slightly higher than those realized as H* among male informants, which met our expectation. However, an opposite trend was found for females, though the difference was minute.

Figure 5: Maximal pitch value of the stressed vowel of Kanakanavu *man'man-'manman* pair.



(a) (b)

Regardless of direction, the difference was small. One possible account is that the informants were asked to produce isolated words. Larger difference might be found in longer prosodic units, *e.g.*, phrases or complete sentences, due to physiological constraint. In addition, compared with spontaneous speech, in which contexts are always provided, the neutral (and thus more boring) property of single word elicitation might also contribute to a relatively lower overall pitch.

4. CONCLUSION

To conclude, the Kanakanavu disyllabic word pair investigated is realized with ample prosodic patterns when the stress is assigned on the penultimate syllable, and the difference observed in the two stress patterns was reflected in the corresponding acoustic correlates. In this language, vowel duration seemed not to be the major cue for realizing lexical stress, as stressed vowel of the final stress were not longer than the unstressed ones; in the case of penultimate stress, the stressed vowel showed even shorter duration than its unstressed counterpart. In terms of F0, peak alignment patterns conformed to the findings of previous studies on other languages in that temporal coordination for L+H* was demarcated later than H*. Last, though opposite trends were found between males and females, the maximal pitch value of H* and L+H* tones did not differ much from each other.

5. FUTURE STUDIES

This study was the first step for investigating word stress pattern in Kanakanavu. The impressionistic observation on the data at hand has indicated that the lexical stress pattern of Kanakanavu is more similar to Hla'alua. Stress shift is commonly observed and the stress position is more flexible, as compared to Tsou.

One issue worth further research is whether vowel length contrast is phonemic in Kanakanavu. The informants had the perception that the contrast is phonemic based on their native speakers' intuition; however, minimal pairs contrasting in vowel length are yet to be found. More fieldwork is needed to gain a better understanding of this issue.

Another possible direction for future research is syllable structure, as it is believed to be influential for the assignment of the position of lexical stress, and its combination with vowel dropping might further complicate the matter. Previous studies also found that peaks in closed syllables tended to be aligned later than those in open syllables [11]. Thus, all the other things being equal, if there exists any Kanakanavu word pair that only differs in the addition of a final consonant, a more detailed investigation of the effect of syllable structure on lexical stress assignment, or more specifically, on the peak alignment of the stressed vowel, would be possible to carry on in the future.

6. REFERENCES

[1] Arvaniti, A., & Garding, G. (2007). Dialectal variation in the rising accents of American English. *Laboratory phonology*, *9*, 547-576.

[2] Arvaniti, A., Ladd, D. R., & Mennen, I. (1998). Stability of tonal alignment: the case of Greek prenuclear accents. *Journal of phonetics*, *26*(1), 3-25.

[3] Atterer, M., & Ladd, D. R. (2004). On the phonetics and phonology of "segmental anchoring" of F0: evidence from German. *Journal of Phonetics*, *32*(2), 177-197.

[4] Beckman, M. E., & Ayers, G. (1997). Guidelines for ToBI labelling. *The OSU Research Foundation*, *3*.

[5] Beckman, M. E., & Hirschberg, J. (1994). The ToBI annotation conventions. *Ohio State University*.

[6] Boersma, Paul & Weenink, David (2012). Praat: doing phonetics by computer [Computer program]. Version 5.3.32, retrieved 17 October 2012 from http://www.praat.org/

[7] Fry, D. B. (1958). Experiments in the perception of stress. *Language and speech*, *1*(2), 126-152.

[8] Ladd, D. R., Mennen, I., & Schepman, A. (2000). Phonological conditioning of peak alignment in rising pitch accents in Dutch. *The Journal of the Acoustical Society of America*, *107*(5), 2685-2696.

[9] Ladd, D. R., & Schepman, A. (2003). "Sagging transitions" between high pitch accents in English: Experimental evidence. *Journal of Phonetics*, *31*(1), 81-112.

[10] Ladd, D. R., Schepman, A., White, L., Quarmby, L. M., & Stackhouse, R. (2009). Structural and dialectal effects on pitch peak alignment in two varieties of British English. *Journal of Phonetics*, *37*(2), 145-161.

[11] Mücke, D., Grice, M., Becker, J., & Hermes, A. (2009). Sources of variation in tonal alignment: evidence from acoustic and kinematic data. *Journal of Phonetics*, *37*(3), 321-338.

[12] Pan, C-j. (2012). A grammar of Lha'alua, an Austronesian language of Taiwan. PhD thesis, James Cook University.
[13] Pierrehumbert, J. (1980). The phonetics and phonology of English intonation. *Unpublished Ph. D. thesis, MIT.*
[14] Zeitoun, E. (2000). A reference grammar of Tsou. *Formosan Language Series*, (7). Taipei: Yuanliou Pub. Co.
[15] Zhang, Y., Nissen, S. L., & Francis, A. L. (2008). Acoustic characteristics of English lexical stress produced by native Mandarin speakers. *The Journal of the Acoustical Society of America*, *123*(6), 4498-4513.