AN ACOUSTIC-PHONETIC DESCRIPTION OF WORD TONE IN KAGOSHIMA JAPANESE

Shunichi Ishihara

Japan Centre (Asian Studies), and
Phonetics Laboratory, Department of Linguistics (Arts), The Australian National University

ABSTRACT

The Japanese dialect of Kagoshima (KJ) has two different surface pitch patterns, $(L)^0HL$ and $(L)^0H$. In this study, the properties of these two surface pitch patterns of KJ will be acoustically-phonetically described by means of z-score normalisation. Words consisting of two, three, four and five syllables were used in this study (the syllable structure is a CV) as test words, and the F0 of each syllable nucleus was extracted and the raw F0 values were z-score normalised. Two native speakers of KJ (one male and one female) participated in this study. The tonal representation of KJ words will be discussed on the basis of the z-score normalised results.

1. INTRODUCTION

Japanese is rich in dialectal variations, and the pitch patterns of many dialects have been described and phonologically analysed by various scholars. Although much work has been done on the acoustics of pitch-accent of Standard Japanese (SJ) (cf. Poser 1986, Pierrehumbert and Beckman 1988, Kubozono 1993), unfortunately few quantitative acoustic-phonetic descriptive works have been done on any variations. This is the first motivation for this study.

KJ shows two different surface pitch patterns, $(L)^0HL$ and $(L)^0H$. In the former pattern, only the penultimate syllable of a word has a high pitch and every other syllable has a low pitch, and in the latter pattern, only the last syllable of a word has a high pitch and every other syllable before it has a low pitch. Therefore, KJ has often been referred to (cf. Hirayama 1960, McCawley 1977, Haraguchi 1977) as a typical example having only two accentual classes: type A and type B, respectively. Shibatani (1990) considers that these two pitch contours belong to accented and unaccented words, respectively. An alternative analysis (e.g. Hayata 1977) is that there is no accentual contrast in words of KJ, but that words fall into one of two distinctive melody classes (LHL and LH). According to this analysis, since a word itself has a specific pitch contour, either LHL or LH, KJ is considered as a word tone dialect (in contrast Chinese or Thai which are syllable tone languages). In the former analysis, KJ is accentually treated in the same way as SJ (cf. the location of an accent needs to be specified in the lexicon). In the alternative two-melody analysis, the location of the accent is not a lexical property, and what one needs to know is which word melody a word has.

Pierrehumbert and Beckman (1988) demonstrate that the linking and spreading of word-level melodies is not the best analysis for surface tone pattern in SJ, and showed that accentual phrases must be underspecified for tone even at the most surface level. They posit four target tones that characterise the tonal phrase of SJ: a boundary tone (L%) which occurs at the phrase boundary, a phrase High tone which occurs at phrase beginning, an accentual High tone which occurs at accent nucleus, and an accentual Low tone which occurs shortly after the accentual High tone. If we apply the target-tone model to KJ, tonal targets need to be specified at least in four locations for accented words and in three locations for unaccented words as shown in figure 1.

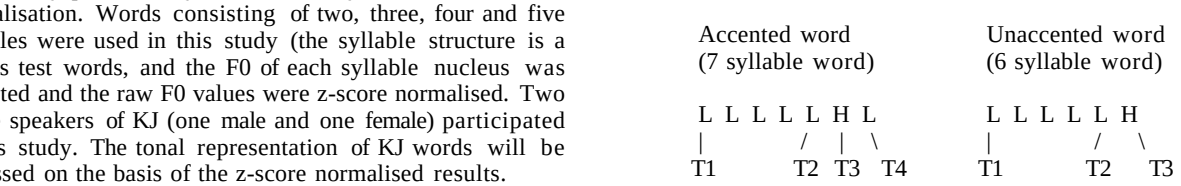

Figure 1: Underspecified tonal representations of accented and unaccented words of KJ. T stands for target tones. Although T1, T2, and T3 are tentatively used for both classes of words on the basis of syllable location, they do not necessarily have the same target values between accented and unaccented words.

In this study, the target-tone model is adopted, and each tonal value will be discussed on the basis of the z-score normalised data. I will discuss this mainly with respect to the initial boundary tone (T1), the low tone syllable immediately before the high tone syllable (T2), and the high tone syllable (T3). As for T4, only a visual impression will be provided.

Although I use the terms; accented and unaccented, to refer to the words that have $(L)^0HL$ and $(L)^0H$ pitch patterns, respectively, it is not clear at this stage that an accented word has a lexical accent and an unaccented word does not. I will discuss this point later.

2. EXPERIMENT PROCEDURE

In this experiment, two native speakers of KJ (one male who is 30 years of age; TI, and one female who is 23 years of age; TS) who were temporarily visiting Australia, were simply asked to read a list of words (approximately 125 words) containing the following test words shown table 1. The reading list was written in standard Kanji orthography and it was read six times including one practice reading. The syllable structure of the following test words is CV. As for the test words, I tried to select natural words which are used on a daily basis and tried to use both voiceless and voiced consonants as evenly as possible and different vowels (especially in terms of vowel height). However, nasals are frequently used in many syllables of the test words. This is because nasals are considered to have a minimal perturbatory effect on the following vowel.

The elicitation sessions were conducted in a acoustically nonreverbeative room, and the reading material was recorded on high-quality normal position types using a SONY TCM-

5000EV tape recorder and a SONY ECM-D8 microphone. The raw material was digitised with Computerized Speech Laboratory (sampling rate = 10000 Hz) at the phonetics Laboratory of the Australian National University and F0 was extracted at 0%, 20%, 40%, 60%, 80% and 100% points of each syllable nucleus by means of the Automatic Pitch Extraction Command of the same machine and then, the raw F0 values were z-score normalised.

	Accented	Unaccented
2	hata 'flag' kagu 'furniture' kubi 'neck' kugi 'nail mune 'chest' siga 'place name'	baka 'stupid' gake 'cliff' kage 'shadow' kagi 'key' kega 'injury' kogu 'to row' kumo 'cloud' mimi 'ear'
3	kakuho 'to ensure' kimono 'cloth' komimi 'ear' minami 'south' sakana 'fish' tomato 'tomato'	gakubu 'department kagami 'mirror' megane 'glasses' mimiwa 'ear ring' namida 'tear' suzume 'sparrow'
4	getabako 'shoe box' hanamizu 'runny nose' kagaribi 'torch' kanimiso 'crab paste' kozakana 'small fish' mamagoto 'cubbyhouse'	demakase 'random answer' kakimono 'writing yamanami 'mountain tops' dokubutu 'poison' miminari 'ringing in one's ear' nanohana 'rape blossom
5	matiyakuba 'town hall' minamigawa 'south side' minamimura 'south village'	kagamimoti 'rice cake' mimizawari 'jarring' miyagemono 'present'

Table 1: Corpus. Numerals stand for the number of syllables. Accented words have a high pitch in the penultimate syllable, and unaccented words in the final syllable.

3. ACCENTED AND UNACCENTED WORDS

The z-score normalisation was used in this study in order to convert the individual speakers' outputs into linguistic-phonetic representations of the two surface pitch realisations of KJ. This is following Rose (1987) who reports its superiority in F0 normalisation. The z-score normalisation procedure is; F0norm = (F0i - x) / SD, where F0i is a sampling point, x is the average F0 from the arithmetic mean of all the sampling points, and SD is the standard deviation around the mean of those points. Table 2 contains the normalisation parameters of the two speakers.

	x	SD
TI	114.4	14.8
TS	203.7	16.3

Table 2: Normalisation parameters (Hz). x stands for the average F0, and SD for standard deviation.

The mean normalised F0 values of accented words (HL, LHL, LLHL, LLLHL) are plotted all together against equalised duration (%) in figure 2, and those of unaccented words (LH, LLH, LLLH, LLLLH) are in figure 3.

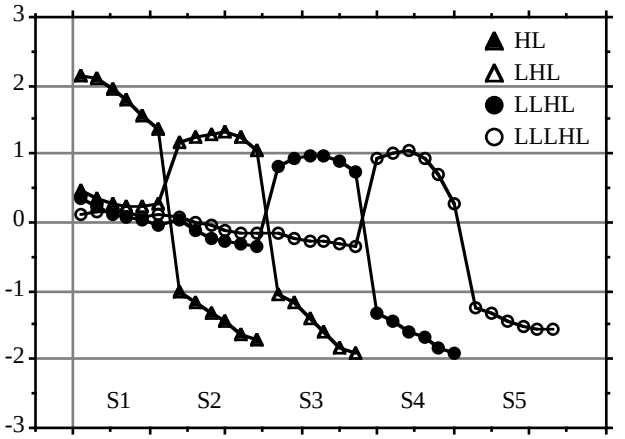

Figure 2: Linguistic-phonetic representations of accented words (HL, LHL, LLHL, LLLHL). Y-axis is z-score and X-axis is equalised duration (%). S1, S2, S3, S4, and S5 stand for first syllable, second syllable, third syllable, fourth syllable and fifth syllable, respectively.

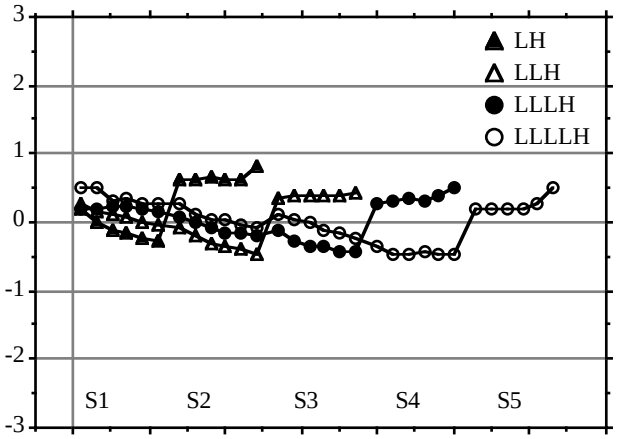

Figure 3: Linguistic-phonetic representations of unaccented words (LH, LLH, LLLH, LLLLH). Y-axis is z-score, X-axis is equalised duration (%). S1, S2, S3, S4, and S5 stand for first syllable, second syllable, third syllable, fourth syllable and fifth syllable, respectively.

In the following subsections, tonal differences between accented and unaccented words will be discussed in terms of word-initial low tone (T1), the low tone syllable immediately before high tone (T2), and high tone syllable (T3). All the statistical comparisons are based on the z-score normalised values. Whenever it is necessary, t-test was used, and 0.05 was set as the level of significance.

One can notice some differences between accented and unaccented words from a visual comparison. The high tone syllables of the accented words appear to be higher in target points than those of the corresponding unaccented words. The final low tone of accented words seems to have a rather low target value compared to those of initial low tones.

However, these visual impressions need to be statistically justified.

3.1. Word-Initial Low Tone Syllable

In this sub-section, the data concerned with the initial boundary low tone (T1) will be presented for both accented and unaccented words. In order to see whether there is a meaningful difference in F0 between accented and unaccented words at the word-initial position, the highest F0 values of accented and unaccented words of the first syllable were statistically compared using t-test. Table 3 contains the mean highest F0 values of the initial syllables, the standard deviations of these, and the results of t-tests for both TI and TS. Those accented and unaccented words that have the same number of low tone syllables before the high tone syllable, were statistically compared.

According to table 3, although there is a significant difference in the highest F0 value between the initial low tone syllable of LH and that of LHL for both speakers ($p < 0.0001$ for TI and $p = 0.0443$ for TS), TS does not show a significant difference in the highest F0 value between the initial low tone syllable of LLH and that of LLHL ($p = 0.4555$). When it comes to the comparison between LLLH and LLLHL, neither TI nor TS shows a significant difference in the highest F0 value ($p = 0.4008$ for TI and $p = 0.3992$ for TS). This result indicates therefore that it is not necessary to posit different tonal values for accented and unaccented words at the initial syllable (T1).

TI	Unaccented	Accented
1	LH	LHL
Mean	0.249	0.964
SD	(0.451)	(0.397)
P value	0.0001	
2	LLH	LLHL
Mean	0.337	0.629
SD	(0.377)	(0.423)
P value	0.0033	
3	LLLH	LLLHL
Mean	0.392	0.417
SD	(0.337)	(0.229)
P value	0.4008	

TS	Unaccented	Accented
1	LH	LHL
Mean	0.003	0.233
SD	(0.467)	(0.602)
P value	0.0443	
2	LLH	LLHL
Mean	0.130	0.117
SD	(0.447)	(0.395)
P value	0.4555	
3	LLLH	LLLHL
Mean	0.152	0.179
SD	(0.284)	(0.409)
P value	0.3992	

Table 3: Comparison between accented and unaccented words in terms of initial low tone syllable (LH vs LHL, LLH vs LLHL, and LLLH vs LLLHL). Bold means that there is a significant difference ($p < 0.05$).

TI	Unaccented	Accented
2	LH	LHL
Mean	-0.769	-0.250
SD	(0.416)	(0.370)
P value	0.0001	
3	LLH	LLHL
Mean	-1.047	-0.769
SD	(0.342)	(0.255)
P value	0.0003	
4	LLLH	LLLHL
Mean	-0.800	-0.617
SD	(0.348)	(0.360)
P value	0.0147	

TS	Unaccented	Accented
2	LH	LHL
Mean	-0.371	0.419
SD	(0.310)	(0.345)
P value	0.0001	
3	LLH	LLHL
Mean	-0.594	-0.166
SD	(0.399)	(0.405)
P value	0.0001	
4	LLLH	LLLHL
Mean	-0.525	-0.308
SD	(0.301)	(0.309)
P value	0.0054	

Table 4: Comparison between accented and unaccented words in terms of high pitched syllable. (LH vs LHL, LLH vs LLHL, and LLLH vs LLLHL). Bold means that there is a significant difference ($p < 0.05$).

3.2. Low Tone Syllable Immediately Before High Tone Syllable

In this sub-section, the data concerning the low tone immediately before the high tone syllable (T2) will be presented for both accented and unaccented words. In order to see whether there is a meaningful difference in target value between accented and unaccented words at the pre-high tone syllable position (T2), the lowest F0 values of the low tone syllable which comes immediately before the high tone syllable are statistically compared. In most cases, the lowest F0 value was observed in the low tone syllable immediately before the high tone syllable, yet sometimes it was observed in the second syllable before the high tone syllable in some five and six syllable words. Table 3 contains the mean lowest F0 values of the low tone syllable immediately before the high tone syllable, the standard deviations, and the results of t-tests for both speakers. Accented and unaccented words that have the same number of low tone syllables before the high tone syllable, were compared using t-test.

TI	Unaccented	Accented
2	LH	LHL
Mean	1.283	1.744
SD	(0.537)	(0.603)
P value	0.0013	
3	LLH	LLHL
Mean	1.298	1.403
SD	(0.586)	(0.686)
P value	0.2678	
4	LLLH	LLLHL
Mean	0.900	1.571
SD	(0.607)	(0.523)
P value	0.0003	

TS	Unaccented	Accented
2	LH	LHL
Mean	0.989	1.370
SD	(0.416)	(0.309)
P value	0.0001	
3	LLH	LLHL
Mean	0.707	0.966
SD	(0.413)	(0.512)
P value	0.0177	
4	LLLH	LLLHL
Mean	0.694	0.833
SD	(0.385)	(0.220)
P value	0.1025	

Table 5: Comparison between accented and unaccented words in terms of high tone syllable. (LH vs LHL, LLH vs LLHL, and LLLH vs LLLHL). Bold means that there is a significant difference ($p < 0.05$).

According to table 3, it was observed for both speakers that the lowest F0 value of the low tone syllable immediately before the high tone syllable is significantly higher for

accented words than unaccented words ($p < 0.0147$). This result indicates therefore that it is necessary to posit different tonal values for accented and unaccented words at the end of low tone sequence (T2).

3.3. High Tone Syllable

Many of the phonetic and phonological studies of SJ assume that accented words and phrases have a higher pitch peak than their unaccented counterparts (Poser 1984, Kubozono 1986). Judging from figures 2 and 3, it is visually clear, as I said before, that the peak of accented words is higher than that of corresponding unaccented words. The peak F0 values (T3) were compared between accented and unaccented words in order to validate this visual impression. If we compare the accented and unaccented words in the light of the location of high tone syllable, three comparisons are possible; LH vs LHL, LLH vs LLHL, and LLLH vs LLLHL from the data. Table 4 contains the mean peak F0 values, the standard deviations of them, and the results of t-tests for both speakers.

As for the comparison between LH and LHL, the peak F0 value of an accented word is significantly higher than that of an unaccented word for both speakers ($p = 0.0013$ for TI and $p < 0.0001$ for TS). As for the comparison between LLH and LLHL, TI did not show a significant difference in peak F0 value ($p = 0.2678$) while TS showed a significant difference (0.0177). As for the comparison between LLLH and LLLHL, on the other hand, the peak F0 values of LLLHL were significantly higher than those of LLLH for TI ($p = 0.0003$) while no significance was observed for TS ($p = 0.1025$).

Although a clear significance can not always be observed between the high tone of an accented word and an unaccented word, it is likely that the high tone of an accented word tends to be realised higher in F0 than that of the corresponding unaccented word. This indicates that different tonal values need to be used to represent the high tone syllable of an accented word and that of an unaccented word.

4. TONAL REPRESENTATION OF KAGOSHIMA JAPANESE

In this section, I will show only a simple sketch of what tones KJ needs to include on the basis of the above mentioned results.

As I have shown in the previous section, the initial boundary tone seems to have the same value regardless of the accentual type of a word. Following this observation, I posit %L (initial low boundary tone) for both accented and unaccented words. When it comes to the low tone which comes immediately before the high, however, the low tone of an accented word appears to have a different value from that of an unaccented word. That is, T2 of an accented word is higher in tonal value than that of an unaccented word. Following this finding, L+ is posited for the T2 of an accented word, and L- for that of an unaccented word. As for the high tone syllable, H is assigned for accented words, and H% (final high boundary tone) for unaccented words. At this stage, it is not certain what tonal value should be assigned to the final low boundary tone (T4) of an accented word. Judging from the visual impression of figure 2, the final low tone syllable of an accented word (T4) seems to have another tonal value which is relatively lower to the other low tone values (%L, L+, and L-) as the F0 fall from the preceding high tone looks rather sharp and the target point looks lower than the general regression of the sequence of low tone syllables. Figure 4 below summarises the tonal representations of accented and unaccented words of KJ. Although I tentatively presented the tonal values of KJ, it needs to be attested as to whether they are real lexical properties as some tones, such as boundary tones, may be associated with a higher intonational level.

It is widely known in SJ that lexical accent instigates intonational phenomena such as accentual fall, accentual boost, downtrend and so on. Although it is still controversial as to whether accent is a lexical property or not in KJ, if one applies the concept of lexical accent to KJ, it is feasible to say that these tonal differences between accented and unaccented words (T2, T3 and possibly T4) are caused by lexical accent. In the case of the alternative two-melody analysis, it is not clear as to the cause of these tonal differences between accented and unaccented words. Further research should address this point.

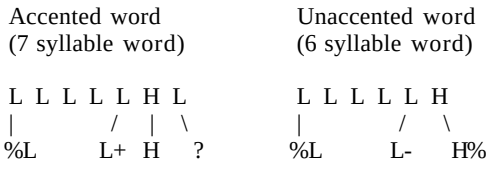

Figure 4: Tonal representation of the KJ (accented and unaccented words)

5. REFERENCES

1. Kubozono, H. The Organization of Japanese Prosody. Kuroshio Publisher. 1993.

2. McCawley, J. The Phonological Component of a Grammar of Japanese. The Hague/Paris: Mouton. 1968.

3. Pierrehumbert, J. and Beckman, M. Japanese Tone Structure. Cambridge, Mass.: MIT Press. 1988.

4. Poser, W.J. The Phonetics and Phonology of Tone and Intonation in Japanese. Ph.D. Dissertation, MIT. 1986.

5. Rose, P. “Considerations in the normalisation of the fundamental frequency of linguistic tone.” Speech Communication 6, 343-351, 1987.

6. Haraguchi, Shosuke. The Tone Pattern of Japanese: An Autosegmental Theory of Tonology. Tokyo: Kaitakusha. 1977.

7. Hayata, Teruhiro. “Seisei akusento ron [A generative theory of accent].” In Ono and Shibata (eds), Iwanami koza nihongo 5: Oninron. 323-360, 1977.

8. Hirayama, Teruo. Zenkoku akusento ziten [All-Japan accent dictionary]. Tokyo: Tokyodo. 1960.

9. Shibatani, M. The Languages of Japan. Cambridge: Cambridge University Press. 1990.