LINGUISTIC-TONETIC DIFFERENCES IN TARGET-TONE REALISATION: STANDARD VS. KAGOSHIMA JAPANESE

Shunichi Ishihara

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

ABSTRACT-Normalised F0 data from four Kagoshima Japanese speakers are used to investigate the relationship between the distance of two target tones and their F0 realisations in utterances of $HL^{(n)}$ sequences. The F0 minima of the $HL^{(n)}$ sequence is shown to vary as a function of the distance of the two target tones. Furthermore, the apparent existence is demonstrated of i) a default contour shape on the basis of which F0 is realised according to the distance of two target tones, and ii) a base-line beyond which F0 does not fall. Implications of the results for Pierrehumbert and Beckman's target-tone model are discussed, and the existence of two linguistically-tonetically different types in terms of target-tone realisation is hypothesised.

INTRODUCTION

In their analysis of the tone structure of Standard Japanese (SJ), Pierrehumbert and Beckman (1988) propose a new model to account for the F0 contours of utterances. They claim that the traditional account (i.e. McCawley, 1977; Haraguchi, 1977; Higurashi, 1983; Vance, 1987; etc) in terms of moraic pitch height does not explain the phonetics. Instead, they introduce an entirely new method of describing Japanese pitch accent by using only a few tones per phrase, with interpolation between them. The tones they assign to a phrase are limited to a boundary low tone (L%) at the beginning of an utterance; a phrasal peak high tone (H) which is normally attached to the second mora; an accent peak high-low tone (H*L) on the accented mora; and a boundary low tone at the end of each phrase.

The main reason for Pierrehumbert and Beckman's rejection of the moraic pitch-height model is the inadequacy of High-spreading and Low-spreading (cf. Haraguchi, 1977). In an anuclear phrase (e.g. a phrase having a $LH^{(n)}$ pitch configuration), the F0 values of the high pitched moras are expected to be the same. But the actual F0 contour is slow falling. This slow fall has been generally interpreted as a non-phonological or physiological downtrend (Kawakami, 1962; Kobayashi, 1963). Pierrehumbert and Beckman, however, found that the rate of this downtrend decreases as the number of the high pitched moras increases. This seems to reflect that the boundary to the next phrase would have a target value to which the contour is interpolating. They also claim that the declination rate of a low pitched mora sequence (i.e. a phrase having a $HL^{(n)}$ pitch configuration) decreases as a function of the number of Low pitched moras. Figure 1 schematically explains how the rate of F0 declination changes depending on the length of low pitched mora sequence.

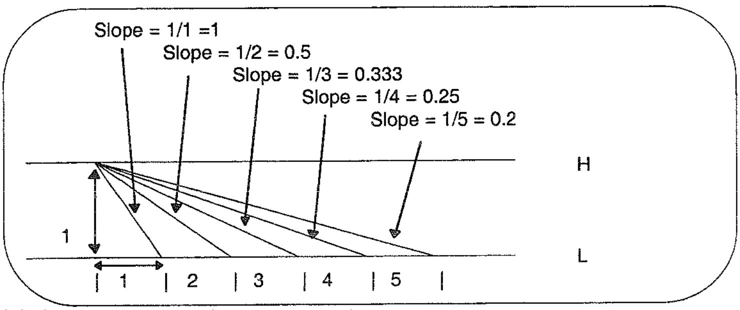

Figure 1: Schematic representation of F0 contour slope in a linear interpolation model of downtrend within unaccented accentual phrases of various lengths from Pierrehumbert and Beckman (1988: 41).

In Figure 1, the degree of F0 slope differs according to the length of low pitched moras in order to achieve the low target value.

If we apply Pierrehumbert and Beckman's target-tone model to Kagoshima Japanese (KJ), the sequence having a $HL^{(n)}$ pitch configuration is considered to be phonetically realised as the interpolation from a high tone to a low tone independent of the number of low pitched syllables, although those target values are different depending on the accentual types. That is, a target tone is accomplished irrespective of the length of the sequences, thus, the F0 minima observed in the $HL^{(n)}$ sequences should be consistent, being independent of the duration of the $HL^{(n)}$ sequences.

In this paper, I will describe how the $HL^{(n)}$ sequence of KJ is phonetically realised, and investigate the relationship between the distance between the two target tones and the F0 realisation of the sequence in question. On the basis of the description and investigation, some implications of Pierrehumbert and Beckman's target-tone model will be discussed.

THE ACCENTUATION OF KAGOSHIMA JAPANESE

KJ exhibits a two way accentual contrast; $(L)^0HL$ and $(L)^0H$ (cf. Hirayama, 1960). In the former pattern, only the penultimate syllable of a word has a high pitch, every other syllable has a low pitch. In the latter pattern, only the last syllable of a word has a high pitch, every other syllable before it has a low pitch. Following Hirayama (1960), the $(L)^0HL$ type is referred to as Type A and the $(L)^0H$ type as Type B. The examples shown in Table 1 exhibit this accentual contrast.

Syllable	Type A		Type B	
2	hana [HL]	"nose"	hana [LH]	"flower"
3	sakura [LHL]	"cherry blossom"	usagi [LLH]	"rabbit"
4	kagaribi [LLHL]	"bonfire"	kakimono [LLLH]	"document"

Table 1: The accentuation of KJ.

EXPERIMENT PROCEDURE AND NORMALISATION

Four native speakers of KJ (two females: TY and YN and two males: TT and NK aged between 25 and 35 years of age) participated in this study. One corpus containing 18 noun phrases differing in syllable number—with approximately 15 dummy phrases which were scattered at random throughout the corpus—was prepared. The pitch configurations of these target noun phrases are presented in Table 2. The syllable structure of these target noun phrases is $C^{(i)}V$. The informants were asked to read this corpus 10 times in the frame given in Table 2 (e.g. Naomi-wa yoka sakana to itta gayo "Naomi said good fish"). The recording was conducted in a room in Sydney with very low ambient noise, and the reading material was recorded onto high-quality normal position tapes using professional equipment. The raw material was digitised with Computerised Speech Laboratory (CSL) (sampling rate = 10000 Hz).

	First component	Second Component	n
Type A + Type A (AA)	LHL	$L^{(n)}HL$	1, 2, 3, 4, 5
Type A + Type B (AB)	LHL	$L^{(n)}H$	2, 3, 4, 5
Type B + Type A (BA)	LH	$L^{(n)}HL$	1, 2, 3, 4, 5
Type B + Type B (BB)	LH	$L^{(n)}H$	2, 3, 4, 5

Frame						
Frame	Naomi-wa	________	to	itta	gayo.	
	LLH L		L	HL	HL	
	Name-TOP	________	QTN.	say-past	SFP.	QTN = Quotation
	Naomi said	________.				SFP = Sentence final particle

Table 2: The pitch configuration of target noun phrases and carrier frame.

All phrases have the same $LHL^{(n)}H(L)$ pitch configuration (n = 1, 2, 3, 4, 5, 6), but have different accentual types (Type A or Type B). In this paper, the phrases having a Type A + Type A pitch configuration are referred to as AA phrase, and the other configurations also follow the same convention (AB, BA and BB Phrases). Another convention which needs an explanation is that, for example, 3a7a stands for the $HL^{(n)}$ sequences which is extracted from an AA phrase whose first component consists of 3 syllables and second component of 7 syllables (= LHL.LLLLLHL).

F0 was extracted from the $HL^{(n)}$ sequences using CSL's Automatic Pitch Extraction. F0 samples were taken at the onset, 50% and the offset of each syllable nucleus except for the initial high pitched syllable from which only the maximum value was sampled.

The logarithmic z-score normalisation—which Zhu (1999) reports the superiority of in F0 normalisation—was used in this study in order to exclude between-speaker differences and specify the invariant features (Rose, 1987). The logarithmic z-score normalisation procedure is: $F0_{norm} = (F0_i - x) / SD$, where $F0_i$ is a sampling point, x is the average F0 from all sampling points, and SD is the standard deviation around the mean of those points, all of which are logarithmic terms. All statistical comparisons are conducted on the basis of logarithmic z-score normalised F0 values. Table 3 contains the normalisation parameters of the four informants.

Speaker	TY	YN	TT	NK
x	2.133	2.013	2.284	2.262
SD	0.053	0.043	0.073	0.043

Table 3: Normalisation parameters in log F0

THE F0 REALISATION OF $HL^{(n)}$ SEQUENCES

In Figure 2 below, the mean normalised F0 curves of the $HL^{(n)}$ sequences (n = 2, 3, 4, 5, 6) collected from AA phrases are plotted against the mean absolute duration. Figure 2 shows that the minimum value seems to decrease as the sequence in question becomes longer until a certain stage (approximately -0.75 of Y-axis) below which F0 does not drop.

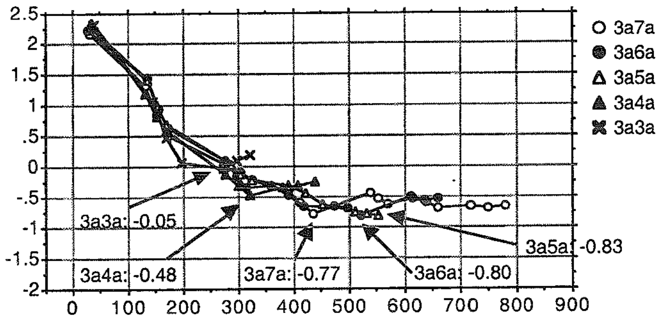

Figure 2: Normalised F0 realisations of KJ's $HL^{(n)}$ sequences (AA). The arrow indicates the F0 minima and the place where it was observed. X-axis is duration and Y-axis is normalised value.

Table 4 below shows the mean F0 minima of each type. Similarly to Figure 2, it can be seen from Table 4 that the mean F0 minima becomes smaller as the syllable number increases (i.e. 3a3b: -0.842 > 3a4b: -1.087 > 3a5b: -1.285 > 3a6b: -1.416).

ANOVA with post-hoc Scheffe F-test were conducted on the F0 minima collected from the $HL^{(n)}$ sequences of AA, AB, BA and BB Phrases, respectively. The results of these statistics are given in Table 5. The results shown in Table 5 statistically justify that the F0 minima decreases as a function of syllable number until a certain point below which it does not further drop. In BA phrases, for example, the relationship between the minimum value and the duration of $HL^{(n)}$ sequences is expressed using an inequality like 2b3a > 2b4a > 2b5a > 2b6a = 2b7a [DF (4, 191) = 91.466, p = 0.0001].

AA	3a3a	3a4a	3a5a	3a6a	3a7a
Mean	-0.159	-0.563	-0.892	-0.890	-0.942
SD	0.405	0.338	0.368	0.340	0.277
AB	3a3b	3a4b	3a5b	3a6b	
Mean	-0.842	-1.087	-1.285	-1.416	
SD	0.341	0.332	0.285	0.435	
BA	2b3a	2b4a	2b5a	2b6a	2b7a
Mean	0.732	-0.085	-0.444	-0.736	-0.796
SD	0.412	0.316	0.437	0.456	0.417
BB	2b3b	2b4b	2b5b	2b6b	
Mean	-0.529	-0.804	-1.018	-1.280	
SD	0.345	0.357	0.369	0.430	

Table 4: The mean minimum value and SD of each combination of accentual types.

AA	DF (4, 188) = 35.554, p = 0.0001			
	3a4a	3a5a	3a6a	3a7a
3a3a	6.511*	21.659*	21.274*	24.726*
3a4a		4.292*	4.192*	5.704*
3a5a			1.263E-4	0.102
3a6a				0.107
AB	DF (3, 154) = 20.036, p = 0.0001			
	3a4b	3a5b	3a6b	
3a3b	3.219*	10.373*	17.652*	
3a4b		2.094	5.869	
3a5b			0.922	
BA	DF (4, 191) = 91.466, p = 0.0001			
	2b4a	2b5a	2b6a	2b7a
2b3a	19.204*	40.881*	63.725*	67.267*
2b4a		3.721*	12.233*	14.223*
2b5a			2.525*	3.573*
2b6a				0.104
BB	DF (3, 156) = 28.648, p = 0.0001			
	2b4b	2b5b	2b6b	
2b3b	3.554*	11.225*	26.497*	
2b4b		2.147	10.642*	
2b5b			3.230*	

Table 5: The results of ANOVA and Scheffe F-test. * stands for significant at 95%.

DISCUSSION

The above statistically justified observation does not seem to be in accordance with the prediction of the target-tone model, in that the observe F0 minima for the $HL^{(n)}$ sequences vary accordingly to the duration of the $HL^{(n)}$ sequences. Besides this, two points need to be pointed out. First of all, as can be seen from Figure 2, the mean contour shapes of the five different types extensively overlap, showing an exponential curve. The same observation can be seen in the other phrase types as well. This implies that there seems to be a certain default contour shape on the basis of which F0 is realised according to the duration of $HL^{(n)}$ sequences. Secondly, the statistical results presented in Table 5 show that there seems to be a certain base-line below which F0 does not continue to fall.

Regarding the first point, Figure 3 below schematically represents a default contour shape. As I have shown, the $HL^{(n)}$ sequences I have investigated are immediately followed by a high pitched syllable, and, as a result of this high pitched syllable, F0 usually starts rising before reaching the offset of the final low pitched syllable of the sequence in question. As with a sequence having an HLLL pitch configuration, for example, F0 is realised on the basis of this default contour shape until the second low pitched syllable (= L2), and F0 is pulled up in the third low pitched syllable (= L3f) by the immediately following high pitched syllable. As a whole, the HLLL sequence has H.L1.L2.L3f pitch realisation. It should be noted that the default contour shape of $HL^{(n)}$ sequence should differ depending on accentual types (Type A or Type B). Therefore, the quantification of this default contour shape enables us to predict the F0 realisation of the $HL^{(n)}$ sequence.

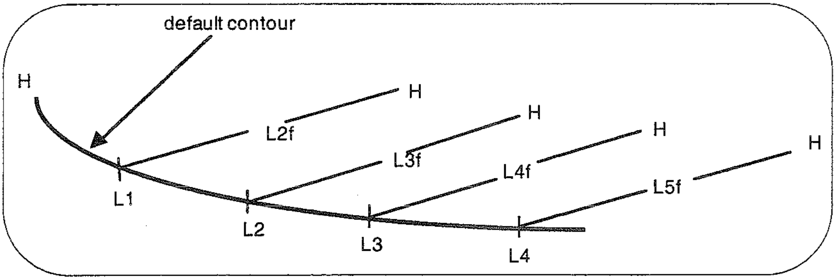

Figure 3: A schematic representation of default curve.

Further regarding the first point, the exponentiality of the default contour shape implies that the degree of F0 slope differs depending on the duration of the sequences concerned. This point certainly conforms to the target-tone model. Moreover, regarding to the second point, the identification of a certain base-line also implies that the base-line is the tone target.

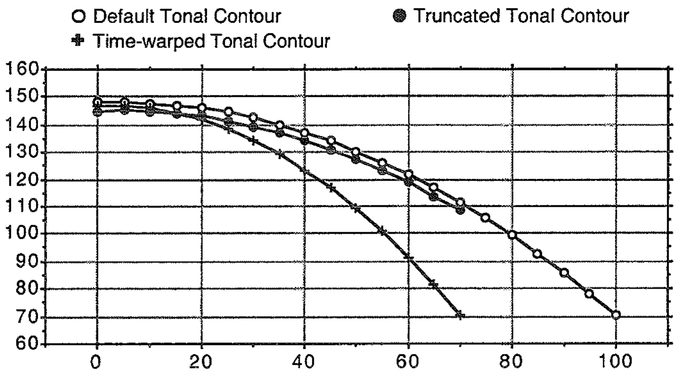

Figure 4: Two possible allotonic patterns.

Various tone languages show apparent allotony of pitch duration conditioned by their phonological structure, such as the absence or presence of a stop consonant in the coda position, the length of the nuclear vowel and so on. As far as the Rhyme length—on which a specific tone is realised—is concerned, mainly two allotonic realisational patterns are possible (Hombert et al, 1979; Rose, 1996).

These two patterns are schematically presented with an 'underlying' full-scale tonal contour in Figure 4 above. The F0 contour shape represented by empty circles is the 'underlying' full-scale realisation of a falling tone. The F0 contour shapes represented by filled circles and + symbols are possible allotonic realisations of the same falling tone.

One of the possible allotonic patterns is that a tonal contour can be truncated depending on the length of the Rhyme (filled circles). In this pattern, the truncated portion is realised congruently with the first portion of the contour representing the underlying tonal contour. The other pattern is that the underlying contour shape is time-warped so that the whole contour shape can be distributed over the shorter duration (+ symbols). Therefore, it can be said that the second pattern is a miniature case of the full-size realisation. It is plausible to consider that these two allotonic realisational patterns which are observed in tone languages can be extended to the phrase or sentence level tonal realisations.

It is assumed that all of the five F0 contour shapes presented in Figure 2, for example, are phonetically realised on the basis of the interpolation from a high target tone to a low target tone. The difference becomes apparent at the phonetic level where the two realisation patterns are possible depending on the typology of a particular language or a variation.

According to the data presented by Pierrehumbert and Beckman, SJ typologically belongs to the latter type where target values are achieved regardless of the duration. On the other hand, KJ appears to belong to the former type in which the F0 realisation of a certain tone is realised on the basis of a default contour shape.

ACKNOWLEDGMENTS

I would like to thank Dr Phil Rose and two anonymous reviewers for their valuable comments.

REFERENCES

Haraguchi, S. (1977) *The Tone Pattern of Japanese: An Autosegmental Theory of Tonology.* (Tokyo: Kaitakusha).
Higurashi, Y. (1983) *The Accent of Extended Word Structure in Tokyo Standard Japanese.* (Tokyo: Educa).
Hirayama, T. (1960) *Zenkoku akusento jiten.* (Tokyo: Tokyodo).
Hombert, J. M., Ohala, J. and Ewan, W. (1979) *Phonetic explanations for the development of tones.* Language 55: 1, 37-58.
Kawakami, S. (1962) *Japanese Accent Observed by the Aid of Pictograph.* Study of Sounds 10, 115-129.
Kobayashi, R. (1963) *On the Relation between Accent Forms and Speech Melodic Curves.* Bulletin of Phonetic Society of Japan 112, 5-8.
McCawley, J. D. (1977) *Accent in Japanese.* Studies in Stress and Accent: Southern California Occasional Papers in Linguistics 4, 261-302.
Pierrehumbert, J. and Beckman, M. (1988) *Japanese Tone Structure.* (Cambridge, Massachusetts, London, England: The MIT Press).
Rose, P. (1987) *Considerations in the Normalisation of the Fundamental Frequency of Linguistic Tone.* Speech Communication 6, 343-351.
Rose, P. (1996) *The Realisation of Stopped-Syllable Tones in Hua Sai and Pakphanang.* Proceedings of Sixth Australian International Conference on Speech Science and Technology , Adelaide, 605-610.
Vance, T. (1987) *An Introduction to Japanese Phonology.* (Albany: State University of New York Press).
Zhu, X. (1999) *Shanghai Tonetics.* Lincom Studies in Asian Linguistics 32. (Lincom Europe).