An Acoustic-Phonetic Descriptive Analysis of Pitch Realisations in Kagoshima Japanese

SHUNICHI ISHIHARA
Australian National University
Shunichi.Ishihara@anu.edu.au

1. Introduction

This paper has two aims. One is to acoustically-phonetically describe the word-level fundamental frequency (F0) realisations of Kagoshima Japanese (KJ) by means of z-score normalisation, the other is, on the basis of this description, to propose a tonal representation of KJ words, using the target-tone model. In the target-tone model, the surface pitch contour is considered to be produced by phonetic interpolation between target tones, including not only lexically stipulated tones but also tones associated with higher prosodic levels (cf. Beckman and Pierrehumbert, 1986; Pierrehumbert and Beckman, 1988; Venditti 1997, and more).

2. The Accentuation of Kagoshima Japanese

KJ exhibits a two way accentual contrast; $(L)^0HL$ and $(L)^0H$ (cf. Hirayama, 1960; Uwano, 1989). In the former pattern, only the penultimate syllable of a word has a high pitch, and every other syllable has a low pitch. 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. Following Hirayama (1960), $(L)^0HL$ type is referred to as Type A and $(L)^0H$ type as Type B in this paper. The examples shown in table 1 exhibit this accentual contrast.

Table 1: The accentual contrast of KJ.

Syllable Number	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"

The concept of mora is essential to explain the accentuation of standard Japanese (SJ). However, in KJ, the notion of syllable plays an important role in accounting for its accentuation. Thus, consider the following examples:

Table 2: Pitch realisations of KJ words. H and L are assigned to each mora. Periods are syllabic boundaries.

	Type A		Type B	
CVV.CVV	tai.zyuu HH.LL	"weight"	gai.too LL.HH	"light"
CVN.CVN	kin.sen HH.LL	"money"	kan.tan LL.HH	"simplicity"
CV.CVC.CV	ka.zat.ta L.HH.L	"decorated"	a.rat.ta L.LL.H	"washed"

In the above examples, a pitch change occurs not over a moraic boundary but over a syllabic boundary. This is why KJ is called a syllable-counting dialect (cf. McCawley, 1977; Kubozono, 1999; Shibata, 1962), and the tone bearing unit of KJ is syllable.

3. Framework

Traditional frameworks (McCawley, 1968, 1977, 1978; Haraguchi, 1977; etc) treat the accent of SJ lexically using a *diacritic* on a mora, and these traditional frameworks also share the same output, except for some minor derivational differences. That is, only the location of the characteristic pitch pattern is contrastive at the word level.

Pierrehumbert and Beckman (1988) introduced an entirely new method of describing Japanese pitch accent by using only a few tones—which are integrated in prosodic structure—with interpolation with them. The tones that Pierrehumbert and Beckman assign a phrase are limited to a boundary low tone (L%) which is assigned at the beginning of an utterance and at the end of each phrase, a phrasal high tone (H) which is normally assigned to the second mora of a phrase, and lexical accent tone (H*L) which links to a specific mora.

In traditional frameworks, a pair of words presented in (1) have the same pitch pattern at the word level, and the difference becomes apparent only when something (e.g. topic marker *-wa*) comes after those words. However, several instrumental experiments (cf. Pierrehumbert and Beckman, 1988; Kubozono, 1993; Poser, 1984; etc) reported that actual F0 contour between accented and unaccented words is significantly different even at the word level. As shown in (1), Pierrehumbert and Beckman's theory makes different representations for the words presented in (1) even at the word level.

(1)

	Accented	Unaccented
	"head-TOP"	"capital-TOP"
Traditional model	atama'-wa / | | \\ L H H L	mijako-wa / | | \\ L H H H
Target-Tone model	atama'-wa / | | \\ L% H H*L L%	mijako-wa / | \\ L% H L%

The target-tone model is adopted in this study for the tonal representation of KJ words.

4. Experiment

A main focus was put on two points to satisfy the first aim of this paper. The first point is the relationship between syllable structure and F0 realisation. The second point is differences in F0 realisation between Type A and Type B words. A corpus was compiled in light of these two points.

4.1. Experimental Procedure

Two native speakers of KJ (one female who is 25 years of age; FG and one male who is 27 years of age; AU) participated in this study. One corpus containing 85 target words (Refer to appendix for actual words) with approximately 60 dummy words which were scattered at random throughout the corpus was prepared. Table 3 shows the syllable structures and the number of these target words. I tried to select words which are used on a daily basis and I also tried to use both voiceless and voiced consonants and different vowels as evenly as possible to neutralise the consonantal perturbatory effect (cf. Gandour, 1974; Hombert, 1978) and intrinsic vowel F0 height (cf. House and Fairbanks, 1953; Lehiste and Peterson, 1961). The corpus was written in the usual Japanese orthography.

The informants were simply asked to read this corpus 5 times. They were also instructed to make sufficient pauses between each word. The recording was conducted in a reasonably quiet room in Sydney, and the reading material was recorded onto high-quality normal position tapes using a SONY TCM-5000EV tape recorder and a SONY ECM-D8 microphone. The raw material was digitised with Computerised Speech Laboratory (CSL) (sampling rate = 10000 Hz) at the Phonetics Laboratory of The Australian National University.

Table 3: The syllable structures and the number of target words. N is a syllable-final nasal.

Word	Type A	Type B
(C)VV	7	7
CVN	5	5
CVVCVV	6	6
CVNCVN	5	5
CVCV	4	4
CVCVCV	4	4
CVCVCVCV	5	5
CVCVCVCVCV	5	5
CVCVCVCVCVCV	1	1
CVCVCVCVCVCVCV	1	

4.2. Measurement Procedure

Wide band spectrograms of the resultant digitised tokens and aligned audio waveforms were created using the CSL. F0 was extracted by means of the Automatic Pitch Extraction Command of the same machine. As for those words having CV syllable structure, F0 samples were taken at every 20 percentage point of each syllable nucleus. For words having (C)VV or CVN syllable structure, F0 was sampled at every 10 percentage point of each syllable rhyme.

5. Normalisation

Different speakers have different acoustic outputs for what is perceived to be the same linguistic information. Therefore, in linguistic-phonetic studies, it is necessary to exclude 'between-speaker' differences and specify the invariant phonetic features of a certain language/variety by means of normalisation. The z-score normalisation—which Rose (1987, 1991) reports the superiority of in F0 normalisation—was used in this study in order to convert the individual speakers' outputs into linguistic-phonetic representations (cf. Ladefoged and Maddieson, 1986; Disner, 1980). The 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. Table 4 contains the normalisation parameters of the two informants.

Table 4: Normalisation parameters (Hz).

	x	SD
AU	133.01	21.8
FG	210.90	12.9

6. Syllable Structure and F0 Realisation

The influence of syllable structures (e.g. length, complexity, and so on) on the realisation of pitch targets has been examined in many languages (cf. Steel, 1986; Silverman and Pierrehumbert, 1990; House, 1989 for English; Gussenhoven and Rietveld, 1992 for Dutch; and more) In this section, I investigate how different syllable structures affect the F0 realisations of KJ's accentual oppositions. In this paper, disyllable words having CVCV, CVVCVV, and CVNCVN structures were used for actual comparison. Mono-syllable words having either (C)VV or CVN structure are also presented for reference.

6.1. Type A

Figure 1 contains the overall average normalised F0 values of Type A mono-syllable and disyllable words plotted against equalised duration.

Judging from figure 1, disyllable words having CVVCVV and CVNCVN structures show an almost identical F0 contour. What appears interesting is the difference in F0 range of the first syllable between disyllable words having CVCV structure and those having CVVCVV or CVNCVN structure. Disyllable words having CVCV structure appear to have a lower F0 range in the first syllable than those having either CVVCVV or CVNCVN structure. Although a visually clear difference is observed in the F0 range of the first syllable depending on syllable structures, the second syllables show an identical F0 contour.

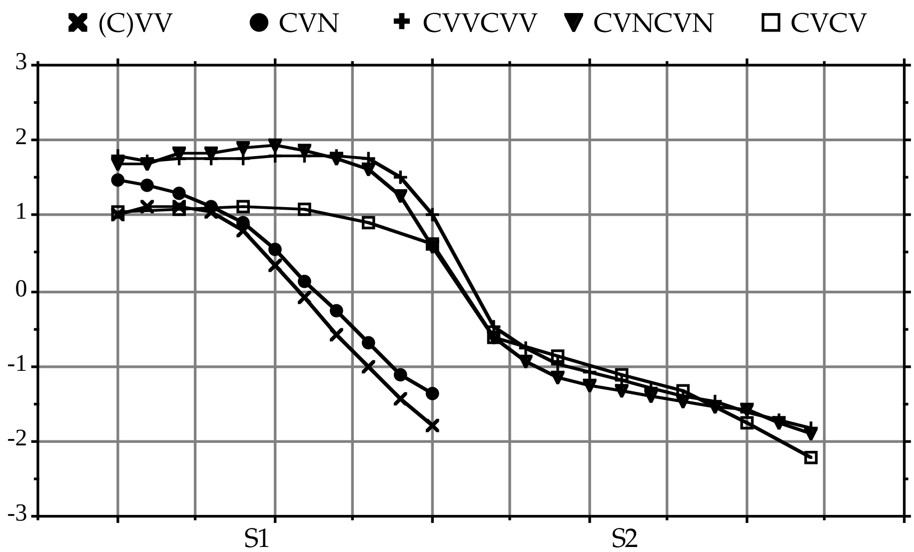

Figure 1: Average normalised F0 contours of Type A mono-syllable and disyllable words. X-axis is equalised duration and Y-axis is z-score value. S1 and S2 stand for the first and second syllable, respectively.

A one-way ANOVA followed by Scheffe F-test (level of confidence = 95%), was done on the peak F0 values of the syllables in question (and also the peak F0 values of mono-syllable words) to justify the above visual impression. The results are presented in table 5, in which shading means a significant difference. It is confirmed from Table 5 that long syllables have a significantly higher F0 realisation in the first syllable position than short syllables for both informants [F (4, 118) = 16.01, p < 0.0001 for FG; F (4, 112) = 35.25, p < 0.0001 For AU].

Table 5: The results of ANOVA and Scheffe-F test between (C)VV, CVN, CVCV, CVVCVV and CVNCVN at the peak value. The top table is for FG and the bottom one is for AU. Shading means significant difference.

FG	Count	Mean	SD
(C)VV	15	1.25	0.28
CVN	34	1.57	0.46
CVCV	16	1.72	0.36
CVVCVV	27	2.26	0.25
CVNCVN	25	2.32	0.32
F = 16.01			
p = 0.0001			

Scheffe F	CVN	CVVCVV	CVNCVN	CVCV
(C)VV	4.89	10.6	10.3	0.66
CVN		0.475	0.63	1.46
CVVCVV			0.23	9.94
CVNCVN				4.12

AU	Count	Mean	SD
(C)VV	31	0.95	0.70
CVN	20	1.66	0.52
CVCV	18	1.22	0.47
CVVCVV	30	1.88	0.57
CVNCVN	24	1.93	0.47
F = 35.25			
p = 0.0001			

Scheffe F	CVN	CVVCVV	CVNCVN	CVCV
(C)VV	2.02	18.88	20.53	3.24
CVN		13.82	15.55	0.463
CVVCVV			0.08	5.68
CVNCVN				6.76

6.2. Type B

Figure 2 contains the overall average normalised F0 values of Type B monosyllable and disyllable words plotted against equalised duration.

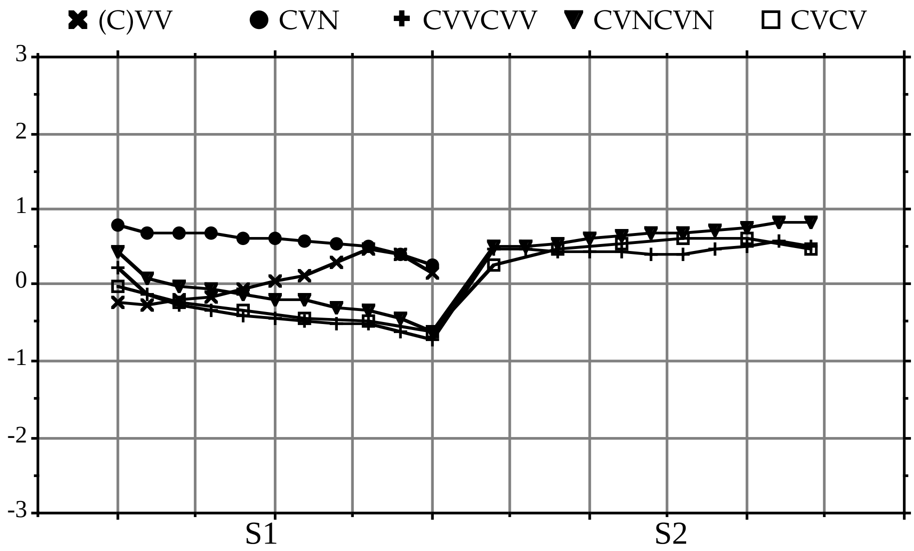

Figure 2: Average normalised F0 contour of Type B mono-syllable and disyllable words. The presentation procedure is the same as figure 1.

What you can observe from figure 2 is the difference in F0 realisation between mono-syllable words having (C)VV structure and those having CVN structure. Although these syllables seem to share the same offset value, the onset value appears to be different. A statistical analysis (unpaired, two-tailed t-test) confirms that CVN has a higher F0 than CVV from the onset to the 60% point for FG and from the onset to the 70% point for AU ($p < 0.05$). As a result, (C)VV shows a rising contour whereas CVN has a level contour.

Unlike Type A disyllable words, Type B disyllable words do not differentiate the F0 range depending on syllable structures.

6.3. Discussion

There are possibly two solutions to incorporate the difference in F0 range caused by syllable structures into the intonational theory of KJ. One solution would be to postulate that a CV syllable induces a slump in F0 range in comparison with CVV and CVN syllables. Alternatively, it would be equally possible to assume that CVV and CVN syllables induce a boost of F0 range in the position in question. The latter solution would be more appropriate mainly because CV syllable structure is the unmarked syllable structure in Japanese (cf. Kubozono, 1999). Given that CV syllable structure is the unmarked syllable structure, it would be natural to assume that F0 range is boosted from the unmarked F0 range by (C)VV and CVN marked syllable structures.

7. F0 Realisations of Type A and Type B

In this section, we look at how differently KJ's accentual oppositions are acoustically-phonetically realised.

The overall mean normalised F0 values of Type A and Type B poly-syllable words (which are in comparison) are graphically presented in figures 3, 4, and 5.

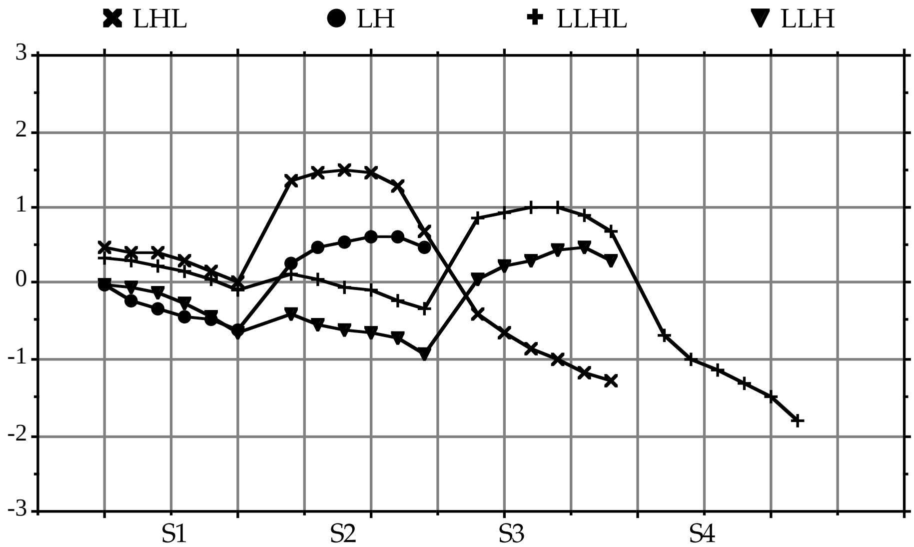

Figure 3: Average normalised F0 contour of LHL, LLHL, LH, and LLH. The presentation procedure is the same as figure 1.

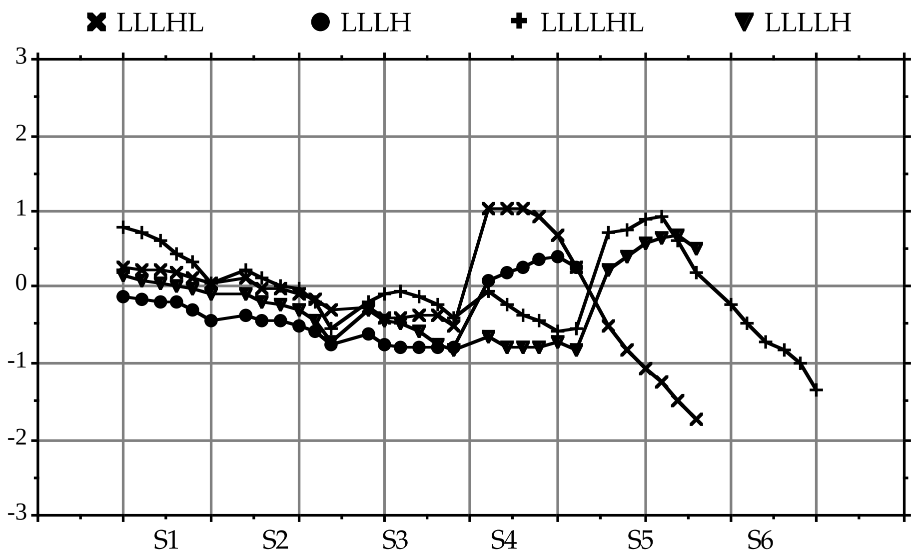

Figure 4: Average normalised F0 contour of LLLHL, LLLLHL, LLLH, and LLLLH. The presentation procedure is the same as figure 1.

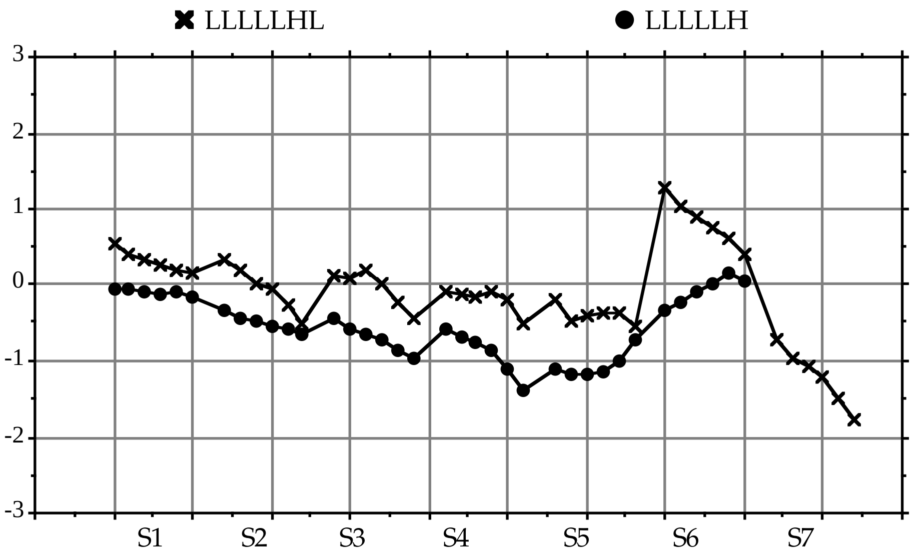

Figure 5: Average normalised F0 contour of LLLLLHL and LLLLLH. The presentation procedure is the same as figure 1.

A visual inspection of figures 3, 4 and 5 shows that Type A words are realised higher in F0 than Type B words.

The F0 realisations of Type A and Type B words were statistically compared at the points where tonal target values need to be specified. These points are graphically presented in (2). Tonal values need to be specified at four points (T1, T2, T3 and T4) to derive the F0 contour of Type A words and at three points (T1, T2 and T3) for Type B words. T1 is the F0 peak of a sequence of low pitched syllables and T2 is the F0 valley of the same sequence. T3 is the F0 peak of a high pitched syllable. In actual comparisons, Type A and Type B words having the same number of low pitched syllables before the high pitched syllable were compared at T1, T2, and T3 in order to factor out declination effect (cf. Vaissière, 1983). The values of T4 are presented only for reference.

(2) Type A: 7 syllable word Type B: 6 syllable word

```
 L L L L L H L          L L L L L H
/        |  |  \       /        |   \
T1       T2 T3  T4     T1       T2   T3
```

Table 6 contains the mean normalised F0 values and the SDs of them at T1, T2, T3 and T4, and also includes the individual results of t-test (unpaired, two-tail) conducted between those Type A and Type B words in comparison at T1, T2 and T3. According to Table 6, the mean values of Type A words are higher without any exceptions than those of corresponding Type B words at T1, T2 and T3. The difference between Type A and Type B words in F0 realisation was confirmed statistically significant ($p < 0.05$) in most comparisons at most comparing points (See table 6 where shading indicates a significant difference).

Table 6: The mean normalised F0 values of each comparative point (T1, T2, and T3), SD, and the individual results of t-test between Type A and Type B. Shading means significant difference. The values of T4 are only for reference.

	AU				FG			
	T1	T2	T3	T4	T1	T2	T3	T4
A 3 s	-0.12 (0.31)	-0.46 (0.36)	1.64 (0.24)	-0.86 (0.49)	1.33 (0.45)	0.56 (0.49)	1.60 (0.60)	-1.98 (0.84)
B 2 s	-0.57 (0.34)	-0.88 (0.31)	0.81 (0.33)		0.54 (0.52)	-0.42 (0.47)	0.70 (0.55)	
p	< 0.05	< 0.01	< 0.01		< 0.01	< 0.01	< 0.01	
A 4 s	-0.32 (0.25)	-0.73 (0.29)	1.24 (0.25)	-1.39 (0.47)	1.07 (0.62)	-0.10 (0.47)	1.14 (0.49)	-2.42 (0.80)
B 5 s	-0.53 (0.24)	-1.01 (0.21)	0.66 (0.31)		0.62 (0.44)	-0.96 (0.46)	0.48 (0.47)	
p	< 0.01	< 0.01	< 0.01		<0.05	<0.01	<0.01	
A 5 s	-0.21 (0.20)	-0.77 (0.17)	1.21 (0.36)	-0.84 (0.48)	0.95 (0.51)	-0.42 (0.43)	1.11 (0.52)	-2.71 (0.80)
B 4 s	-0.60 (0.29)	-1.04 (0.21)	-0.64 (0.34)		0.59 (0.73)	-0.83 (0.54)	0.56 (0.82)	
p	< 0.01	< 0.01	< 0.01		0.064	< 0.01	< 0.05	
A 6 s	-0.36 (0.30)	-0.95 (0.14)	1.02 (0.09)	-0.33 (0.30)	1.98 (0.18)	-0.36 (0.20)	1.17 (0.35)	-2.47 (1.12)
B 5 s	-0.42 (0.21)	-1.13 (0.22)	0.83 (0.35)		0.89 (0.60)	-0.94 (0.54)	0.75 (0.56)	
p	0.618	0.091	0.254		< 0.01	< 0.05	0.171	
A 7 s	-0.11 (0.36)	-0.76 (0.31)	1.24 (0.28)	-1.19 (0.53)	1.21 (0.83)	-0.50 (0.34)	1.31 (0.81)	-2.41 (0.99)
B 6 s	-0.21 (0.18)	-1.26 (0.19)	0.48 (0.30)		0.27 (0.18)	-1.41 (0.22)	-0.02 (0.18)	
p	0.595	< 0.05	< 0.01		< 0.05	< 0.01	< 0.01	

8. Intonational Structure of KJ words

In this section, a possible tonal representation of KJ words is presented on the basis of the above description, using the target-tone model. Since we have not known yet about the prosodic structure of KJ, only tonal values which are needed to distinguish Type A and Type B words at surface level are presented and discussed.

In SJ, the differences in F0 realisation between accented and unaccented words are considered to result from the lexical accent. Lexical accent boosts the F0 range of accented mora and pre-accented mora (if there is one), and moreover, causes a sharp F0 drop in post-accented moras. These phenomenon were named accentually-induced "F0 boost", "initial F0 rise", and "F0 fall", respectively by Kubozono (1993). It is interpreted in the target-tone model that these accentually-induced F0 phenomenon are caused by the lexically specified H*L tone.

It can be seen from the above description of Type A and Type B words that Type A words exhibit similar characteristics to the "F0 boost" and "initial F0 rise" of SJ accented words. Following this similarity between KJ and SJ, I tentatively posit four tones (H*L, H, L% and L) as shown below in the example.

(3) Type A: 7 syllable word / Type B: 6 syllable word

```
Type A: 7 syllable word                 Type B: 6 syllable word

   s  s  s  s  s  s  s                     s  s  s  s  s  s
  /              |  |   \                 /              |   \
 T1             T2 T3  T4                T1             T2   T3
 /               |  |     \              /               |    | \
L%               L  H*L    L%           L%               L    H  L%
```

Bitonal H*L, which is possibly a lexical tone, links to the penultimate syllable of Type A words. Similar to SJ, it is assumed that H*L tone is the one which induces the overall rise in F0 of Type A words. H tone is assigned to the final syllable of Type B words. L% is associated with the initial syllable and the final syllable of a word.

The value of T2 also needs to be specified to derive the contour shapes shown in figures 3, 4 and 5. The necessity of a target tone for T2 can be also understood from figure 6 as well.

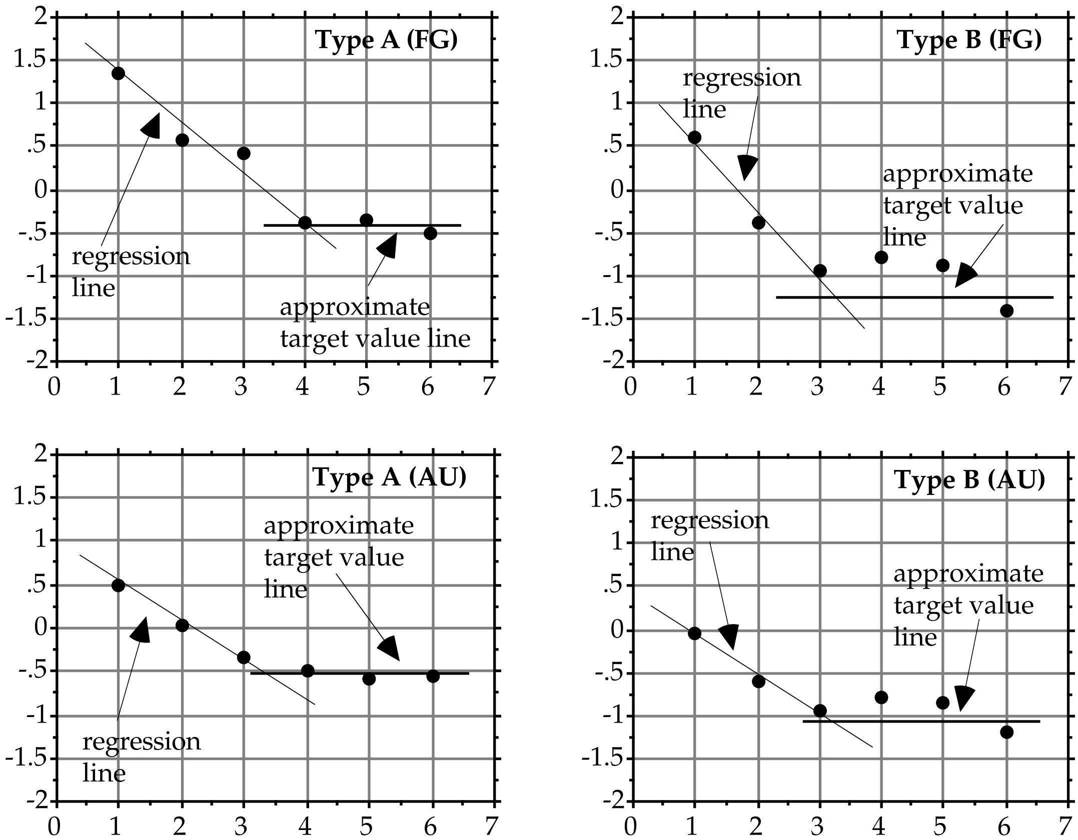

Figure 6: F0 transition of low pitched syllable sequence. X-axis is syllable number and Y-axis is normalised value. The top graphs are for FG and the bottom ones are for AU.

The four graphs of figure 6 show how F0 changes in the sequence of low pitch syllables. In the graphs, the average lowest normalised F0 values of each syllable nucleus before the high pitched syllable are plotted against syllable number. According to these graphs, F0 seems to decrease as a function of syllable number to a certain stage and it remains rather level after this stage. That means that there is a certain pitch target and it takes some amount of syllables to reach the target value. Therefore, I posit L tone as the target value for T2.

It has been reported in many languages that some factors, such as syllable structure, segmental duration, prosodic condition, rhythmic structure and so on, affect tonal alignment and the phonetic realisations of target tones (cf. Caspers and van Heuven, 1993a, 1993b; Arvaniti, 1994; Prieto, et al., 1995; Rietveld and Gussenhoven, 1995, and more). The syllable structure of the test words used for the comparison between Type A and Type B words is a simple CV, and several different consonants and vowels were used for the segments. Therefore, the influence resulting from different syllable structure is not relevant to this study, and possible effects on F0 realisation from segmental differences must have been neutralised. Yet, there is still a possibility that prosodic condition might have influenced the results of this study, particularly regarding the T3 value of Type B words. The T3 value of Type B might have been realised lower than that of Type A because the final syllable of Type B words is associated with a tone which causes sentence-final lowering, namely L% tone. Although we can not solely deny the possible interference of the L% tone, it seems to be still plausible to posit H*L tone in contrast with H tone so as to account for the fact that Type A words are realised higher in F0 than Type B words, not only at T3 but also at T1 and T2.

In this study, the above findings and discussion are limited to the word-level F0 realisation of KJ. In order to understand the comprehensive intonational structure of KJ, it is necessary not only to re-examine the above findings in a larger prosodic unit, but also to investigate other factors which possibly affect the tonal alignment patters of KJ.

9. Appendix

Type A				Type B			
(C)VV		CVN		(C)VV		CVN	
boo	stick	gin	silver	a i	love	dan	paragraph
ei	ray	kan	instinct	ao	blue	gun	local district
goi	vocabulary	kin	gold	au	to meet	ken	prefecture
ou	to chase	san	three	dai	topic	hon	book
suu	to suck	yon	four	gei	art	man	ten thousand
ue	up			ie	house		
uo	fish			i i	good		

Type A				Type B			
CVVCVV		CVNCVN		CVVCVV		CVNCVN	
zyuudoo	Judo	genkan	front door	gaitoo	street light	hansen	sail boat
keezai	economy	zinbun	humanity	kaitoo	answer	kantan	simple
kyooyuu	common owner	kansen	infection	keekai	lightly	konban	tonight
meekai	clear	yonban	No 4	keetoo	system	sinbun	news-paper
taiguu	treatment	yonman	forty thousand	kooryuu	exchange	sinkon	new couple
taizyuu	weight			kyoodai	brothers		

	Type A		Type B	
2	buta	pig	baka	stupid
	kagu	furniture	kagi	key
	mune	chest	kogu	to row
	take	bamboo	mimi	ears
3	kakuho	security	gakubu	department
	kimono	clothes	megane	glasses
	komimi	ears	tamago	egg
	sakana	fish	yubiwa	ring
4	gatabako	shoe box	demakase	lie
	kanimiso	crab paste	kakimono	document
	kagaribi	torch	miminari	ear noise
	kimusume	maid	nanohana	flower
	mamagoto	cubby house	yamanami	mountain tops
5	kamibasami	scissors	hanamizake	drinking while watching flowers
	kamiyasuri	sandpaper	kamikazari	hair-ornament
	matiyakuba	town hall	katatumuri	snail
	minamigawa	south side	mimizawari	irritating
	minamimura	south village	miyagemono	present
6	terebikamera	TV camera	niwatorigoya	hen cage
7	kamikiribasami	paper scissors		

References

Arvaniti, A. (1994) Acoustic features of Greek rhythmic structure. *Journal of Phonetics 22*. 239-268.

Beckman, M. E. and Pierrehumbert, J. B. (1986) Intonational Structure in Japanese and English. *Phonology Yearbook 3*. 255-309.

Caspers, J and van Heuven, V. J. (1993a) Perception of low-anchoring versus high-anchoring in Dutch accent-lending pitch rises. In House, D. and Touati, P. (eds.), *Proceedings of an ESCA Workshop on Prosody*. 188-191.

Caspers, J and van Heuven, V. J. (1993b) Effects of time pressure on the phonetic realisation of the Dutch accent-lending pitch rise and fall. *Phonetica 50*. 161-171.

Disner, S. (1980) Evaluation of vowel normalisation procedures. *Journal of the Acoustical Society of America 67*. 253-261.

Gandour, J. T. (1974) Consonant types and tone in Siamese. *Journal of Phonetics 2*. 337-350.

Gussenhoven, C. and Rietveld, T. (1992) A target-interpolation model for the intonation of Dutch. *In Proceedings of the International Conference on Spoken Language Processing 92*, University of Alberta, Canada. 1235-1238.

Haraguchi, S. (1977) *The tone pattern of Japanese: An autosegmental theory of tonology*. Tokyo: Kaitakusha.

Hirayama, T. (1960) *Zenkoku akusento jiten*. Tokyo: Tokyodo Shuppan.

Hombert, J. M. (1978) Consonant types, vowel quality and tone. In Fromkin, V. A. (ed.), *Tone: A linguistic survey*. New York: Academic Press. 77-111.

House, A. S. and Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journal of the Acoustical Society of America 25*. 105-113.

House, J. (1989) Syllable structure constraints on F0 timing. Poster presentation, *Second Conference of Laboratory Phonology*, Edinburgh.

Kubozono, H. (1993) *The organisation of Japanese prosody*. Tokyo: Kuroshio Publishers.

Kubozono, H. (1999) *Nihongo no onsei: Gendai gengogaku nyumon 2*. Tokyo: Iwanami Shoten.

Ladefoged, P and Maddieson, I. (1986) Some of the sounds of the world's languages. *UCLA working papers in phonetics 64*. 1-12.

Lehiste, I and Peterson, G. E. (1961) Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America 33*. 419-425.

McCawley, J. D. (1968) *The phonological component of a grammar of Japanese*. The Hague: Mouton.

McCawley, J. D. (1977) Accent in Japanese. In Hyman, L. (ed.), *Studies in stress and accent: Southern California Occasional Papers in Linguistics 4*. 261-302.

McCawley, J. D. (1978) What is a tone language? In Fromkin, V. A. (ed.), *Tone: A linguistic survey*. New York: Academic Press.

Pierrehumbert, J. and Beckman, M. (1988) *Japanese Tone Structure*. Cambridge, Massachusetts, London, England: The MIT Press.

Poser, W. J. (1984) *The Phonetics and Phonology of Tone and Intonation in Japanese*. PhD Dissertation, MIT.

Prieto, P., van Santen, J. and Hirschberg, J. (1995) Tonal alignment pattern in Spanish. Journal of Phonetics 23. 429-451.

Rietveld, T. and Gussenhoven, C. (1995). Aligning pitch targets in speech synthesis: effects of syllable structure. *Journal of Phonetics 23*. 375-385.

Rose, P. J. (1987) Considerations on the normalisation of the fundamental frequency of linguistic tone. *Speech Communication 6*. 343-351.

Rose, P. J. (1991) How effective are long term mean and standard deviation as normalisation parameters for tonal fundamental frequency? *Speech Communication 10.* 229-247.

Shibata, T. (1962) Shirabimu hogen mora hogen. In Kokugogakkai (ed.), *Hogengaku gaisetsu*. Tokyo: Musashinoshoin. 137-161.

Silverman, K. E. A. and Pierrehumbert, J. B. (1990) The timing of pre-nuclear high accents in English. In Kingston, J. and Beckman, M. E. (eds.), Papers in Laboratory Phonology I. Cambridge: Cambridge University Press. 72-106.

Steele, J. (1986) Nuclear accent F0 peak location: effect of rate, vowel, and number of syllables. *Journal of the Acoustical Society of America Supplement 1, 80*. s51.

Uwano, Z. (1989) Accent in Japanese. In Sugito, M. (ed.), *Koza nihongo to nihongo kyoiku 2: Nihongo no onsei, onin 1*. 178-205.

Vaissière, J. (1983) Language independent prosodic features. In Cutler, A. and Ladd, D. R. (Eds.), *Prosody: Models and Measurements.* Berlin: Springer-Verlag. 53-66.

Venditti, J. J. (1997). Japanese ToBI labelling guidelines. In Ainsworth-Darnell, K. and D'lmperio, M. (eds.), *OSU Working Papers in Linguistics 50: Papers from the Linguistic Laboratory*. 127-162.