

Surface Tonal Representations of Kagoshima Japanese Accentual Contrast

Shunichi Ishihara

Japan Centre (Asian Studies) & Phonetics Laboratory (Linguistics, Arts),
The Australian National University
Shunichi.Ishihara@anu.edu.au

Abstract

On the basis of the acoustic–phonetic descriptive results using polysyllabic words of Kagoshima Japanese (KJ), surface tonal representations of KJ accentual contrast are discussed and presented in this paper. The fundamental frequency realisation differences of KJ accentual contrast are also accounted for from the surface tonal representations presented in this paper, and the account is justified.

1. Introduction

Kagoshima Japanese (KJ) exhibits a two way accentual contrast; L^nHL and L^nH where $n =$ positive integers or zero (Hirayama 1960). In the former pattern, only the penultimate syllable of a word has a high pitch, and every other syllable has a low pitch (i.e. *sakura* LHL, 'cherry blossom' and *kagaribi* LLHL, 'watch fire'). 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 (i.e. *usagi* LLH, 'rabbit' and *kakimono* LLLH, 'document'). Following Hirayama (1960), the L^nHL type is referred to as Type A and the L^nH type as Type B in this paper.

As can be seen from KJ's Type A and Type B contrast, although unlike Standard Japanese (SJ), there is not a lexical contrast in terms of the location of an accent in KJ, the main contrast between Type A and Type B is whether there is a fall in pitch at the end or not. Like the accentual behaviour of SJ (Poser 1984; Pierrehumbert and Beckman 1988; Kubozono 1993), however, recent instrumental approaches to KJ prosody revealed that Type A initiates not only global rises but also drops in F0 (Kubozono and Matsui 1996). The results of the above instrumental research indicate that the mapping from phonological pitch description, whereby a sequence of conventional (H)igh and (L)ow are used, to phonetic F0 realisation is not straightforward.

Taking Kubozono and Matsui's (1996) findings—which are summarised in §1.1.—as the first step, the current study further investigates the F0 realisation difference between Type A and Type B at the word level.

It will be demonstrated in this paper that Type A words are associated with higher F0 values than corresponding Type B words throughout its entire time course.

On the basis of the descriptive results, surface tonal representations of Type A and Type B words are discussed using the Autosegmental–Metrical (AM) theory of intonation (Pierrehumbert and Beckman 1988). Furthermore, how the F0 realisations of KJ's accentual contrast can be derived from the surface tonal representations is also accounted for.

1.1. Previous studies on Kagoshima Japanese

Kubozono and Matsui (1996) reported from their experiment on KJ using several sentences and phrases, that the syllable having a high phonetic pitch in a Type A word was realised with a higher F0 than that of a Type B word

(i.e. Type A [LL↑HL] and Type B [LLLH]). However, since KJ's accentual contrast was compared only at intonational target points in their experiment, it is not yet known what the whole F0 time courses of Type A and Type B words look like. In addition to this, in contrast to Kubozono and Matsui, in which phrases and sentences were used for comparisons, this study will investigate whether the difference in F0 realisation between Type A and Type B that Kubozono and Matsui reported is also observable at the word level. Another point that needs to be made regarding Kubozono and Matsui's experiment is that the possible effect of declination (Vaissière 1983) was not controlled in their experiment because the F0 sampling points were not time aligned. The present study rectifies this.

1.2. Autosegmental–metrical model

Pierrehumbert and Beckman (1988) argue that the surface pitch contour is produced by phonetic interpolation between tone targets, including not only the lexically stipulated tone but also tones associated with higher levels of prosodic structure. On the basis of this idea, they introduced an entirely new model of describing Japanese pitch accent by using a few tones per phrase with interpolation between them. The tones assigned to a phrase are: the low boundary tone (L%) at the onset of an utterance, the high phrasal tone (H) that links the second sonorant mora, and the boundary tone (L%) that links the last mora. At the word level, a lexical accent (H*L) is attached to the designated mora. Besides the notion of target tones, prosodic hierarchy has also been agreed upon by many linguists working on Japanese intonation (Pierrehumbert and Beckman 1988; Selkirk and Tateishi 1988).

1.3. Experimental procedure and normalisation

Twenty Type A polysyllabic words (4 disyllabic, 4 trisyllabic, 5 four-syllabic, 5 five-syllabic, 1 six-syllabic and 1 seven-syllabic words) and nineteen Type B polysyllabic words (4 disyllabic, 4 trisyllabic, 5 four-syllabic, 5 five-syllabic, and 1 six-syllabic words) were used in this study. The syllable structure of these words is CV.

A corpus was compiled using these polysyllabic words and monosyllabic words together with some dummy words which were scattered at random throughout the corpus. The informants (2 males and 2 females) were asked to read this

corpus 5 times with a sufficient pause between each word. The recording was conducted in a room in Kagoshima City 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 using Computerised Speech Laboratory (sampling rate = 10 kHz).

Three F0 measurement procedures were used for the analysis. In the first procedure (Procedure 1), F0 was sampled at the onset and every 20% point of the duration of each syllable nucleus. This procedure makes it possible to show in detail the F0 time course in each word. In Procedure 2, F0 was sampled at the 50% point of the duration of each syllable nucleus. Procedure 2 enables us to collect F0 values that are the least consonantly perturbed. In Procedure 3, the minimum F0 value was sampled from each low pitched syllable nucleus and the maximum F0 value from the high pitched syllable nucleus. The F0 values sampled in Procedure 3 are possible target values associated with the syllables.

The logarithmic z-score normalisation was used in this study in order to exclude between-speaker differences and specify the invariant features of KJ's accentual contrast (Rose 1987). The log z-score normalisation procedure is: $F0_{norm} = \frac{F0_i - \bar{x}}{sd}$, where $F0_i$ is a sampling point, \bar{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 in logarithmic terms. Table 1 contains the normalisation parameters for each informant.

Table 1: *The average log F0 (\bar{x}) and standard deviation (sd) from all sampling points for each informant.*

informant	sex	\bar{x}	sd
1	Male	2.012	0.060
2	Male	1.994	0.037
3	Female	2.263	0.088
4	Female	2.211	0.054

The F0 realisations of Type A and Type B on polysyllables were statistically compared on the basis of the log z-score normalised F0 values which were converted from the raw F0 values sampled by means of Procedures 2 and 3. In actual comparison, Type A and Type B polysyllabic words having the same number of L pitches before the H pitch were compared in order to factor out the declination effect.

That is, five different comparisons—Type A three syllable word vs. Type B two syllable word; Type A four syllable word vs. Type B three syllable word; Type A five syllable word vs. Type B four syllable word; Type A six syllable word vs. Type B five syllable word; and Type A seven syllable word vs. Type B six syllable word—are possible from the data. In the case of comparing a Type A seven syllable word and a Type B six syllable word, for example, a t-test (unpaired, two-tailed) was conducted at six different locations as shown in Fig. 1 for both sets of values collected by means of Procedure 2 and Procedure 3. The values of the syllables having the identical index number were compared between Type A and Type B. That is, a pair of comparable

Type A and Type B words was compared one syllable to one syllable starting from the left.

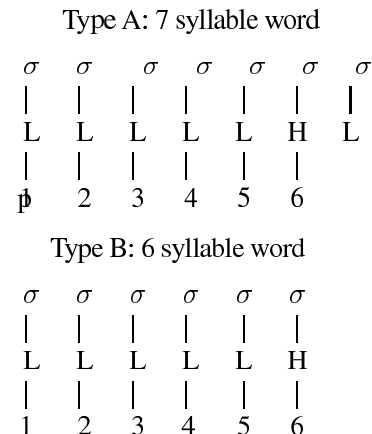


Figure 1: *Statistical comparison points between Type A and Type B words.*

In this paper, the syllable associated with a high pitch (= the penultimate syllable for Type A and the ultimate syllable for Type B) is referred to as the peak syllable, and the syllables appearing before the peak syllable as pre-peak syllables. The final syllable of Type A words is referred to as the post-peak syllable.

2. Results of normalised data

Fig. 2 contains the mean normalised F0 contours of Type A five syllable (●) and Type B four syllable words (○). In addition to the fact that Type A falls in F0 at the end and Type B does not, other differences between Type A and Type B are not only that as demonstrated by Kubozono and Matsui (1996), the peak syllable of Type A is associated with higher normalised F0 values than that of Type B, but also that all the pre-peak syllables of Type A are associated with higher normalised F0 values than those of Type B. The same visual observation can be made from the other pairs which are in comparison.

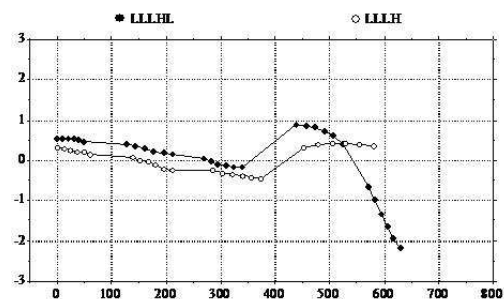


Figure 2: *Mean normalised F0 contours of Type A 5 syllable (●) and Type B 4 syllable (○) words.*

An interesting point that can be recognised in Fig. 2 is that normalised F0 declines around the mean (= 0) in the pre-peak syllables. That means that although the pre-peak syllables are phonologically associated with a low pitch, it is more accurate to say that they are associated with a mid

phonetic pitch from the narrow transcription and perceptual points of view.

The above visually observed differences between Type A and Type B need to be confirmed statistically. The F0 realisations of Type A and Type B on polysyllables were statistically compared on the basis of the log z -score normalised F0 values which were converted from the raw F0 values sampled by means of Procedures 2 and 3.

The statistical comparisons confirmed that Type A words have a significantly higher normalised F0 realisation than the corresponding Type B words ($p < 0.05$) at all comparison points (except a couple of points in the longest words). At the majority of the comparison points, Type A and Type B words are significantly different at $p \leq 0.001$ level. That is, although the traditional view is that the difference between Type A and Type B is only whether there is a pitch fall or not at the end, the results show that the accentual contrast is realised throughout the entire duration of a word.

On the basis of the above descriptive results, surface tonal representations of KJ's accentual contrast will be discussed in the following section.

3. Surface tonal representation of Kagoshima Japanese

By referring to the descriptive results, the location and the values of the target tones that are necessary to represent Type A and Type B words in the AM theory will be discussed in this section.

In order to derive the F0 contours which are typical to Type A and Type B polysyllabic words, such as those given in Fig. 2, target tones need to be specified at (at least) four points (T1, T2, T3 and T4) as shown in Fig. 3.

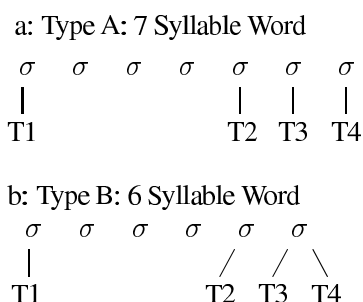


Figure 3: *Target tone locations* using Type A 7 syllable and Type B 6 syllable words as examples.

T1 and T2 correspond to the onset and the offset of the pre-peak syllables and T3 corresponds to the peak syllable. T4 is the offset boundary tone, and it needs to be associated with the ultimate syllable of both Type A and Type B words.

T2 needs to be specified because the descriptive results—such as in Fig. 2—do not show a linear contour from T1 to T3, but show a descending contour from T1 to T2 before it starts rising from T2 to T3.

On the basis of the descriptive results of Type A and Type B polysyllabic words, I posit the ones given in Fig. 4 as the surface tonal representations of KJ's accentual con-

trast. In Fig. 4, Type A 7 syllable and Type B 6 syllable words are used as examples.

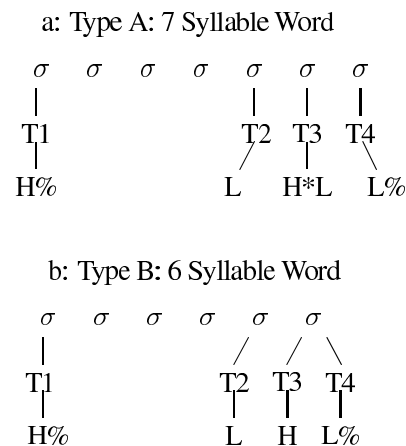


Figure 4: *Surface tonal representations of Type A 7 syllable and Type B 6 syllable words.*

In the following subsections, these proposed target tones are justified. The justification of each target tone is discussed in the order of T3, T4, T2 and T1 in the following subsections.

3.1. T3 tone

It is apparent that the most important tones that determine the F0 realisation difference between Type A and Type B are the target-tones of T3 (peak syllable) for Type A and Type B. As far as the paradigmatic F0 realisation differences between Type A and Type B words are concerned, it is not necessary to posit different target tones for Type A and Type B as T3 because the realisation differences can be phonetically accounted for without assigning different target tones (Refer to §3.5.). However, as Kubozono and Matsui (1996) demonstrated, the sequence of Type A and Type A words (i.e. LHL.LHL, period = accentual boundary) is realised differently in F0 from that of Type B and Type A words (i.e. LH.LHL) in that the second accentual unit (Type A) is realised lower in the former combination (AA) than the latter combination (BA). This point needs to be accounted for as well.

Therefore, two different tones: H*L tone for Type A and H tone for Type B have been decided to posit as T3. The H*L tone represents the falling nature of Type A and the H tone the non-falling nature of Type B. Furthermore, the falling associated with the H*L tone phonetically induces the F0 realisation differences between Type A and Type B (Refer to §3.5.).

One may think of the possibility that the lower F0 value of the Type B peak syllable compared to the Type A peak syllable may be due to the association of a boundary tone (L%) to the Type B final syllable. Namely, the L% tone associated with the peak syllable of Type B pulls the F0 realisation down. If this is true, there may be no need to assign different T3 tones for Type A and Type B words because the F0 realisation difference between Type A and Type B words at the peak syllable is considered to have resulted from the final L% that is attached to the final syllable of

Type B words. However, it has been decided to assign different tones as T3 between Type A and Type B words because the difference observed at the pre-peak syllables also needs to be accounted for.

3.2. T4 tone

Although the current data does not contain relevant information to discuss the L% tone, an offset L boundary tone is in fact easily observable from the F0 contour of a long utterance. Fig. 5 contains the audio speech waveform and F0 contour of a long utterance. In Fig. 5, the low pitched syllables (which are circled) before a pause have significantly lower F0 values than other low pitched syllables even if the effect of declination is taken into consideration. Therefore, the L% tone is posited as T4 in KJ.

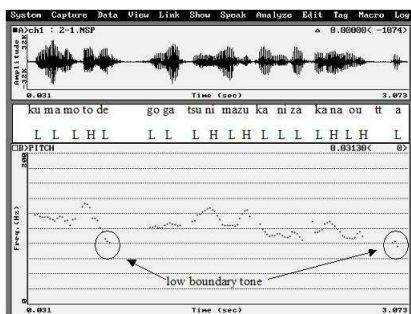


Figure 5: Audio speech waveform and F0 contour of 'kumamoto-de gogatsu-ni mazuka nizakana-o utta' (I sold badly cooked fish in Kumamoto in May).

3.3. T2 tone

Judging from the typical F0 contour shapes of Type A and Type B polysyllabic words, such as given in Fig. 2, I have discussed above that T2 needs to be specified at the offset of the pre-peak syllables. I will further argue for the necessity of T2 with some evidence based on the descriptive data in this study.

Fig. 6 shows how the normalised F0 time course is realised in the pre-peak syllables and also how the length of a word affects the F0 realisation on the peak syllable and the post-peak syllable. In order to draw the scatter graphs of Fig. 6, the mean normalised F0 values were calculated from the F0 samples collected by means of the syllables of each target word for each informant, and these mean normalised F0 values were then plotted against the mean absolute sampling points. These mean normalised F0 values were plotted separately between the pre-peak syllables, the peak syllable and the post-peak syllable. This was repeated for all target words of the informants for both normalised F0 values based on Procedures 2 and 3.

Since the results based on Procedures 2 and 3 do not differ in any essential ways and Type A and Type B show essentially the same result, only the one from Procedure 2 are given in Fig. 6 for Type A. However, as with the post-peak syllable, the normalised F0 value based on Procedure 3 are used because the minimum value of the Type A final syllable is considered to be associated with the final target value. This is because a clear F0 falling shape is observed in the final syllable of Type A words.

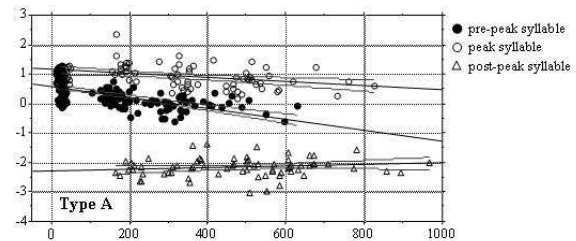


Figure 6: Scatter plot showing normalised F0 values appearing in pre-peak, peak and post-peak syllables as a function of word length.

It is noticeable from Fig. 6 that the most decay is shown in the pre-peak syllables (●); a little in the peak syllables (○); and none in the final syllables (△). An important point is that the rate of normalised F0 decay observed in the pre-peak syllables is greater than that observed in the peak syllables. If it had been simply the case that the normalised F0 decay observed in both the pre-peak syllables and the peak syllables was due to the mechanical declination effect, the observation made from Fig. 6 should have been reversed. That is, larger normalised F0 decay would have been observed in the peak syllables than the pre-peak syllables (Vaissière 1983). This implies that the normalised F0 decay observed in the pre-peak syllables is not simply due to the declination effect, but is possibly due to some other effect. If it is considered that the sharper F0 decay appearing in the pre-peak syllables than in the peak syllables is understandable. Therefore, L tone is posited as T2.

3.4. T1 tone

The value of T1 appears to need extensive discussion. Simply following the phonological pitch configurations of Type A and Type B words (i.e. LLLLHL and LLLLH), one might posit L% tone as T1. However, a few points need to be addressed regarding this.

First of all, the onset normalised F0 value of Type B words is around the mean (= 0), as can be seen in Fig. 2 in all results. That means that the onset F0 value of Type B words is around the average F0 of the informants' F0 distribution.

The F0 realisation of intonational tones is one of the main issues in the phonology of intonation. The question I would like to address here is 'Is it phonologically plausible to assign L% to the syllable whose F0 realisation is around the average of a speaker's F0 distribution?'

It is possible to interpret the F0 associated with the onset of Type B words as the most neutral F0 value that is around the mean F0 value of a speaker's F0 distribution. Considering this, it might not be necessary to assign a particular tone, such as L% for T1 because the F0 realisation of the initial syllable of Type B words is predictable from the information regarding a speaker's F0 distribution. A very similar interpretation can be seen in Dainora (2001) in the intonation of English. Referring to the analysis on the intonation of English by Goldsmith (1978) in which the

intonation was analysed in terms of H, M and L, Dainora (2001) mentions that 'the M does not represent a tone at all, but, rather, represents the neutral frequency used from the beginning of an utterance to the first pitch accent.' In Goldsmith's analysis, M tone is always the first tone in the sequence and is not associated with an accented syllable. If the same discussion applies to KJ, it can be argued that no particular tone needs to be associated with the T1 position.

Prima facie, the above 'no tone analysis' has a degree of plausibility. However, it causes a problem in the light of the report made in Ishihara (2004). I reported regarding the F0 realisation of Type B in disyllabic words that although they share the same offset value in the first rhyme, a moderate F0 falling contour is observed in the first rhyme of Type B heavy disyllabic words—((C)VV.(C)VV and (C)VN.(C)VN where a period stands for a syllabic boundary)—while a level F0 contour is observed in the first rhyme of Type B light disyllabic words—((C)V.(C)V). I discussed that the F0 realisation observed in the light rhyme type is an undershot realisation (\downarrow T%) of the full (=unmarked) realisation of the initial boundary tone (T%) due to the shortness of the light rhyme. This is because 1) the undershooting of a tone due to the weight (= shortness) of a syllable is a commonly observed phenomenon in KJ (Ishihara 2004) and 2) no plausible reason can be provided for the F0 boosting associated with heavy syllables when the F0 realisation of light syllables is assumed to be the unmarked F0 realisation of Type B on disyllables.

In this study, all of the polysyllabic target words have light syllables, and as observed, Type B polysyllabic words having light syllables start around a speaker's average F0 value. Based on this observation, it has been discussed above that Type B polysyllabic words having light syllables do not have to have any particular tone at the onset. If it is so assumed, the F0 realisation of Type B light disyllabic words can be considered as the unmarked F0 realisation of Type B on disyllables, and the F0 realisation of Type B on heavy disyllables needs to be considered as a deviation from the unmarked F0 realisation. If the F0 realisation of Type B in light disyllabic words is assumed to be the unmarked F0 realisation of Type B on disyllables, the deviated (or boosted) F0 realisation of Type B on heavy disyllables in the first rhyme needs to be accounted for. However, as mentioned above, no plausible reason for the boosting can be found. Furthermore, even if there were a plausible explanation for the boosting, we would still encounter difficulty in explaining how something that does not exist can be boosted. Therefore, it appears to be more plausible at this stage to assign a tone for T1.

At the beginning of the discussion regarding T1, I questioned the plausibility of assigning the L% tone to the onset syllable. Although Type B polysyllabic words start with around the average of a speaker's F0 distribution, the average F0 value at the onset is quite high judging from the informants' F0 range.

Fig. 7 contains the histogram of the normalised F0 values based on Procedure 3. As can be seen from Fig. 7, the histogram has a longer skirt to the negative side from the mean value (= 0). That is, the actual mid point of a speaker's F0 range is lower than the mean value of their F0

distribution. As can be seen by this, the H% tone rather than the L% tone is considered to be appropriate for the tone assigned at the T1 position.

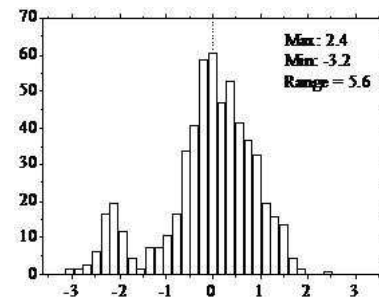


Figure 7: Histogram of the normalised F0 values measured in Procedure 3.

3.5. Phonetic treatment of F0 realisation differences between Type A and Type B

The F0 realisation difference between Type A and Type B words can be phonologically accounted for by assigning an F0 boosting role to, for example, the abstract concept of an accent (Haraguchi 1977). That is, an accent, which is a property of Type A, boosts the F0 realisation of Type A words (i.e. [\uparrow L \uparrow H'L] Type A vs. [LLH] Type B, where an apostrophe stands for the location of an accent).

Paying attention to the word-final position of KJ words, the F0 realisation difference between Type A and Type B are very similar to Tone 1 (high level = H) and Tone 4 (high falling = HL) of Mandarin Chinese (Ladefoged 2003). Tone 4 starts with higher F0 values compared to Tone 1 and also Tone 4 falls in F0 drastically towards the offset. In Mandarin Chinese, this F0 difference between Tone 1 (H) and Tone 4 (HL) is a phonetic realisation difference, and it can be explained that Tone 4 starts with a higher F0 compared to Tone 1 in order to enhance the perceptual salience of the falling contour.

Since the prime difference between Type A and Type B is whether there is a pitch fall or not after a pitch peak, it would be sensible to consider that as in the case of Mandarin Chinese, the F0 realisation of Type A at the peak syllable is raised in order to perceptually enhance the salience of the falling contour of Type A. If it is so assumed with the peak syllable, it can be considered that the higher F0 realisation of Type A observed at the pre-peak syllables is an anticipatory assimilation to the resultant enhanced F0 realisation at the peak syllable of Type A. This point is explained in Fig. 8.

Arrow 2 is the vector of the enhancement of the F0 realisation at the peak syllable of Type A for the perceptual salience of the falling contour. Arrow 1 is the vector of the assimilation of the F0 realisation at the pre-peak syllables of Type A to the enhanced F0 realisation at the peak syllable. Therefore, the realisation difference between Type A and Type B words can be phonetically accounted for without assigning the boosting role to, for example, accent.

One possible argument against the assimilation analysis of the pre-peak syllable to the enhanced F0 realisation at the peak syllable can be drawn from a well-established

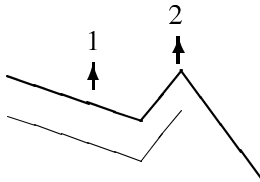


Figure 8: *Phonetic explanation for the realisation difference between Type A and Type B from the viewpoint of the enhancement of perceptual salience.*

tonal phenomenon of tone languages. Using disyllabic words as examples, in tone languages, the tone of the second syllable assimilates to the tone of the first syllable, and the tone of the first syllable dissimilates from the tone of the second syllable with respect to F0 height (Gandour et al. 1994; Thompson 1997). The rules given in Fig. 9 explain this.

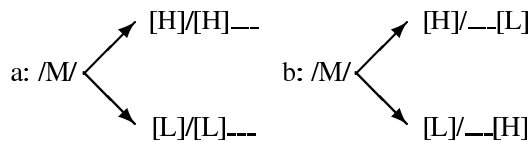


Figure 9: *Tonal assimilation and dissimilation.*

In Fig. 9a, the /M/ is realised as [H] if the preceding tone is [H] whereas it is realised as [L] if the preceding tone is [L]. In Fig. 9b, the /M/ is realised as [H] if the following tone is [L] whereas it is realised as [L] if the following tone is [H]. If the above rules are applied to KJ, the F0 realisation of Type A at the pre-peak syllables should have been dissimilated from the enhanced F0 realisation at the peak syllable. However, one may possibly counter-argue that the dissimilation rule of Fig. 9b does not apply to KJ (and SJ) because those rules given in Fig. 9 are only applicable when two lexical tones are involved. Although at this stage we cannot be certain of the lexical status of KJ's accentual contrast, judging from the similarities between KJ and SJ, it appears to be that only the H*L tone is the lexical tone in KJ. Therefore, one could possibly argue that the dissimilation rule given in Fig. 9b does not apply to the tones (H% and L) associated with the pre-peak syllables of Type A words because those tones are not lexical tones.

4. Conclusion

In this paper, I have demonstrated that in contrast to the traditional assumption, the F0 realisation of Type A and Type B words is globally different throughout its time course. Type A words are associated with higher F0 not only at the peak but also the pre-peak syllables compared to Type B words. On the basis of the descriptive results, I have proposed how Type A and Type B words can be represented at the surface level. I have also argued that the F0 realisation differences between Type A and Type B can be phonetically explained in the light of the perceptual salience without assigning any roles to accent, which is a phonological abstract concept.

5. Acknowledgments

This research was supported by the Faculty of Asian Studies' Research Support Funds (ANU). I thank Dr Phil Rose at the ANU for giving valuable comments for this research.

References

- Dainora, A. (2001). *An Empirical Based Probabilistic Model of Intonation in American English*. Ph.D. dissertation, University of Chicago.
- Gandour, T., S. Potisuk, and S. Dechongkit (1994). Tonal coarticulation in Thai. *Journal of Phonetics* 22, 477–492.
- Goldsmith, J. (1978). English as a tone language. *Communication and Cognition* 11, 453–476.
- Haraguchi, S. (1977). *The Tonal Pattern of Japanese: An Autosegmental Theory of Tonology*. Tokyo: Kaitakusha.
- Hirayama, T. (1960). *Zenkoku akusento jiten [The Accent Dictionary of Japanese Dialects]*. Tokyo: Iwanami shoten.
- Ishihara, S. (2004). *An Acoustic-Phonetic Descriptive Analysis of Kagoshima Japanese Tonal Phenomena*. Ph. D. thesis, The Australian National University.
- Kubozono, H. (1993). *The Organization of Japanese Prosody*. Tokyo: Kuroshio Publishers.
- Kubozono, H. and M. Matsui (1996). Nihongo no akusentojakka to intoneeshon [The downstepped accent of Japanese and intonation]. In *Proceedings of the Tenth General Meeting of the Phonetic Society of Japan*, pp. 101–106. The Phonetic Society of Japan.
- Ladefoged, P. (2003). Phonetic fieldwork. In J. S6le, D. Recasens, and J. Romero (Eds.), *Proceedings of the Fifteenth International Congress of Phonetic Sciences*, pp. 203–206.
- Pierrehumbert, J. and M. Beckman (1988). *Japanese Tone Structure*. Cambridge, Massachusetts, London, England: MIT Press.
- Poser, W. (1984). *The Phonetics and Phonology of Tone and Intonation in Japanese*. Ph. D. thesis, MIT.
- Rose, P. (1987). Considerations in the normalisation of the fundamental frequency of linguistic tone. *Speech Communication* 6, 343–351.
- Selkirk, E. and K. Tateishi (1988). Constraints on minor phrase formation in Japanese. In *Proceedings of the Chicago Linguistic Society* 24, pp. 316–336.
- Thompson, N. (1997). *Tone and Tone Sandhi in Southern Thai: An Acoustic Description of the Seven Tone Dialect of Ron Phibun*. Honours thesis, The Australian National University.
- Vaissi6re, J. (1983). Language independent prosodic features. In A. Cutler and R. Ladd (Eds.), *Prosody: Models and Measurements*, pp. 53–66. Berlin: Springer-Verlag.