The Flat Type and Flattening in Japanese

The Flat Type and Flattening in Japanese

Nobuo Yuzawa

Keywords: Accent, Flat type, Declination, F0 adjustment, Flattening, Tonic mora

0. Introduction

Accents play an important role in spoken Japanese[1], and they are divided into two types: the flat type and the non-flat type. The latter type is subdivided into three categories: the end-high type, the mid-high type and the head-high type[2]. The phonological difference between the flat type and the end-high type lies in the pitch movement of the following post-position ('joshi'). When the post-position is spoken with the same pitch as the last mora of the preceding noun, this noun belongs to the flat type. In contrast, when the pitch of the post-position is lowered in relation to the last mora of the preceding noun, this noun belongs to the end-high type. The purpose of this paper is to examine the f0 contour of the flat type to see whether its phonological description is true to the phonetic reality. This approach is expanded to other types to look into their phonetic reality, especially at utterance-ends. Comparison is also made of this type with the non-flat types, especially the end-high type. The type of Japanese we study here is standard Tokyo Japanese, and all the data used in this paper are taken from NHK Nihongo Centre (1994) and NHK Hosobunka Kenkyujo (2002). This means that these data are spoken by professional NHK announcers.

1. The Flat Type

The flat type is widely used in nouns, adjectives and verbs in Japanese. The basic feature of this type is that the first mora is low and the remaining morae are high level

[1] Accents used in Japan are not universal. According to NHK Hosobunka Kenkyujo (1998), they are divided into four types. Tokyo Japanese belongs to Type II. In some parts of Japan (Tochigi, Ibaraki, Fukushima, Miyagi, Miyazaki, etc.), such accentual differences are lost.

[2] The flat type, the non-flat type, the head-high type, the mid-high type and the end-high type are the author's translation of 'heiban-gata', 'kifuku-gata', 'atamadaka-gata', 'nakadaka-gata' and 'odaka-gata', respectively.

including the following post-position. This type is commonly displayed as in Figure 1.

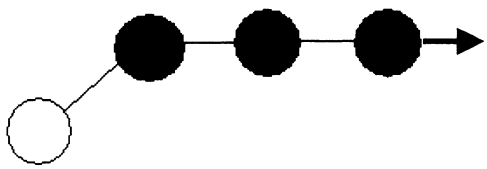
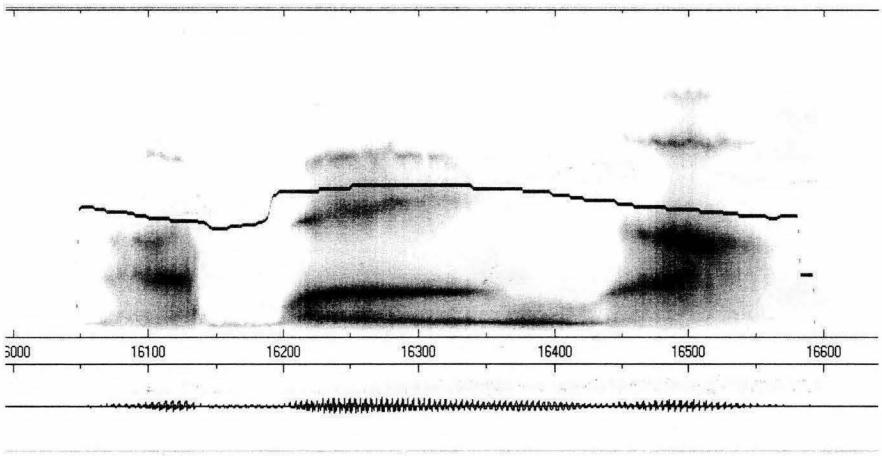


Figure 1 This shows a phonological description of the flat type.

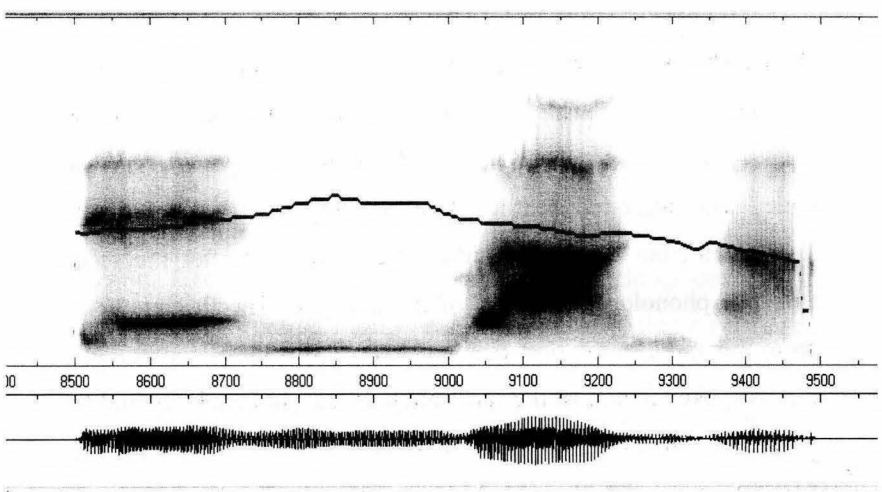
This represents the pitch pattern of four-mora words, as in 'aijoo' (love) and 'sansee' (agreement). The black dots and the white dot indicate high-pitched morae and a low-pitched mora, respectively. The arrow after the third black dot indicates the pitch of the following post-position. Judging from the angle of this arrow, we can tell that the pitch of the post-position is equal to that of the high-pitched three morae. When this arrow is described as descending, this means that the preceding word belongs to the end-high type, as in 'imooto' (younger sister) and 'otooto' (younger brother). In other words, it is this direction of the angle after the last mora that differentiates the flat type from the end-high type. Without a post-position, words belonging to both types are pronounced identically.

Let us examine the f0 contour of the flat type by using three examples with different numbers of morae. The examples cited here are 'kabe' (wall), 'denwa' (telephone) and 'shingoo' (traffic light). The number of morae is two, three and four, respectively. All of these words are followed by a post-position 'ga'. Contrary to the orthography, 'g' in 'ga' is pronounced as the velar nasal[3]. The analysis results are shown in Figure 2.

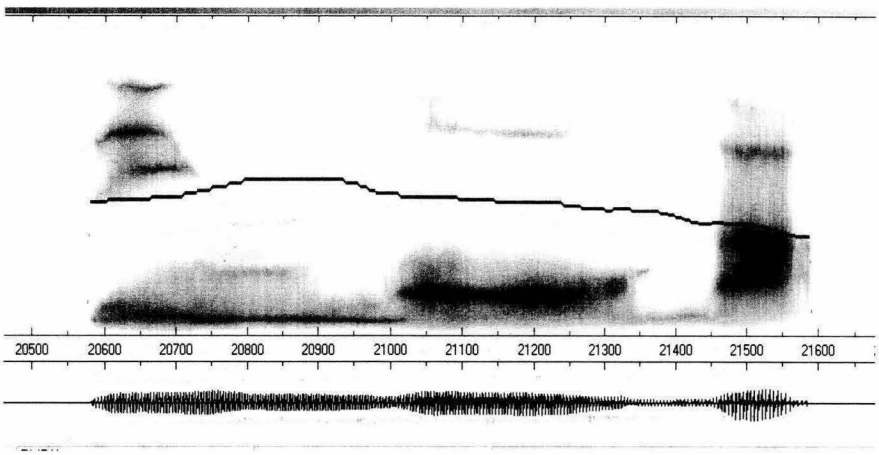
[3] Some Japanese people pronounce this 'g' as the voiced velar stop.


6000
16100
16200
16300
16400
16500
16600

(a)


)0
8500
8600
8700
8800
8900
9000
9100
9200
9300
9400
9500

(b)



(c)

Figure 2 Panels (a), (b) and (c) show the analysis results of 'kabe-ga', 'denwa-ga' and 'shingoo-ga', respectively. The speaker is a female.

This figure shows the three major pieces of information of an utterance: the f0 contour, the spectrogram and the waveform. The first two pieces of information are displayed above the ruler which shows the duration. The spectrogram uses the LPC analysis mode.[4] The figure tells that it is true that the pitch goes up from the first mora to the second mora[5], but the movement of the f0 contour after the second mora is different from the phonological display of the flat type. The f0 contour continuously descends from this mora to the post-position 'ga'. Since 'kabe' consists of two morae, the first mora is spoken in almost the same pitch as 'ga'. In Panels (b) and (c), however, 'ga' is pitched lower than the first mora. It is known that an utterance spoken not with the rise begins in a higher pitch and ends in a lower pitch. This phenomenon is related to human physiological traits and is called *declination*. It seems from Figure 2 that the declination is more influential than the pitch type indigenous to a particular word. This

[4] The reason why we use the LPC analysis mode, not the FFT analysis mode is that the former produces clearer analysis results than the latter. Since this paper is printed in black and white, not in colour, the original superimposed red-coloured f0 contours cannot be clearly distinguished from the spectrogram. The spectrogram in the LPC mode, however, shows the formants more clearly, which helps to read the superimposed f0 contours more easily than the FFT mode.

[5] When the first mora is devoiced and its consonant is voiceless, as in 'kuchi' (mouth) or hikari (light), the f0 contour for this mora is not displayed. But we can perceive the rise of the pitch from the first mora to the second mora. This may be related to our knowledge of the word or the general phonetic feature of Japanese. The first mora and the second mora are not spoken in the same pitch in Tokyo Japanese. When the first mora is low, the second mora should be high, and when the first mora is high, the second mora should be low.

is also clearly shown when a flat type + a post-position is followed by another flat type + a post-position, as in 'denwa-ya tegami-de kimochi-o[6] tsutaeru' ((We) convey our feelings by phone and mail)[7]. All the three nouns ('denwa', 'tegami' and 'kimono') and the verb ('tsutaeru') belong to the flat type. The analysis result of this utterance is displayed in Figure 3.

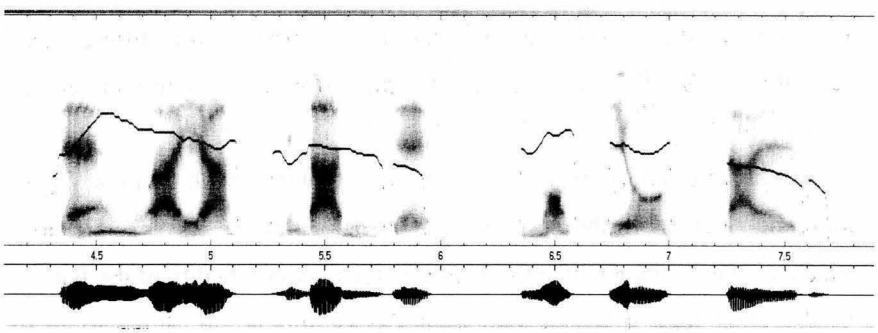


Figure 3 This shows the analysis result of 'denwa-ya tegami-de kimochi-o tsutaeru'. The speaker is a female.

The noticeable feature of the f0 contour is detected in the first part of the utterance 'denwa-ya tegami-de', where the overall declination is more influential than the accent type of each word. This is shown by the fact that the angle of descent between the second mora 'n' and the post-position 'ya' in 'denwa-ya' is identical to the angle of descent between the second mora 'ga' and the post-position 'de' in 'tegami-de', with a slight step-up in the second descent. This may be related to a close semantic link between 'denwa-ya' and 'tegami-de' in that both 'denwa' and 'tegami' are targeted to the post-position 'de' which shows a means of doing something. In contrast, the f0 contour for the first mora 'ki' in 'kimochi-o' is not lower than the previous mora 'de'. It is as high-pitched as the first mora 'de' in 'denwa-ya'. This is because of what we call the *f0 adjustment* to indicate a semantic boundary. This f0 adjustment is useful in conveying the information chunks to the listener and in increasing his/her understanding. It also enables the speaker to continue to speak by signalling the listener that he/she still

[6] The post-position 'wo' is an orthographic description, not a phonetic one. Phonetically, there is no approximant in this post-position. It is also a fact, however, some Japanese pronounce this approximant especially in singing.
[7] The subject is unknown from the utterance, but we put 'We' as a subject here.

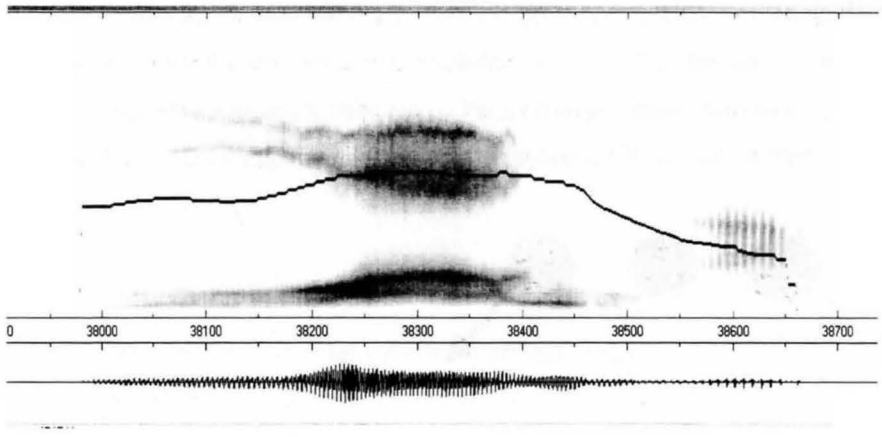
has more to say.[8] Without this adjustment, the f0 contour is likely to reach the speaker's bottom range soon after he/she starts speaking, unless he/she starts speaking at a very high pitch, which is practically impossible. It should also be noticed that there is no f0 contour for the vowel 'u' in the first mora of 'tsutaeru', because it is devoiced. This means that we cannot see how the f0 contour changes between the first mora and the second mora of this word. In this case, however, we can tell that this is spoken in the flat type for two reasons. First, in Tokyo Japanese, the pitch of first mora is different from that of the second mora. Second, the descent of the pitch from the second mora and the third mora is very mild.

In sum, the phonetic reality of the flat type is not identical with its phonological representation. The pitch after the second mora is not level. It descends toward a semantic boundary[9] due to the declination. In a way, the phonetic reality of the flat type is similar to the phonological representation of the end-high type. It is necessary, therefore, to examine the phonetic reality of the end-high type to find the difference between these two types.

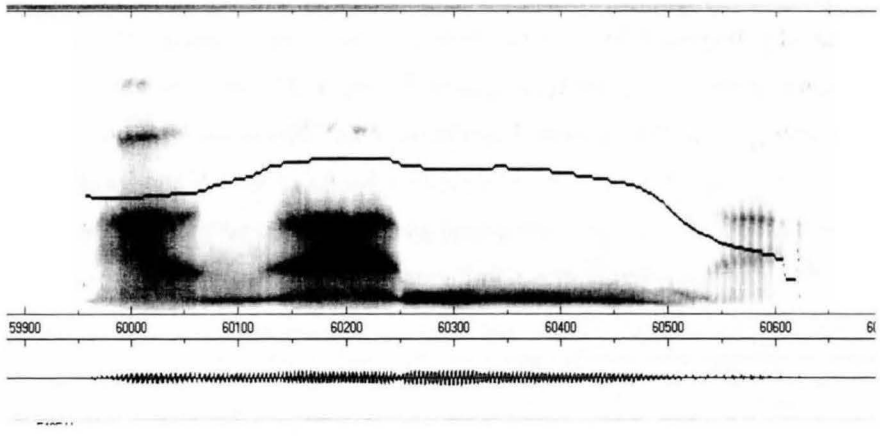
2. The End-high Type

The following three examples belonging to the end-high type are examined: 'iE' (house), kagaMI (mirror) and 'rekishiSHO' (history book).[10] As in the flat type, each of the three words is followed by the post-position 'ga'. The analysis results are shown in Figure 4.

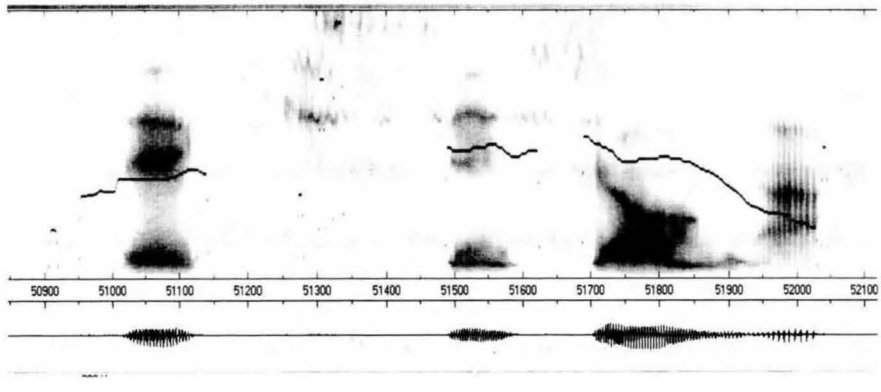
[8] In addition to this phonetic cue, syntactic and semantic cues are also helpful in successful exchanges in communication.
[9] It may be possible to equate this semantic boundary with the tone unit.
[10] The uppercase mora indicates the accented mora. This convention is created by the author.



(a)



(b)



(c)

Figure 4 Panel (a), (b) and (c) show the analysis results of 'iE-ga', 'kagaMI-ga' and 'rekishiSHO-ga', respectively. The speaker is a female.

This figure shows that, just like the flat type, the descent of the f0 contour from the second mora is detected.[11] The very steep descent is also noticeable between the last mora of each word and the post-position 'ga'. This appears to show the validity of the phonological representation of the end-high type, which is diagrammed as in Figure 5.

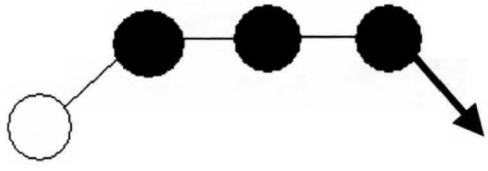


Figure 5 This shows a phonological description of the end-high type.

Let us see the f0 contour for the end-high type in a longer utterance and compare it with the f0 contour for shorter utterances, as in Figure 4. The example for the longer utterance is 'kaeRI-ni minNA-de hanaMI-o shiyoo' ('Let's go and see cherry blossoms all together on the way home'). All the three nouns ('kaeRI', 'minNA' and 'hanaMI'), as the uppercase morae show, belong to the end-high type, but the verb ('suru') belongs to the flat type. The analysis result is shown in Figure 6.

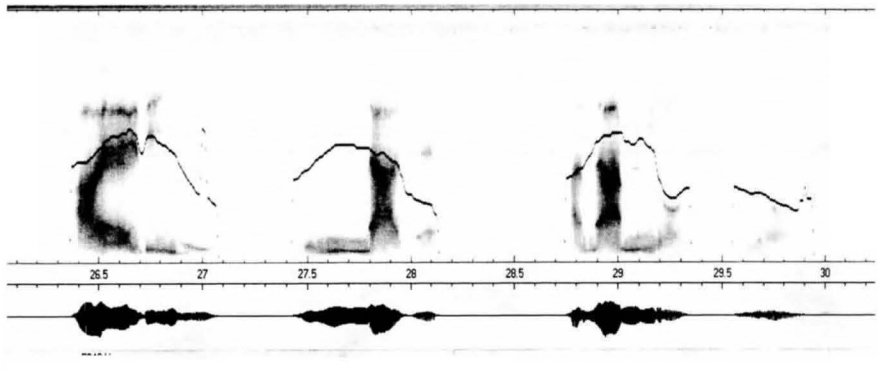


Figure 6 This shows the analysis result of 'kaeRI-ni minNA-de hanaMI-o shiyoo'. The speaker is a female.

As this figure shows, the end-high type is clearly shown in the three nouns, as embodied by the difference in pitch between the third mora of each noun and the

[11] In the case of 'rekishiSHO-ga', the f0 contour for the second mora is missing, due to the voiceless consonant and the devoiced vowel.

post-position. The influence of the declination is also detected in 'kaeRI-ni minNA-de', where the peak of the former is higher than that of the latter. The duration between 'minNA-de' and 'hanaMI-o' is relatively long due to a semantic boundary, and the f0 adjustment is also detected at the onset of 'hanaMI-o' as the peak of this noun is very similar to that of the first noun 'kaeRI'.

3. Comparison of the Flat Type with the End-high Type

We pointed out in Figure 4 that there is a steep descent between the last mora of each word and the post-position in the end-high type. It is worth examining here whether this steep descent is statistically significant or not, and the difference in the vertical value of the f0 contour is calculated between the last mora of the word and the post-position, both in the flat type and in the end-high type. It is necessary, in the beginning, to choose the measurement point, as the vertical value of the f0 contour changes in sonorant segments. Our idea is to measure the value in the middle of the vowel, because the vowel is the core of the mora as well as the syllable. Another important factor is that the calculation should not be performed in Hz, because what is important in this comparison is the relative difference that may be applicable to speakers in general. The semitone conversion[12] is introduced for this purpose. The result is shown in Table 1.

Type	Words	Difference
Flat Type	kabe-ga	3.0
	denwa-ga	3.3
	shingoo-ga	2.7
End-high Type	iE-ga	12.8
	kagaMI-ga	13.9
	rekishiSHO-ga	11.7

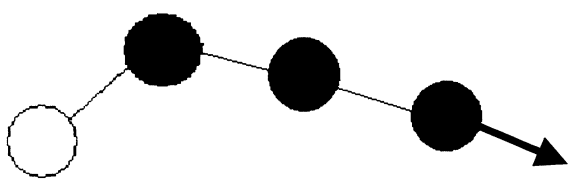
Table 1 This shows the difference in the vertical value of the f0 contour between the last mora of the word and the post-position. The unit of the numbers is ST.

As this table shows, there seems to be a clear difference between the flat type and

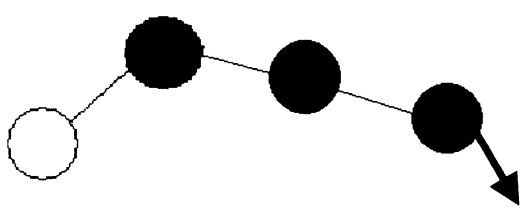
[12] The formula used in this conversion is $D = 12 \times \log_2 (f_1 / f_2)$, where D means a distance between two given points in the f0 contour. See 't Hart et al. (1990: 24) for more detail.

the end-high type, as the mean value is very different: 3.0 ST for the former and 12.8 ST for the latter. A t-test also supports this result ($t = -14.89$, $df = 4$, $p <= 0.001$ (two-tailed)). This result indicates that the flat type and the end-high type are categorically different. In other words, the steep descent found in the end-high type between the last mora of a noun and the post-position is important.

It should be remembered that one common feature found in the two types is the gradual descent of the f0 contour from the peak at the second mora. This is caused by the fundamental physiological trait and, for this reason, this phenomenon may be considered to be irrelevant to the phonological representation and conventionally Figures 1 and 5 are used for years. In learning the Japanese accents, however, we think that the learners should be informed of this gradual descent of the f0 contour, by telling them clearly that the pitch is not completely level as the representation of the accent types shows. This is because those who have a keen ear for changes in pitch are likely to be confused between the phonological representation and the phonetic reality. In this sense, Figures 1 and 5 should be also displayed as in Figure 7 (a) and (b), respectively, as supplementary aides, adding the information of declination.



(a)



(b)

Figure 7 This shows a narrow phonological description of the flat type, shown in (a), and the end-high type, shown in (b).

This revised description can be classed as a narrow phonological description. Some may think that the original description of the flat type and the end-high type is better and easier to understand. However, we think that for learners with a sensitive ear for changes in pitch, this new description will be more useful and true to their perception.

One interesting example, which is confusing to such learners, is found in the pronunciation of 'tookai', a place name in Japan. The speaker, who is a professional NHK announcer, pronounces this word both naturally and slowly. Interestingly or surprisingly, he does not use an identical pitch pattern in the two speaking styles, as Figure 8 shows. Note that in this slow speaking style, he pronounces each mora very carefully.

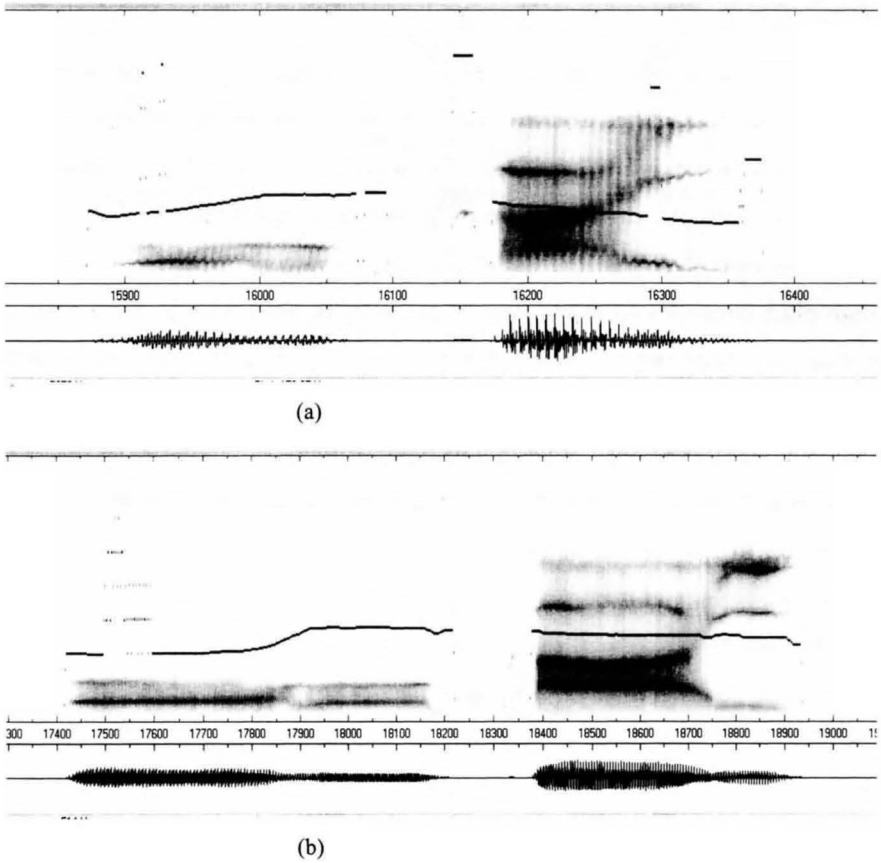
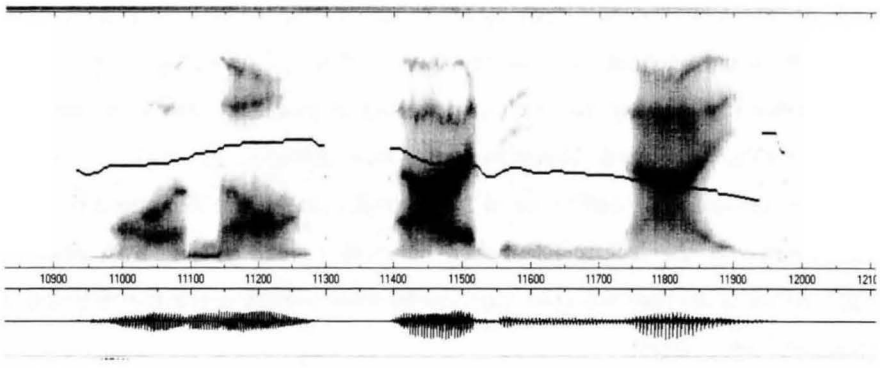


Figure 8 Panel (a) shows the analysis result of 'tookai' spoken naturally, and Panel (b) shows the analysis result of the same word spoken slowly. The word 'tookai' belongs to the flat type. The speaker is a male.

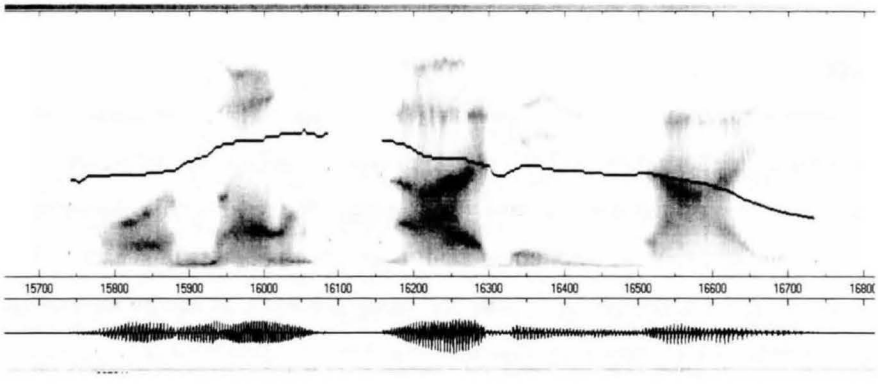
The word is spoken in 549 ms in Panel (a) and 1,612 ms in Panel (b). The former is spoken approximately 2.9 times faster than the latter. Panel (a) shows a descending f0 contour after the second mora 'o', and this corresponds to all the examples we have seen in the flat type, and hence to Figure 7 (a). In contrast, Panel (b) does not show any descending f0 contour after the second mora. The f0 contour shows the phonological pattern of the flat type, as shown in Figure 1. It seems that the phonological representation may be applicable phonetically if the speaker pronounces slowly and carefully. More importantly, it seems that at least for this speaker, the difference in the f0 contour found in the two speaking styles is irrelevant to his perception.[13] In other words, the declination may not be related to his perceptual ability of spoken Japanese. However, this observation poses a mystery because changes in pitch are important in standard Japanese.[14]

As we have seen, the flat type and the end-high type are clearly distinguished based upon the difference in the vertical value of the f0 contour for the post-position. A similar phonetic distinction is also found in some adjectives. To take 'monotarinai' (not satisfactory) as an example, this adjective is spoken both in the flat type and in the mid-high type 'monotariNAi'. In the mid-high type, the second mora is high-pitched and this high pitch continues up to the fifth or penultimate mora 'na', as the uppercase mora indicates. The final mora 'i' is low-pitched. This adjective of the mid-high type corresponds to nouns of the end-high type followed by a post-position, as in 'iE-ga', 'kagaMI-ga' and 'rekishiSHO-ga'. It is interesting to examine how the final mora 'i' is spoken differently. The analysis results of the two types of this adjective are shown in Figure 9.

[13] We think, however, that this presentation of an example in the two speaking styles is misleading to the learner, because the two types of the f0 contour are far from identical.
[14] This may raise another important research question, but we will not pursue this in this paper.



(a)



(b)

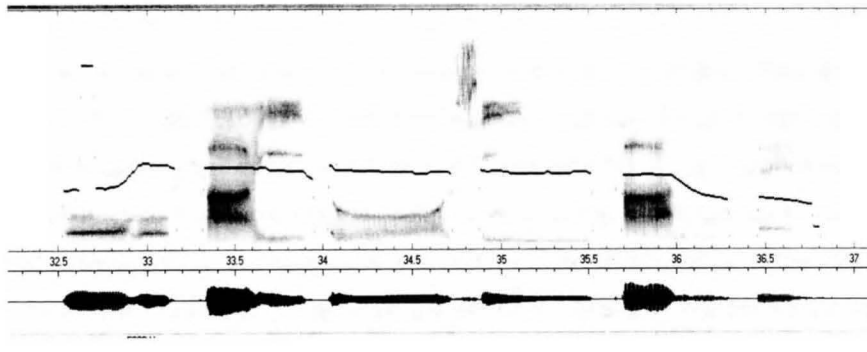
Figure 9 Panel (a) and (b) show the analysis results of 'monotarinai' spoken in the flat type and in the mid-high type, respectively. The speaker is a female.

One noticeable feature found in this figure is that the vertical value of the final mora 'i' is not very different between Panel (a) and Panel (b): 155.7 Hz for the former and 140.5 Hz for the latter. This difference between the two types is not as great as we expect. As far as the difference in the vertical value between the penultimate mora 'na' and the final mora 'i' is concerend, however, the two types are differentiated more clearly: 2.1 ST for the flat type and 5.9 ST for the mid-high type, and this difference is manifested by the difference in the slope of the f0 contour from the penultimate mora to the last mora. The f0 contour for the flat type is much less steep than that for the mid-high type. It should be noticed that, in spite of the difference, the value for this mid-high type is much smaller than the values calculated for the end-high type earlier in this section. This difference may be related to the role of the accented mora. The accented mora is the last mora of a noun in the end-high type, while it is the penultimate

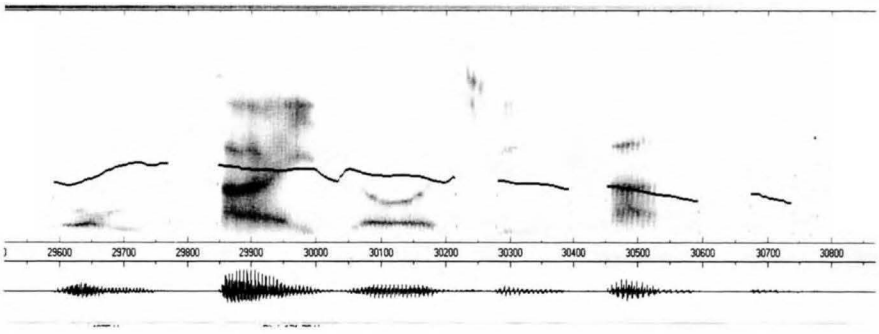
mora of an adjective in the mid-high type. In the former case, there is a clear morphological boundary between the accented mora and the following mora, which is a post-position, while in the latter, no such morphological boundary exists between the accented mora and the final mora. This presence or lack of the morphological boundary may be the main reason for the difference in the vertical value of the f0 contour between the accented mora and the following mora. The same tendency is also detected in 'muzukashii', which is another adjective that can be spoken both in the flat type and in the mid-high type 'muzukaSHIi'.

4. Flattening

The examples cited so far clearly show that the flat type is carefully distinguished from the end-high type for nouns and the mid-high type for adjectives. However, there are cases in which the non-flat type becomes the flat type. We refer to this phenomenon as *flattening*. One example is the compound noun 'tookaidoo shinkansen'. This utterance belongs to the mid-high type, with 'ka' being accented, as can be represented as 'tookaidoo shinKAnsen'. However, this accented mora is noticeable only when it is spoken slowly, as shown in Figure 10.



(a)

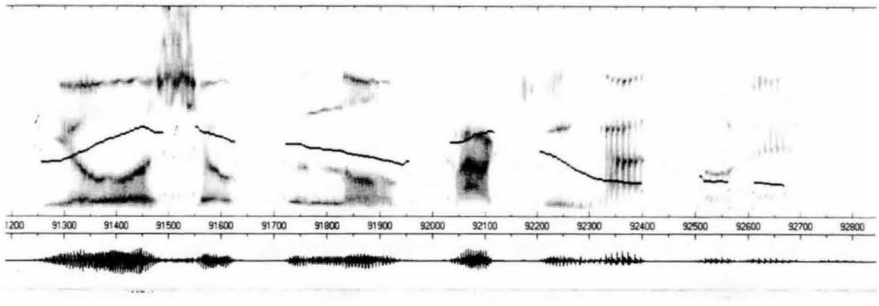


(b)

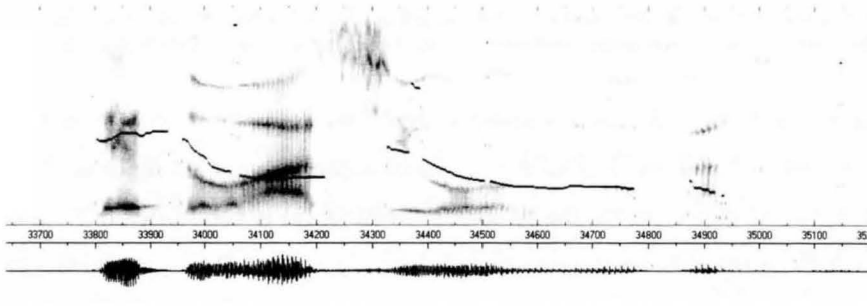
Figure 10 Panel (a) shows the analysis result of 'tookaidoo shinKAnsen' spoken in slow speed, and Panel (b) shows the analysis result of the same compound noun spoken in natural speed. The speaker is a male.

There is a difference in duration between these two examples: 4,489 ms in Panel (a) and 1,266 ms in Panel (b). The former is spoken approximately 3.5 times as slowly as the latter. As this figure shows, the f0 contour in Panel (a) is very different from the f0 contour in Panel (b). Besides the presence of the declination in Panel (b) and the lack of it in Panel (a), the mid-high type is clearly shown in Panel (a), probably due to careful pronunciation. The f0 contour ascends at the second mora 'o' and descends at the antepenultimate mora 'n', which is true to the phonological representation, as indicated by 'tookaidoo shinKAnsen'. In contrast, in Panel (b), a step-up in pitch is detected at the second mora, but no noticeable descent can be seen at the antepenultimate mora 'n' as a proof of the mid-high type. The whole f0 contour is identical with the f0 contour for the flat type. 'KA' and the following mora 'n' simply show a mild decline in the f0 contour due to the declination. Judging from the f0 contour, this mid-high type tends to behave like the flat type in natural speed, in spite of this speaker's knowledge that this compound noun belongs to the mid-high type. It seems that the phonological difference in accent types cannot always be observed in phonetic reality.

Another case of flattening can be detected at the end of utterances. Let us examine two examples: 'yooshoku-no KAki-ga toREru' (We can get farmed oysters) and 'TE-to aSHI-wo ugoKAsu' (I move my arms and legs). We focus on the last verbs: 'toREru' and 'ugoKAsu'. As these two words show, they belong to the mid-high type. The analysis result of these two utterances is shown in Figure 10.



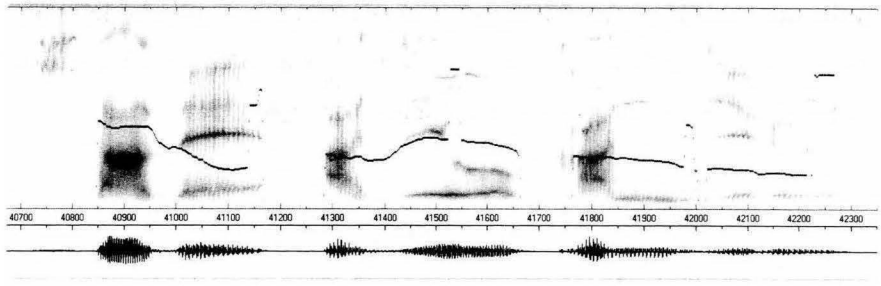
(a)



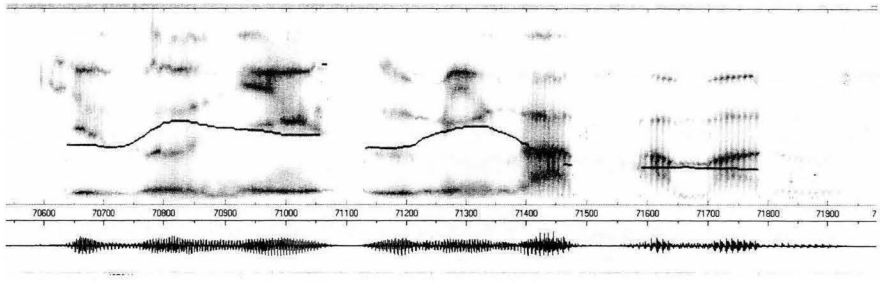
(b)

Figure 11 Panels (a) and (b) show the analysis results of 'yooshoku-no KAki-ga toREru' and 'TE-to aSHI-wo ugoKAsu', respectively. The speaker is a male.

As this figure shows, there is no sign of stepping up in pitch in 'RE' in 'toREru' and 'KA' in 'ugoKAsu'. They are spoken in the flat type, not in the mid-high type. To be more exact, they may not be spoken in the flat type, and this is detected in 'ugoKAsu', where the second mora 'go' is not stepped up. This verb is simply spoken in a mildly descending f0 contour. As can be guessed from this example, it is not only the mid-high type that is influenced by the flattening, but also the flat type itself is. The stepping up in the second mora is also missing in such utterances as 'shiTA-de aji-o kanjiru' (I taste with my tongue) and 'kizu-ni uMI-ga tamaru' (Pus gathers in the wound). Their analysis results are shown in Figure 12.



(a)



(b)

Figure 12 Panels (a)[15] and (b) show the analysis results of 'shiTA-de aji-o kanjiru' and 'kizu-ni uMI-ga tamaru', respectively. The speaker is a male.

As this figure shows, there is no sign of stepping up in pitch in the second mora 'n' in 'kanjiru' and 'ma' in 'tamaru', just as in the case of 'ugoKAsu'. It is known that due to the declination, the voice range becomes narrower as an utterance goes on, and it becomes very narrow at an utterance-end or a tone-unit boundary. For this reason, it may be easily inferred that changes in pitch becomes less noticeable in such places than at the beginning of an utterance or a tone-unit. However, such complete loss of accentedness, which may be referred to as *de-accentedness*, may weaken the importance of accents in Japanese at utterance-ends or tone-unit boundaries. One explanation about this phenomenon is found in NHK Nihongo Centre (1994: 50-51). It says that some words in an utterance are combined into a semantic chunk, and that each chunk carries the accent pattern suitable to it. This phenomenon is known as the *stability of accent*, and is realized unconsciously on the part of the speaker. One example that it picks up for the explanation is 'kinoo Yonda zasshi-ni oishii oSUshi-no tsukuriKAta-ga notteite totemo sankoo-ni NAtta' (The magazine that I read yesterday explains how to make

[15] The pitch jump at the utterance-end is believed to be a program error. Sometimes the measurement of the f0 contour at the utterance-end is not successful, mainly because of its low signal levels.

delicious sushi, and it was very useful to me), and this utterance is divided into three chunks: 'kinoo Yonda zasshi-ni', 'oishii oSUshi-no tsukuriKAta-ga notteite' and 'totemo sankoo-ni NAtta'. Its analysis result is displayed in Figure 13.

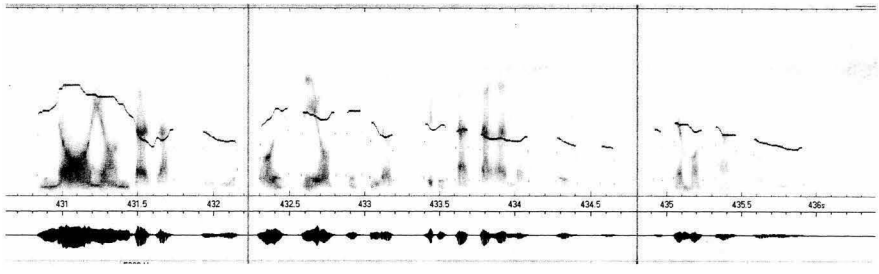
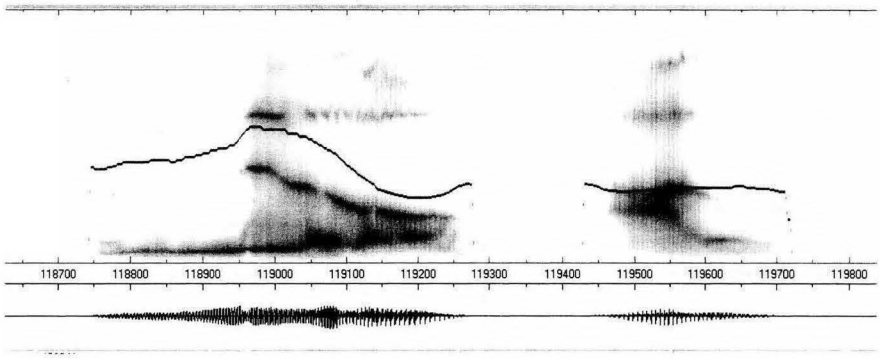


Figure 13 This shows the analysis results of 'kinoo YOnda zasshi-ni oishii oSUshi-no tsukurikata-ga notteite totemo sankoo-ni NAtta'. The speaker is a female.

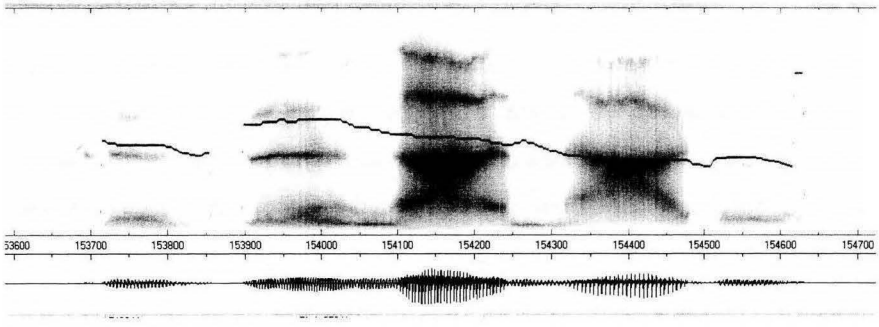
The two vertical lines indicate the boundary of the three chunks. As this figure shows, the f0 adjustment is detected at the beginning of the second and the third chunks, and the f0 contour in each chunk tends to descend under the influence of the declination. Moreover, the initial peak in each chunk becomes lower as the utterance proceeds, and this again is influenced by the declination. Coupled with the explanation by NHK Nihongo Centre (1994: 51), we can tell from this figure that the accent pattern of these three chunks are 'kinoo YOnda zasshi-ni', 'oishii oSUshi-no tsukurikata-ga notteite' and 'totemo sankoo-ni NAtta'. These chunks are informational units and may be considered as tone-units, and the accented three morae ('YO', 'SU' and 'NA') may be referred to as *tonic morae*.[16] In Japanese, it seems that what is important is that the division of an utterance into informational or semantic units (i.e. tonality) and the location of the tonic mora in each tone unit (i.e. tonicity).

The stability of accent is also noticed in short utterances such as 'iNU-o kau' (I buy a dog) and 'suzu-ga naru' (A bell rings), where the two verbs 'kau' and 'naru' are spoken in a completely flat, as shown in Figure 14.

[16] This is the author's coined term in reference to the tonic syllable used in the stress-timed languages such as English.



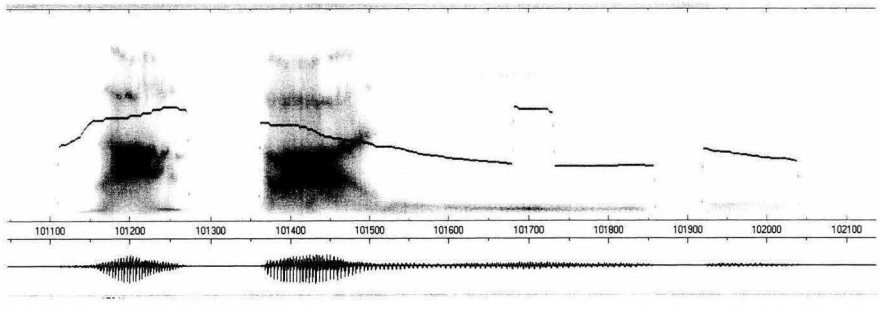
(a)



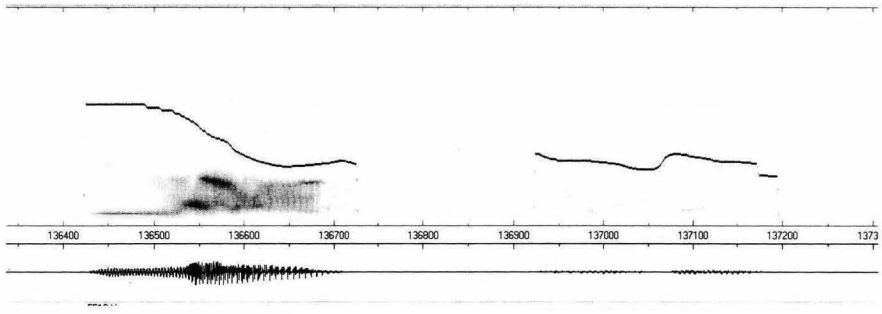
(b)

Figure 14 Panels (a) and (b) show the analysis results of 'iNU-o kau' and 'suzu-ga naru', respectively. The speaker is a female.

However, in short utterances with the same syntactic structure such as 'NAka-ni iru' (I am inside) and 'fuKU-o kiru' (I put on clothes), the stability of accent is not noticed in the two verbs 'iru' and 'kiru', as shown in Figure 15.



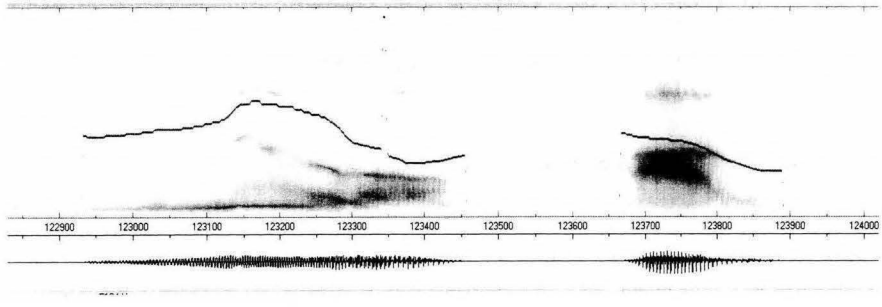
(a)



(b)

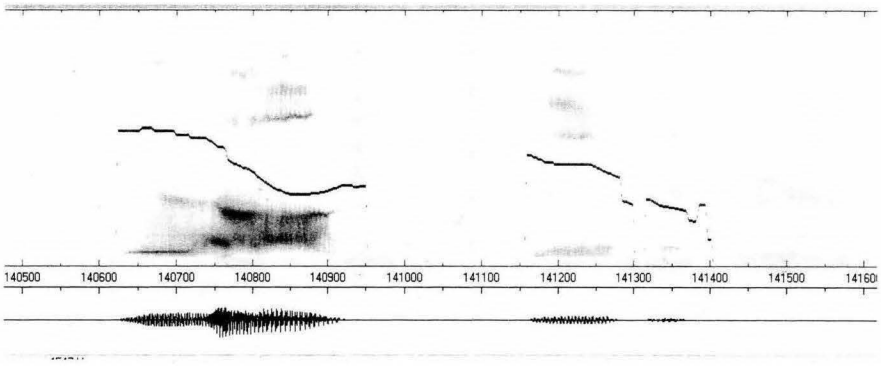
Figure 15 Panels (a)[17] and (b) show the analysis results of 'NAka-ni iru' and 'fuKU-o kiru', respectively. The speaker is a female.

The step-up in pitch is relatively noticeable in these two verbs, indicating the flat type more faithfully. However, when such two-mora verbs belong to the head-high type, this initial height appears to be retained clearly. We take two examples 'iNU-o KAu' (I keep a dog) and 'fuKU-o KIru' (I cut clothes), and the analysis results are shown in Figure 16.



(a)

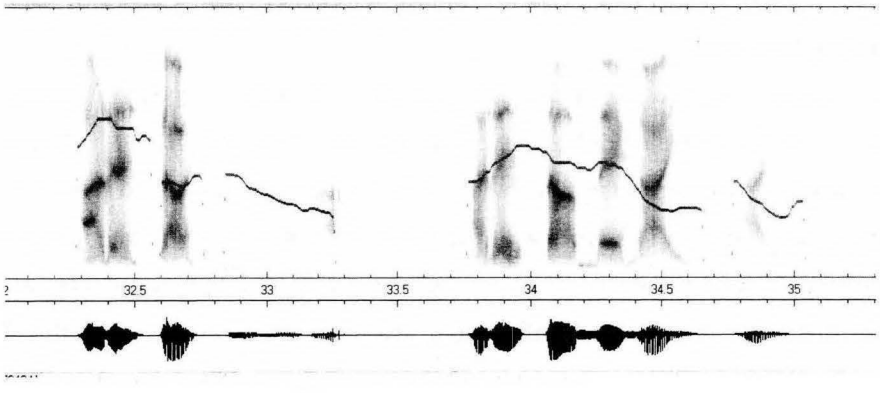
[17] The sudden step-up in pitch in the middle of the utterance is believed to be a program error.



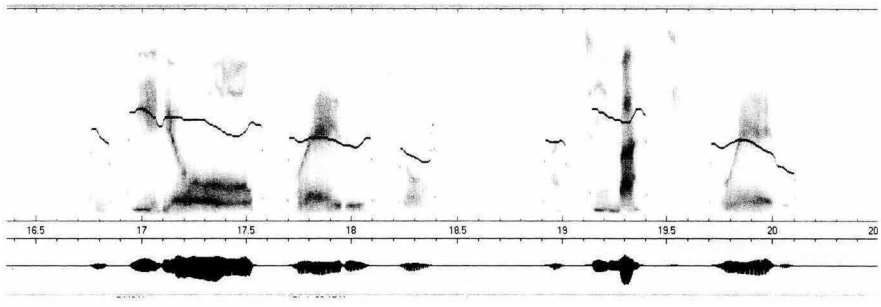
(b)

Figure 16 Panels (a) and (b) show the analysis results of 'iNU-o KAu' and 'fuKU-o KIru', respectively. The speaker is a female.

It seems that the head-high type is less influenced by the stability of accent rather than the flat type as well as the mid-high type. It should be noted here that the two utterances in Figures 14-16 are spoken by the same speaker in the same list of such utterances, which means that the result is irrelevant to influences by different speakers. Moreover, as examples belonging to the flat type show, the stability of accent may be fluctuated, and this fluctuation is also noticeable in the mid-high type, such as in 'HAreta HI-niwa hareta ME-ga iTAi' (On sunny days I have a pain in my eyes) and 'tekiryoo-o koeru-to tachimachi koEru' (I become fat when I eat too much). Their analysis results are shown in Figure 17.



(a)



(b)

Figure 17 Panels (a) and (b) show the analysis results of 'HAreta HI-niha hareta ME-ga iTAi' and tekiryoo-o koeru-to tachimachi koEru', respectively. The speaker is a female.

As this figure shows, the stability of accent is not influential in the two verbs 'iTAi' and 'koEru', where the second mora becomes higher than the first mora, and the accent realization is more faithful to the inherent phonological pattern. As mentioned above, it seems that this stability of accent is not always observed. NHK Nihongo Centre (1994: 18) states that the utterance 'HA-ga iTAi' (I have a toothache) should be pronounced as 'Ha-ga itai' by de-accenting 'TA', but as we saw in Figure 17, it seems that this violation is made even by professional announcers. One of the great difficulties in spoken Japanese is that how much we should be observant of the phonological pattern at the end of an utterance or a tone-unit. The important thing is that the declination is more influential than accent inherent in words in such places.

5. Concluding Remarks

In this paper, we focused on the flat type and flattening in spoken Japanese, including comparison with the head-high type, the mid-high type and the end-high type. In the beginning, we examined the f0 contour of the flat type and compared it with the phonological description of this type, with the result that contrary to the phonological description, the f0 contour does not show a straight level line from the second mora and it descends due to the physiological influence called the declination. In a way, this f0 contour looks like that of the end-high type. However, we learned by comparing the f0 contour of the flat type with that of the end-high type that the angle of the latter type is very steep at the last mora, which corresponds to the post-position. This feature was

supported in that the difference in pitch between the penultimate mora and the final mora, which corresponds to the post-position, is statistically significant by using the semitone conversion. In this sense, the phonological description of the flat type and of the end-high type displays the core feature of these two types correctly. We think, however, that especially for those who have an ear sensitive to differences in pitch, it would be better to describe those two types by using a slight descending line from the second mora, as shown in Figure 7. This could be referred to as the narrow phonological description.

A very noticeable feature in Japanese accent is that the accents near the end of an utterance or a tone-unit tend to be flattened. The non-flat types tend to be the flat type, with no step-up in the second mora, except for the head-high type. This includes the flat type itself, where the second mora is not spoken in high pitch. This phenomenon can be understandable in the sense that the voice range becomes very narrow near the end of an utterance or a tone-unit due to the declination, and for this reason, a step-up in pitch cannot be realized noticeably as in the utterance-onset. Moreover, each word is likely to lose its original pitch type as an utterance becomes longer. A good example is 'kinoo YOnda zasshi-ni oishii oSUshi-no tsukurikata-ga notteite totemo sankoo-ni NAtta' that we saw in Figure 13, and this utterance is likely to be divided into three parts when spoken: 'kinoo YOnda zasshi-ni', 'oishii oSUshi-no tsukurikata-ga notteite' and 'totemo sankoo-ni NAtta'. There is only one accented core in each of the three parts or three tone units: 'YO' in the first tone-unit, 'SU' in the second and 'NA' in the third. It should be remembered that these accented cores are referred to as tonic morae in this paper, in reference to the tonic syllable in a stress-timed language such as English. The pitch after these tonic morae becomes low and level phonologically. Phonetically, it becomes low and continues to descend mildly until it reaches the end of a tone-unit. This low part is relatively long. For instance, in the immediate above example, the number of morae is six in the first tone-unit, thirteen in the second, and two in the third. The reason why it is small in the third tone-unit may be that the final word 'natta' belongs to the head-high type, which may work as a good trigger to keep the pitch of the preceding morae high. The remaining morae after the tonic mora within the tone-unit become de-accented. In short, the narrower voice range and the de-accentedness make the pitch at or near the end of an utterance or a tone-unit less noticeable, or usually flattened. This may be one

of the reasons why it becomes more difficult to perceive the tonic mora correctly in longer utterances, especially at the final tone-unit. However, the location of the tonicity in each tone-unit may be far from unpredictable once the tonality is correctly decided as the speaker intends, because the candidates are only the tonic morae in each word in the given tone-unit. In this sense, it may be easier to locate the tonicity in spoken Japanese than in spoken English. It is an inevitable fact, however, that utterance accents are much more difficult to study than word accents and phrase accents, because, as can be inferred from the above discussion, utterance accents cannot be explained by the accumulation of studies of word accents and phrase accents. Moreover, we noticed that the de-accentedness after the tonic mora is not always realized, as we saw in Figure 17, which means that such de-accentedness may have variations, ranging from non-existence to clear presence.

Accentedness can be perceived by the change in pitch. Except for the head-high type, the second mora is always higher than the first mora, unless it is influenced by the stability of accent near or at the end of an utterance or a tone-unit. A question may arise as to the reason why this step-up in pitch cannot be perceived as accented as the step-down from the tonic mora. This may be explainable from what Tench (1996: 77) calls *inevitable jump*. He states that in English a novice tends to consider a rise before a falling tone to be a rising tone. The pitch should jump up to a higher level in order to realize a falling tone, which means that this initial rise before the falling tone is an inevitable jump. We think that this idea of inevitable jump can be successfully used to explain the reason why the step-up in pitch at the second mora cannot be perceived as accented in Japanese. In this sense, the perception of accentedness may entail more of subjective quality.

References:

Amanuma, Y., Ootsubo, K., and Mizutani, O. (1978). *Nihongo Onseigaku (Japanese Phonetics)*. Tokyo: Kuroshio-shuppan.

Kawawa, T. (1988). *Nihongo-no Hassei Ressun (Pronunciation Lessons in Japanese)*. Tokyo: Shinsui-sha.

Matsuzaki, H. and Kawano, T. (1998). *Yokuwakaru Onsei (Easy-to-understand Sounds)*.

Tokyo: ALC.

NHK Hosobunka Kenkyujo. (1998). *Nihongo Hatsuon Akusento Jiten (Dictionary of Japanese Pronunciation and Accent)*. New Edition. Tokyo: NHK.

NHK Hosobunka Kenkyujo. (2002). *Nihongo Hatsuon Akusento Jiten (Dictionary of Japanese Pronunciation and Accent)*. CD-ROM. Tokyo: NHK.

NHK Nihongo Centre. (1994). *Akusento Hatsuon (Accent and Pronunciation)* with three cassette tapes. Tokyo: NHK.

Tench, P. (1996). *The Intonation System of English*. London: Cassell.

't Hart, J., Collier, R. and Cohen, A. (1990). *A Perceptual Study of Intonation: An Experimental-Phonetic Approach to Speech Melody*. Cambridge: Cambridge University Press.

日本語における平板型と平板化

湯澤　伸夫

本論文では，日本語のアクセントに関して論じた。まず，平板型の音声特徴を尾高型と比較し，学習上の観点から自然下降現象も表示できる精密音韻表記を提案した。日本語ではアクセントは重要な要素だが，自然下降現象と発話速度との微妙な関係から，表記と音声実体には見過ごすことのできない相違があることを例示した。また，発話速度が増すとアクセント核が曖昧になったりすることや，音調単位の後方では，アクセント型の違いによる影響はあるものの，アクセントの弁別能力が曖昧になる傾向があることも例示した。しかし，ここで注目すべきは，こうした物理的変化は発話者の意識にはなく，この点から日本語のアクセントには認知的要素がかなり影響していると考えられる。

（高崎経済大学経済学部 助教授）