Journal of Phonetics (1997) **25,** 43–60

Japanese final-accented and unaccented phrases

Natasha Warner
Department of Linguistics, University of California, Berkeley, 2337 Dwinelle Hall, Berkeley, CA 94720, U.S.A.

Received 29th January 1996, and in revised form 17th September 1996.

The most common theory of Japanese pitch accent predicts that words with final-accent and those with no accent have the same f_0 contour within the word, and diverge only in the mora following the word. Pierrehumbert & Beckman's (1988) theory, however, predicts a difference even within the word for the two types. Previous instrumental studies of final-accented and unaccented words, the results of which have not been conclusive, used only one- or two-mora words, which introduce a confound for Pierrehumbert & Beckman's theory. In this experiment, longer words are used, and reiterant speech is used to lower variance. All four speakers produced the words with a difference between the two accent types in the direction predicted by Pierrehumbert & Beckman's theory, and this difference is statistically highly significant. However, the speakers produced an additional difference between the two types which is not predicted by either theory. The results have implications for other theories of Japanese pitch accent. The experiment also tests, and confirms, the validity of reiterant speech as a method.

1. Introduction

1.1. *Comparison of the competing theories*

Within the considerable literature on Japanese pitch accent, one often encounters the opinion that words with the accent on their final mora (final-accented) and words with no accent at all (unaccented) have the same f_0 pattern within the word. The difference between the two types is said to become apparent only on the mora following the word. However, the theory of Japanese pitch accent presented by Pierrehumbert & Beckman (1988) makes a different prediction about these types of words. The present experiment investigates this difference in predictions.

What I will call the ''traditional'' theory includes a great deal of work by both Japanese and Western scholars in a variety of frameworks. However, all of the treatments included in this group produce basically the same string of surface tones. They agree that the first mora of a phrase has a low tone (unless it is the accented mora or the first syllable is long), every mora from the second to the accented mora has a high tone, and every mora after the accented one has a low tone. The effects

0095-4470/97/010043 + 18 $25.00/0/jp960033

of long syllables and initial accent will not be relevant for this paper. McCawley (1968) derived this output within an early generative phonology framework; Haraguchi (1977) offered an autosegmental analysis with the same output, except for a small difference in non-initial phrases. The traditional theory assigns the following tones for the commonly cited nouns below (McCawley, 1968: 132, 191; Vance, 1996, personal communication), where an apostrophe follows the accented mora:

(1)

H L L		L H L	
| | |		| | |	
i'noti	''life''	ito'ko	''cousin''
L H H		L H H	
| | |		| | |	
atama'	''head''	pmijako	''capital''

As demonstrated by /atama'/ and /mijako/, the traditional theory claims that, within the word, a final-accented word has the same tones as an unaccented one, and that the difference appears only when something follows the word. McCawley made this explicit when he stated that ''a final-accented phrase . . . is indistinguishable from an accented phrase: each is pronounced entirely on a high pitch, except for the first mora, which is low-pitched'' (1968:139). Statements to the effect that the difference between final-accented and unaccented words is neutralized unless something follows the word occur frequently in the literature (Sugito, 1968:1; Weitzman, 1970:18; McCawley, 1977:262; Higurashi, 1983:12; Vance, 1987:81). Any theory which assigns these same tones, without discussing details of the phonetic f_0 contour, can only predict that final-accented and unaccented words will be the same: there is simply no mechanism in such theories for giving the same LHH string different f_0 contours.

If another syllable, such as a particle or the copula, follows the final-accented or unaccented word within the same prosodic phrase, then the difference between the two types becomes evident, but only in the mora following the word itself:

(2)

L H H **L**		L H H **H**	
| | | |		| | | |	
atama'-wa	''head-topic''	mijako-wa	''capital-topic''

That is, a final-accented word will be followed by a low tone mora, since any mora after the accent is low, while an unaccented word will be followed by a high tone mora, since there is no accent to trigger the fall to low pitch. The tones assigned to the two types are still the same for the words themselves, and diverge only on the following mora.

The neutralization of final-accented and unaccented words has always been somewhat controversial. Sakuma (1929:359–76) presented early experimental data to show that the two types are different, in that final-accented words have a higher f_0 peak that unaccented words. He used a three tone system to distinguish them, where /atama'/ has the tones LHH, but /mijako/ has LMM. This is the basis of a long debate among Japanese researchers as to whether Japanese requires three tones in order to account for a difference between final-accented and unaccented types, or just two tones, for a more abstract system. The arguments for and against the two-

and three-level theories are reviewed in Kawakami (1995:143–148). Uwano (1977:289) said there is only a tendency for the final-accented and unaccented types to become identical in isolation, and that they are different for some speakers, some of the time, but did not explain this in terms of the three-level theory.

The traditional theory makes no explicit predictions about what fundamental frequency will do during any of the tones assigned by the theory. Of course no one would claim that f_0 progresses in sudden jumps between low and high, stair-step fashion. However, without instrumental methods, linguists working in the traditional theory could do little toward a more explicit description. Although one could ascribe everything involving f_0 to phonetic realization rules, Poser (1984) as well as Pierrehumbert & Beckman have shown that any mapping from high and low tones to f_0 is not straightforward.

Several more recent approaches to Japanese pitch accent use instrumental methods and deal with the f_0 contour itself. Higurashi (1983) and Poser (1984) investigated f_0 extensively, but still worked within some form of the traditional theory. Pierrehumbert & Beckman (1988) introduced an entirely new method of describing Japanese pitch accent by using only a few tones per phrase, with interpolation between them. The tones they assign a phrase are limited to a boundary low tone (L%) at the beginning of an utterance, a ''phrasal peak'' high tone (H) which is normally attached to the second mora, an ''accent peak'' high–low tone (HL) on the accented mora, and a boundary low tone (L%) at the end of each phrase. An accent peak (at the HL tone) is slightly higher than a phrasal peak (H tone only). The HL accent tone does not occur in an unaccented phrase. Their theory's basic predictions for /atama'-wa/ and /mijako-wa/, based on phonological effects on f_0 only, are given in Fig. 1.

As shown in Fig. 1, Pierrehumbert and Beckman's theory predicts that final-accented and unaccented words will be different even within the word. In a sequence of final-accented word plus particle, such as /atama'-wa/, f_0 rises from the initial boundary low tone to the phrasal high tone, then continues to rise to the slightly higher accent HL tone, and finally falls sharply in the particle after the word to the phrase-final low tone (see Pierrehumbert & Beckman, 1988:124 on the timing of the fall). In the unaccented /mijako-wa/, f_0 rises initially to the phrasal peak in the same way as the final-accented word, but then falls steadily to the boundary low tone at the end of the phrase. The particle after the word is also included in that fall. Even without reference to the final particle, the theory predicts a difference:

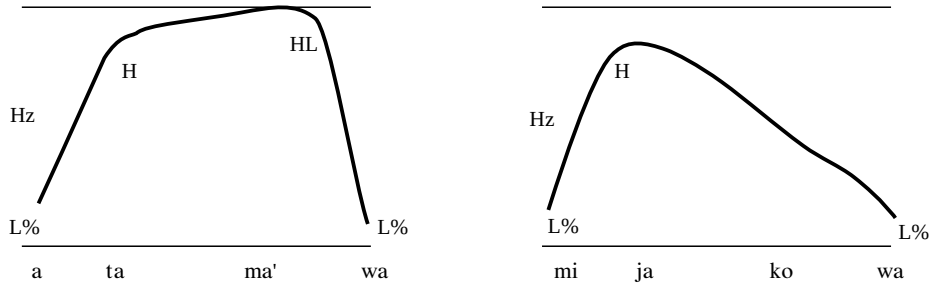

Figure 1. f_0 curves predicted by Pierrehumbert and Beckman's theory for final-accented /atama'/ and unaccented /mijako/, each followed by the topic particle /wa/. Only phonological effects on pitch are included.

between the second mora, at which the phrasal peak occurs, and the last mora of the word itself (the mora before /wa/), f_0 should rise slightly for final-accented words, but fall for unaccented words.

However, Pierrehumbert and Beckman's theory also includes several purely phonetic factors which affect f_0 and which have not been included in the predictions in Fig. 1. These factors influence both final-accented and unaccented words equally, regardless of whether they are followed by a particle or not. One such factor is final lowering; Pierrehumbert and Beckman (1988:74) showed that the last several moras of a declarative utterance have lower f_0 than in a corresponding question, and therefore posit a final lowering effect for statements. This should affect the end of both final-accented and unaccented statement utterances equally, but it could be strong enough to make the high stretch in the final-accented utterance flat or slightly falling instead of slightly rising. Still, the two types should be different, since unaccented utterances are also subject to this effect.

A second phonetic factor which affects f_0 is declination. Declination, in Pierrehumbert and Beckman's terminology, is a completely phonetic unconditioned lowering by which f_0 falls by some small number of Hertz per second in all utterances [although this fall need not take place in a linear fashion (1988:70)]. This should also affect both final-accented and unaccented words equally, but will also contribute to making the high stretch in a final-accented word slightly falling instead of slightly rising.

One additional factor not included in Fig. 1 is phonologically conditioned. Pierrehumbert and Beckman's data showed that the value of a final boundary low tone depends on whether there is an accent in the prosodic phrase it ends or not: a final L% is at considerably higher f_0 at the end of an unaccented phrase than at the end of a phrase containing an accent (Pierrehumbert & Beckman, 1988:83–84). They attributed this to the effect of catathesis in the accented utterance. Accent HL tones trigger catathesis, which decreases the pitch range after the HL tone by lowering the high line. This has the effect of making everything after an accent lower than otherwise expected until pitch range is reset at the next intermediate phrase boundary (similar to a major phrase boundary). This difference in final L% values is also apparent in my data, but the present experiment does not test whether catathesis is the cause of this difference.

The higher value for L% at the end of an unaccented phrase means that unaccented utterances will not be expected to fall as far as accented ones. This further reduces the difference between the two types. However, Pierrehumbert and Beckman stated that within an accentual phrase, no L tone can be higher than a H tone (1988:189). Therefore, their theory still predicts a difference between final-accented and unaccented words within the word, but it will not be as obvious as in the basic prediction shown. Both types are expected to fall after the phrasal peak, but the unaccented utterance should fall slightly more than the accented one.

Pierrehumbert and Beckman's theory makes the explicit prediction shown in Fig. 2, including phonetic effects. In Fig. 2, the top and bottom lines, which represent the limits of the pitch range, are slanted to account for the overall lowering effect of declination. The tones are the same as in Fig. 1, but the f_0 curve has been lowered during the last three moras of each phrase to account for final lowering. The f_0 at the final boundary low tone (the end of the particle) has been adjusted so that the final L% is higher at the end of the unaccented phrase than at the end of the phrase

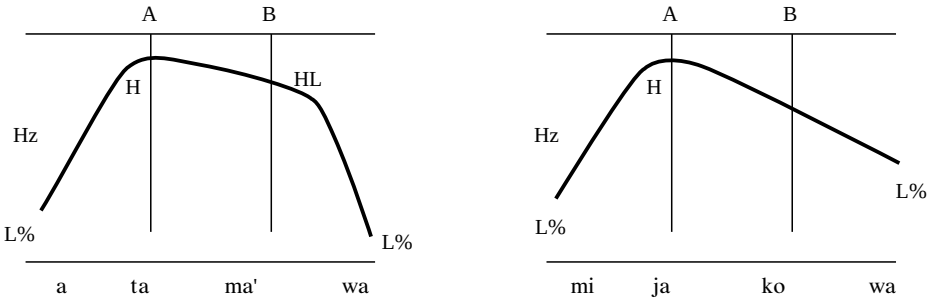

Figure 2. f_0 curves predicted by Pierrehumbert and Beckman's theory for /atama'-wa/ and /mijako-wa/, including phonetic effects on pitch.

containing an accent. The vertical lines labeled A indicate the point of the phrasal peak, and the B lines show a point early in the final mora of the word itself. Between these two points, f_0 falls in both words, but falls somewhat more in the unaccented word, since its final L% tone is lower than the HL tone of the final-accented word.

The traditional theory never reaches the true phonetic level, and it makes no explicit prediction. It predicts only that whatever the f_0 contour for final-accented and unaccented words is, it should be the same for both types within the word and then diverge only in the mora following the word.

The question of whether final-accented and unaccented words are the same or not has been experimentally investigated before (Sakuma, 1929; Sugito, 1968, 1982; Poser, 1984; Vance, 1995). Sakuma found higher peak f_0 for accented one-mora words than for unaccented when the words were read with a following particle, but data for the same words in isolation was inconsistent (1929:361–68). Measuring the f_0 peak on the second mora of two-mora words, Poser found that final-accented words had a higher f_0 peak only if the words were produced in a frame sentence; there was no difference when produced in isolation (1984:137–38). Sugito (1968) found no difference in f_0 of the second mora of two-mora words produced in isolation using only one speaker, but found a difference for three out of 14 speakers in her 1982 work, even though the words were produced in isolation. Vance (1995) found a difference for one out of four speakers in an experiment similar to Sugito's. It is important to note that all of these researchers used very short words, with no more than two moras; /hana'/"flower" and /hana/"nose" form a commonly tested pair.

If one is trying to test Pierrehumbert and Beckman's theory against another theory, the use of very short words introduces a confound. In Pierrehumbert and Beckman's theory, a two-mora word with final accent would have both a phrasal H tone and an accent HL tone on top of each other (Pierrehumbert & Beckman, 1988:132–133). Since there is no space in time between the phrasal peak and the accent peak, the predicted difference in slope described above cannot be tested. In addition, the effect of having the H and HL tones overlapping is unknown. Finally, I have found that because of the difficulties pitch trackers have with the cessation of voicing, two-mora words produced in isolation often do not reach an f_0 peak by the point at which the pitch tracker stops producing reliable data. Therefore, data from words such as /hana'/ and /hana/ produced without a frame sentence may be

suspect, especially when the measurement point is defined as the f_0 peak of the syllable, as in most of the previous experiments.

It is also helpful to examine absolute f_0 at the phrasal peak, in addition to change in f_0 between the phrasal peak and the end of the word. Many researchers have found that peak f_0 in accented words (not necessarily final-accented) is higher than in unaccented words (Kubozono, 1993:87–90; Poser, 1984:138–144; Pierrehumbert & Beckman, 1988:46), but these results are generally based on words with first or second mora accent, such as /uma'i/"tasty" *vs.* /amai/"sweet" (used by Poser, Pierrehumbert and Beckman, and Kubozono). In these cases, the second mora of an accented word (/uma'i/) has, in Pierrehumbert and Beckman's theory, both a phrasal H tone and an accent HL tone (Pierrehumbert & Beckman, 1988:132–133), so the difference can be accounted for by assuming that the HL tone is higher than the H tone. However, in longer words, the HL and H tones are separated, and Pierrehumbert and Beckman's theory would predict no difference in the phrasal peak depending on the presence or absence of an accent later in the word. This is related to the confound for short final-accented words discussed above.

2. Design of the experiment

2.1. *Materials and speakers*

I recorded six native speakers of Japanese reading the following 21 sentences in both regular and reiterant speech.

(3)

1.	[otʃa-kumi'da]	"It's a tea-pouring."
2.	[baba-nuki' da]	"It's Old Maid (game)."
3.	[imooto' da]	"It's a little sister."
4.	[otooto' da]	"It's a little brother."
5.	[ano kawa' da]	"It's that river."
6.	[ano natsu' da]	"It's that summer."
7.	[ano muɾa' da]	"It's that village."
8.	[utʃi-no-çito' da]	"It's my husband."
9.	[bake-no-kawa' da]	"It's a false skin."
10.	[ano hanaʃi' da]	"It's that discussion."
11.	[ano çikaɾi' da]	"It's that light."
12.	[nana-satsu-me' da]	"It's the seventh book."
13.	[kokonoka-me' da]	"It's the ninth day."
14.	[haha-oja da]	"It's the mother."
15.	[ne-sagaɾi da]	"It's a lowering in price."
16.	[tomodatʃi da]	"It's a friend."
17.	[ano momo da]	"It's that peach."
18.	[tobi-agaɾi da]	"It's a jumping upward."
19.	[çiki-otoʃi da]	"It's a pulling over (wrestling move)."
20.	[tamago-jaki da]	"It's a fried egg."
21.	[ano sakana da]	"It's that fish."

The noun phrases in sentences 1 to 13 were chosen as the final-accented items: they have final accent (for the noun) as either the only possibility or the preferred pronunciation according to the Nihon Hoosoo Kyookai accent dictionary (1985).

Because accent placement varies even within the Tokyo dialect, I also consulted with Professor Yoko Hasegawa, a native spaker of the Tokyo dialect, and used noun phrases she thought were likely to be pronounced with the accent as marked. The noun phrases in sentences 1 and 2 were recorded despite her prediction that they would not be final-accented. However, none of the speakers produced them with final accent, so these sentences were eliminated from the database. Unaccented words are very common, and have more stable accent assignment. Sentences 14 to 21 contain unaccented noun phrases. Thirteen final-accented examples and only eight unaccented examples were used on the assumption that some final-accented examples would be produced with an unexpected accent placement.

The noun phrases in 1–7 and 14–17 have four moras in the noun phrase itself; 8–13 and 18–21 have five moras. Four- or five-mora examples were used because shorter phrases do not have enough distance in time between the phrasal peak and the end of the word to show the effect being tested. Words longer than five moras, plus the frame sentence, are difficult to produce in reiterant speech. The very short frame sentence X *da* "It is X" was used because it is semantically appropriate for almost any noun, so that all noun phrases can have the same frame sentence. /da/ has no effect on accent placement. Longer frame sentences are more difficult to produce in reiterant speech, and would also have allowed for various phrasings, which would have affected the f_0 curves and increased error variance.

The noun phrases in some examples consist of /ano/ plus a noun, "that noun." /ano/ does not affect accent placement, but in strongly hyperarticulated speech, it can form a separate accentual phrase (or minor phrase) from the noun. In such cases, the noun itself becomes an accentual phrase too short to show the effect being tested. Although this was a problem in a pilot version of this experiment, none of the speakers in the present experiment used separate accentual phrases. (When separate accentual phrases are used, a dip in the f_0 curve is clearly visible between the two words.)

/ano/ was used to create longer noun phrases because of the difficulty of finding four- to five-mora words with stable final accent. Although some sentences consist of /ano/ plus a monomorphemic noun, some of a compound noun, and some of monomorphemic nouns alone, all of these types should be the same for prosodic purposes, because all the sentences were read as tightly connected single phrases. Also, all conditions contain examples both of /ano/ plus a noun and of a noun alone, so any unexpected effects should be controlled.

The speakers read the sentences both normally and as reiterant speech, in which they substitute [ma] for each syllable of the original sentence and attempt to keep all other factors unchanged. This method has been described and tested by Larkey (1983). Reiterant speech was used to remove the variation caused by segmental influences. It is almost impossible to find minimal pair sentences of four or five moras for final-accented *vs.* unaccented words, so the intrinsic pitch of differing vowels creates a problem. Reiterant speech also allows for the use of words in the modeled utterances containing obstruents and voiceless sounds. Measurements were made on both original and reiterant versions to check the validity of reiteration as a method. Figs. 3(a) and 3(b) illustrate segmental influences on f_0 in original speech (left side of figures), and the removal of those effects in reiterant speech (right side of figures): in Fig. 3(a), the original production of [ano hanaʃi' da] has of course no f_0 during [ʃ], and a depression of f_0 for [d]. In Fig. 3(b),

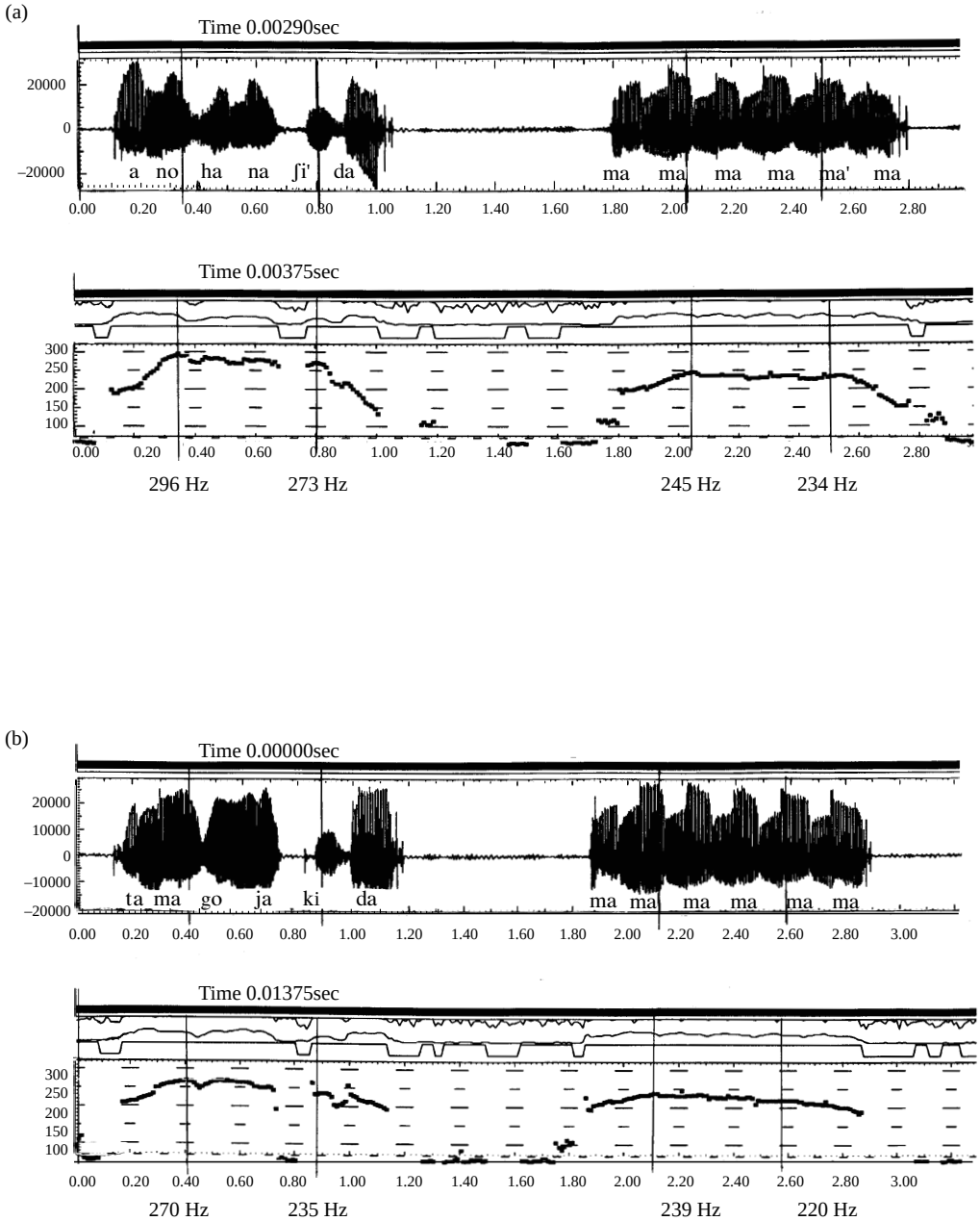

Figure 3. (a) Waveform, pitch track, and measurements for Speaker 3's production of sentence 10, containing the final-accented noun phrase [ano hanaʃi'], both original and reiterant productions. The labeled vertical lines show measurement points and f_0 in Hz at those points. As predicted, f_0 rises initially, falls slightly to the accented mora, then falls sharply after the accented mora. (b) Waveform, pitch track, and measurements for Speaker 3's production of sentence 20, containing the unaccented noun phrase [tamagojaki], both original and reiterant productions. The labeled vertical lines show measurement points and f_0 in Hz at those points. As predicted, f_0 rises to a peak in the second or third mora, then falls at a steady rate to the end of the phrase.

[tamagojaki da], there are depressions for both voiced stops, and a disturbance in the f_0 track when voicing resumes after [k]. In both figures, the reiterant versions are clearly free of such effects.

In summary, the sentences were chosen to include final-accented and unaccented noun phrases with relatively stable accent placement, to have either four or five moras, to be appropriate in the frame sentence, and to be common enough that subjects would be familiar with their pronunciation (and thus with their accent placement). Long syllables were avoided when possible so that speakers would not have to decide whether to substitute one long or two short *ma*'s. (Sentences 3 and 4 do include long syllables, because of the difficulty of finding stable final-accented words which met all the criteria. Speakers were prompted to use /maa/ for the long syllable, but not corrected if they used /mama/.) Otherwise, no attention was paid to segmental content because of the use of reiterant speech.

Six speakers participated in the experiment, but two were excluded from the analysis because they could not produce accurate reiterant speech: those two speakers frequently produced reiterations with the wrong number of *ma*'s. Of the other four, Speakers 1 and 2 produced reasonably good reiterations, while Speakers 3 and 4 were very good at reiterant speech. All speakers are female native speakers of Japanese. Speakers 3 and 4 speak strictly Tokyo dialect, Speaker 2 grew up partly in Tokyo, and Speaker 1 grew up in Tokyo but acquired a different dialect, which still has a Tokyo-type accentual system, from her mother. Speaker 1 is bilingual in English. Speakers 1 and 2 are linguists, but not phonologists. Speakers 3 and 4 are linguistically naive. None of the speakers knew the purpose of the experiment.

2.2. *Procedure*

The sentences in (3) were written on index cards in the usual Japanese orthography, a mixture of Chinese characters and the *hiragana* syllabary. The pronunciations of all Chinese characters were written above them in the *hiragana* syllabary, which disambiguates any confusions in the readings of the Chinese characters. The reiterant speech sequence (one *ma* per syllable, long mark after *ma* for long syllables) was written in the *katakana* syllabary below the words.

Each speaker sat in a recording booth and listened to a tape recording made by Professor Yoko Hasegawa demonstrating reiterant speech for sentences similar to those used in the experiment. They were then given a chance to practice using comparable sentences. They then read each of the 21 sentences, in randomized order, first in regular and then in reiterant speech. The reiterant reading of each sentence followed the regular reading immediately because it is easier to produce reiterant speech if one has just heard oneself say the same sentence in normal speech. Each speaker produced five to eight repetitions of the entire stack of cards.

I digitized the data on the CSL system, then used ESPS/Xwaves to produce f_0 contours and measure the f_0. Measurement points were the phrasal peak, usually in the second mora, and the beginning of the vowel in the last mora of the noun phrase itself (that is, the last mora before /da/). The vertical lines in Fig. 2 show the measurement points in the hypothetical phrases, and the labeled vertical lines in Figs. 3(a) and 3(b) show the actual measurement points for two typical tokens. In Figs. 3(a) and 3(b) one can see that the first measurement point has been placed where the f_0 contour turns and becomes flat or falling. Because of slight f_0 variations during the high stretch of the contour, locating the phrasal peak is not always

straightforward. The point at which the slope of the f_0 curve changes sharply was chosen, even if it was not the absolute peak. The second measurement point was placed shortly after the beginning of the vowel in the last mora of the noun phrase, at a point at which the f_0 had become stable.

In addition to evaluating the absolute f_0 at the first measurement point (the phrasal peak), f_0 at the first measurement point was also subtracted from f_0 at the second measurement point, so that a fall in f_0 between the two gives a negative number and a rise between the two gives a positive number. The f_0 at the phrasal peak and the difference between the two points are the two measures to be evaluated in this experiment. Pierrehumbert and Beckman's theory predicts that f_0 at the phrasal peak will be the same for both final-accented and unaccented noun phrases, and that both types will have negative values for the difference measure, but the unaccented noun phrases will have larger negative values. The traditional theory predicts that whatever they are, both the first measurement point and the difference measure should be the same for both types.

Whether final-accented and unaccented noun phrases are the same or different is the main point of interest in this study, but the experiment can give us more information. The number of moras in the noun phrase can be treated as a factor, with four moras and five moras as the two levels. If the number of moras has a significant effect on phrasal peak or change in f_0, it would be interesting, but not important to the central question. Original *vs.* reiterant reading provides another factor, which can be used to test the validity of reiterant speech. If reiteration successfully maintains the prosody, one would expect that the factor of reiteration would not have a significant effect, but that the variance for original readings of the sentences would be higher than the variance for reiterant readings, which have eliminated segmental influences.

3. Results

As previously mentioned, sentences 1 and 2 were excluded for all subjects because they all placed the accent somewhere other than the expected final mora, usually on the third mora. For Speakers 1 and 2, additional data were excluded because these subjects consistently had difficulty producing reiterant speech for certain sentences (as judged by the fact that they verbally expressed doubt about their productions of those sentences), or they placed the accent on an unexpected mora for a particular sentence. For these reasons, all productions (both original and reiterant, all tokens) of the following sentences were excluded: 7, 10, and 21 in the data of Speaker 1; and 6, 9, 14, and 17 in the data of Speaker 2. No data were excluded simply because the researcher felt that the reiterant production was inaccurate.

For each subject, the remaining scores were averaged across token number and sentence number within each condition. (A condition consists of all sentences which have the same number of moras, the same accent type, and are all either original or reiterant. Reiterant readings of sentences 8 to 13, for example, compose the five mora, final-accented, reiterant condition.) This gave a total of 8 scores (2 mora lengths $\times$ 2 accent types $\times$ 2 reading types) for each subject, each of which was derived from between 15 and 42 utterances, depending on how many tokens of each sentence the subject produced and on how many sentences were in the condition. The averaged scores were analysed as a three-factor within-subjects design, where

TABLE I. Average f_0 in Hz at the first measurement point (the phrasal peak) for each speaker (S1–S4) in each condition, with standard deviations in parentheses

Final accent original speech	4 moras	5 moras	Final accent reiterant speech	4 moras	5 moras
S1	245.9 (8.52)	257.6 (12.76)	S1	234.9 (6.90)	246.8 (8.12)
S2	255.7 (9.73)	264.8 (10.38)	S2	240.9 (8.03)	244.5 (7.93)
S3	264.2 (13.36)	271.1 (14.30)	S3	238.5 (12.69)	244.0 (9.22)
S4	251.5 (6.64)	258.0 (8.78)	S4	244.3 (6.80)	249.3 (6.07)
Unaccented original speech	4 moras	5 moras	Unaccented reiterant speech	4 moras	5 moras
S1	234.5 (10.08)	244.8 (8.42)	S1	224.2 (9.77)	232.9 (8.52)
S2	253.4 (8.30)	258.1 (8.87)	S2	235.5 (7.23)	240.6 (7.92)
S3	254.6 (10.44)	267.8 (12.01)	S3	235.0 (8.86)	242.9 (9.57)
S4	245.1 (6.33)	252.7 (5.64)	S4	238.9 (6.14)	245.3 (5.96)

the three factors are: accent type, with the levels final-accented and unaccented; reiteration, with the levels original speech and reiterant speech; and number of moras, with the levels four moras and five moras.

3.1. *f_0 at the phrasal peak*

Table I shows the absolute f_0 at the first measurement point for each subject and each condition. The averages for each condition, collapsed across the other two factors and all speakers, appear in Table II. Table III shows the results of the main analysis of variance. At this measurement point (the phrasal peak), final-accented tokens have significantly higher f_0 than unaccented tokens, original speech has

TABLE II. f_0 in Hz at the first measurement point (the phrasal peak) for each level of each factor, averaged across all speakers, both other factors, and all tokens

	f_0 (Hz)
Final-accented	250.7
Unaccented	244.1
Original speech	255.0
Reiterant speech	239.9
4 moras	243.6
5 moras	251.3

TABLE III. Results of the main analysis of variance for f_0 at the first measurement point. Sources printed in bold are significant

Source	$F(1, 3)$	
Accent type	**12.45**	$p < 0.039$
Reiteration	**16.32**	$p < 0.027$
No. of moras	**47.93**	$p < 0.006$
Accent type by reiteration	1.39	$p > 0.05$
Accent type by no. of moras	0.08	$p > 0.05$
Reiteration by no. of moras	**11.14**	$p < 0.045$
Accent type by reiteration by no. of moras	0.003	$p > 0.05$

significantly higher f_0 than reiterant speech, and five-mora noun phrases have significantly higher f_0 than four-mora noun phrases. In addition, the interaction between reiteration and number of moras is significant: the tendency for original productions to have higher f_0 than reiterant productions is slightly (2 Hz) greater for five-mora noun phrases than for four-mora noun phrases.

Because the interaction of reiteration and number of moras is significant, one must examine the simple effects (the effect of a factor at each level of the other factor separately) to see whether the main effects are representative. Original tokens have significantly higher f_0 than reiterant tokens both for the five-mora examples ($F(1, 3) = 16.14,\ p < 0.028$) and for the four-mora examples ($F(1, 3) = 16.49,\ p < 0.027$). Similarly, the effect of number of moras (five mora higher than four mora) is significant for both original ($F(1, 3) = 72.61,\ p < 0.003$) and reiterant ($F(1, 3) = 27.93,\ p < 0.013$) speech. Therefore, the interaction of reiteration and number of moras, while significant, is very small. The main effects discussed above can be considered reliable.

3.2. *Change between measurement points*

Table IV shows the scores for the difference between measurement points for each subject and each condition. The averages for each condition, collapsed across the other two factors and all speakers, appear in Table V. Looking just at these averages, one sees that f_0 falls for all the conditions, but unaccented noun phrases fall more than final-accented, original more than reiterant, and five mora than four mora. Table VI shows the results of the main analysis of variance for the difference measure. The main effects of accent type (unaccented fall more than final-accented) and number of moras (five mora fall more than four) are significant.

The interactions of accent type by reiteration and reiteration by number of moras are also significant. However, the three-way interaction, the interaction of accent type by number of moras, and the main effect of reiteration, are not. Since two of the interactions are significant, it is again important to examine the simple effects of the factors involved. First, we can examine graphs of the significant interactions, shown in Fig. 4. If there were no interaction, the two lines in the graph would be parallel. These graphs show that although the interactions are significant, they are very small. For the interaction of accent type with reiteration, the tendency of unaccented noun phrases to fall more than final-accented is stronger for original than for reiterant speech. For the number of moras by reiteration interaction, the

TABLE IV. Average change in f_0 in Hz between the two measurement points for each speaker in each condition, with standard deviation in parentheses

Final accent original speech	4 moras	5 moras	Final accent reiterant speech	4 moras	5 moras
S1	2.1 (8.10)	−5.0 (7.81)	S1	0.6 (4.69)	1.8 (4.30)
S2	−11.0 (9.24)	−25.7 (9.71)	S2	−4.4 (4.10)	−12.8 (6.82)
S3	−5.4 (8.13)	−11.8 (8.95)	S3	−4.2 (4.01)	−9.4 (4.27)
S4	2.7 (6.93)	−8.1 (9.13)	S4	−0.6 (4.21)	−9.4 (5.37)
Unaccented original speech	**4 moras**	**5 moras**	**Unaccented reiterant speech**	**4 moras**	**5 moras**
S1	−4.0 (7.35)	−15.2 (6.23)	S1	−5.3 (3.95)	−2.7 (5.00)
S2	−21.7 (9.78)	−38.8 (8.86)	S2	−8.5 (3.87)	−18.9 (7.76)
S3	−10.0 (5.97)	−25.4 (5.06)	S3	−10.1 (4.15)	−15.4 (4.56)
S4	−7.3 (5.15)	−18.8 (5.34)	S4	−8.5 (3.16)	−15.1 (4.70)

TABLE V. Change in f_0 between measurement points for each level of each factor, averaged across all speakers, both other factors, and all tokens

Final-accented	−6.3
Unaccented	−14.1
Original speech	−12.7
Reiterant speech	−7.7
4 moras	−6.0
5 moras	−14.4

TABLE VI. Results of the main analysis of variance for the difference measure. Sources printed bold are significant

Source	$F(1, 3)$	
Accent type	**301.83**	$p < 0.001$
Reiteration	3.03	$p > 0.05$
No. of moras	**20.36**	$p < 0.02$
Accent type by reiteration	**20.67**	$p < 0.02$
Accent type by no. of moras	2.81	$p > 0.05$
Reiteration by no. of moras	**17.39**	$p < 0.025$
Accent type by reiteration by no. of moras	5.92	$p > 0.05$

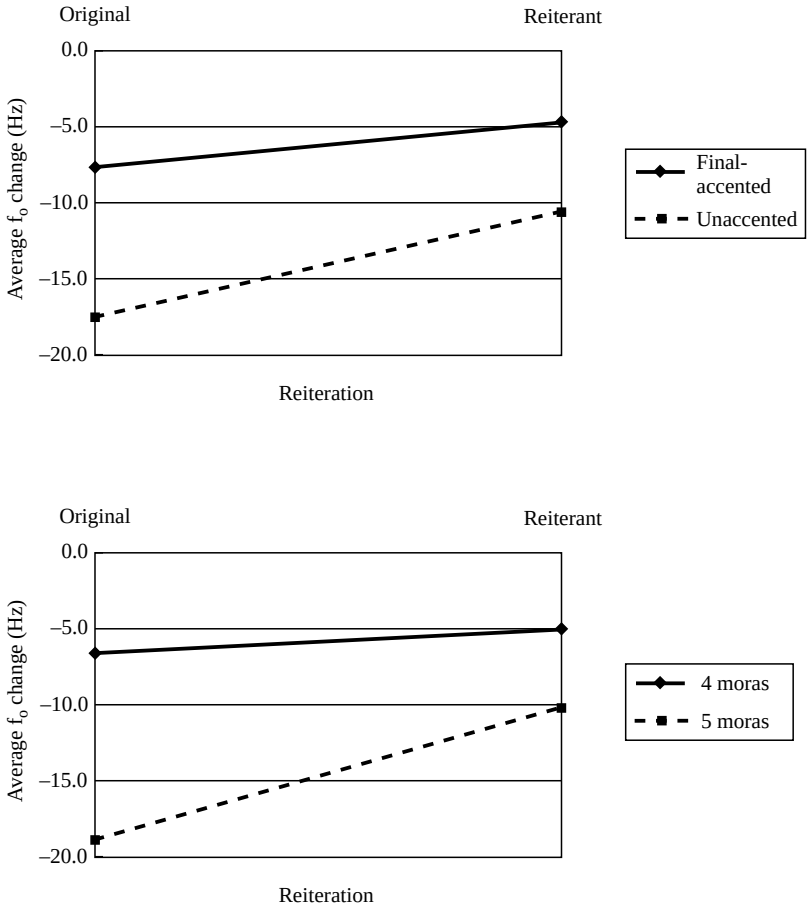

Figure 4. Graphs representing the reiteration by accent-type interaction (top) and the reiteration by number of moras interaction (bottom).

tendency of five-mora noun phrases to fall more than four-mora noun phrases is stronger for original than for reiterant speech.

A look at the simple effects confirms that the interactions are small, though significant. The simple effects of reiteration both at final-accented and at unaccented levels are non-significant, at $F(1, 3) = 1.41$ ($p > 0.05$) and $F(1, 3) = 4.62$ ($p > 0.05$), respectively. Simple effects of accent type are significant for both original ($F(1, 3) = 148.08$, $p < 0.001$) and reiterant speech ($F(1, 3) = 213.45$, $p < 0.001$). That is, the effect of reiteration is not significant for either accent type, while the effect of accent type is significant for both original and reiterant speech. Thus, the significant accent type by reiteration interaction is not very meaningful.

Reiteration is not significant at either level of number of moras, with $F(1, 3) = 0.37$ ($p > 0.05$) for four moras and $F(1, 3) = 6.87$ ($p > 0.05$) for five moras. However, while the effect of number of moras is significant for original speech, with $F(1, 3) = 66.39$ ($p < 0.004$), it is not significant for reiterant speech, with $F(1, 3) = 4.22$ ($p > 0.05$). Since none of the simple effects of reiteration at any level of either other factor is significant, we can trust the non-significant main effect and conclude

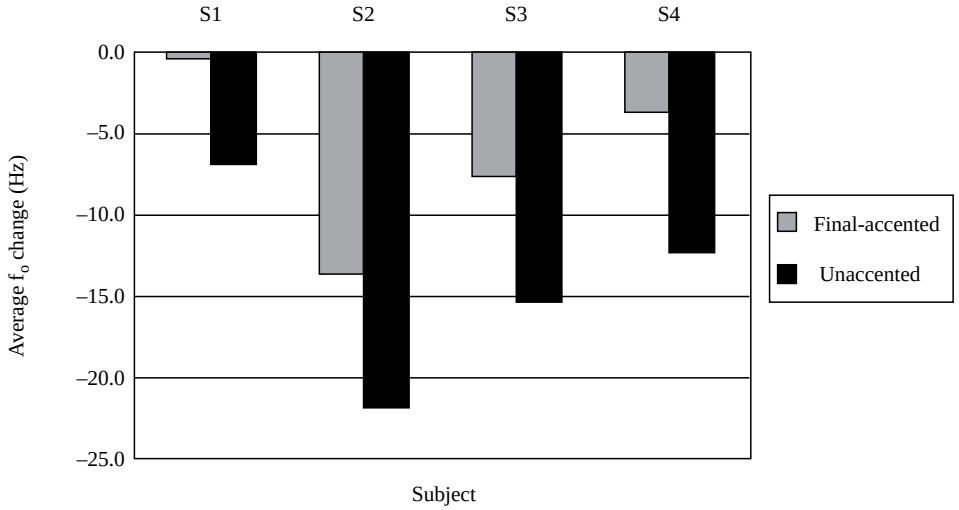

Figure 5. Average f_0 change for final-accented and unaccented phrases for each speaker.

that reiteration *vs.* regular speech does not make a difference in the results. Original speech does have higher variance than reiterant speech; variance for all original speech scores grouped together is 10.84, while for all reiterant tokens it is 6.21.

Fig. 5 shows each subject's average score for the levels of the accent type factor. All four subjects have f_0 falling more for unaccented than for final-accented tokens. Why do the subjects vary so much in the degree of f_0 fall? In Pierrehumbert and Beckman's theory, declination and final lowering are parameters set for each speaker (and viewed as separate effects: Pierrehumbert & Beckman, 1988:70–76). These data fit with their theory if we assume that Speaker 1 has almost no declination or final lowering, while Speaker 2 has a large effect from one or both of those factors. Pierrehumbert and Beckman do not give a range of possible values speakers might have for their various speaker-dependent variables. However, the final lowering data they present for two of their speakers (1988:74) seem to show that in the penultimate mora of the prosodic phrase, final lowering accounts for a drop of about 15–20 Hz. Both of the speakers whose data are shown are male, so an even greater effect could be expected for female speakers. Therefore, even Speaker 2, who has an f_0 fall in final-accented tokens of 13.5 Hz, fits well within the range which can be explained by final lowering and declination. (The final-accented scores are the ones to consider for this point, since the unaccented scores include the effect of interpolation from a high to a low tone as well as final lowering and declination.)

4. Conclusions

Accent type, which is the main topic of this experiment, has a signficant effect both for the first measurement point alone and for the difference between the two measurement points: at the first measurement point (phrasal peak), final-accented noun phrases already have higher f_0 than unaccented ones, and between the two measurement points, f_0 falls more in unaccented noun phrases than in final-accented ones. This difference at the first measurement point, in which the realization of the

phrasal peak depends on the presence or absence of an accent up to three moras later, cannot be predicted by Pierrehumbert and Beckman's left-to-right theory. This confirms Kubozono's observation of problems for a strictly left-to-right theory based on f_0 differences in the first mora of the phrase (1993:93). Pierrehumbert and Beckman state that although they ''have not made use of phonetic realization rules that refer to the rightward phonological context of the element being produced,'' they ''would not be surprised to find that such rules exist'' (1988:162). Still, this higher f_0 for final-accented noun phrases several moras before the accent cannot currently be predicted by their theory. The traditional theory, of course, cannot predict this difference either, since it assigns the same tones to both accent types.

The significant effect of accent type on the difference measure (unaccented noun phrases fall more than final-accented, but both fall) is in accordance with Pierrehumbert and Beckman's theory. This result is not at all compatible with the traditional theory. The fact that many linguists, including many native speakers of Japanese, failed to detect this difference when working by ear may indicate that the different degree of f_0 fall within the word is not a major perceptual cue for distinguishing final-accented from unaccented words. (Some linguists, usually those supporting the three-level theory, did hear the difference. Sugito (1982:193–200) found that some listeners can perceive this difference.) It is almost certain that the main cue to place of accent in Japanese is the sharp fall in pitch on the mora following the accent (Hasegawa & Hata 1992:87,88). However, even if the difference within the word is not an major perceptual cue, my data show that it is produced consistently. These results provide strong support for Pierrehumbert and Beckman's theory on this point, regardless of any significance for perception.

Number of moras also has a significant effect both for the first measurement point and for change between the two points: five-mora noun phrases have a higher phrasal peak than four-mora noun phrases, and fall more between the measurements points. The difference at the first measurement point could be explained by longer utterances generally starting at higher f_0 (Fujisaki & Hirose, 1982:65; Hirai, Iwahashi, Higuchi & Sagisaka, 1996:7). As for the difference in degree of fall, in a five-mora noun phrase, the proportion of the duration of the sentence which is located between the two measurements points is greater than in a sentence with a four-mora noun phrase. Therefore, a greater proportion of the fall gets measured.

The reiteration factor is significant for the first measurement point, with original tokens having higher f_0 than reiterant tokens. This reflects an overall flattening and lowering of the f_0 curve in reiterant tokens, which one can see in Figs. 3(a) and 3(b): the entire pitch range appears to be lower and smaller in reiterant speech (or in the second sentence of a repeated pair, the environment of the reiterant tokens).

For the change between measurement points, the reiteration factor appears to be decidedly non-significant, which is a good sign for the validity of reiterant speech as a method. There are significant interactions with reiteration but, as discussed above, these interactions are small. Both significant interactions might reflect the fact that reiterant tokens tend to be flatter than original speech tokens, as discussed above, so that there is less fall to measure in the reiterant tokens. The tendency for five-mora noun phrases to fall more than four-mora noun phrases, or for unaccented noun phrases to fall more than final-accented, could be stronger for original than for reiterant speech simply because there is less fall in any of the reiterant tokens. Since the simple effect of reiteration at each level of each other factor is also not

significant, the non-significant main effect can be trusted. The reiterant tokens did have lower variance than the original, as expected. Thus, reiteration is probably a safe method for such studies. It has essentially the same results as original speech, but with lower variance.

Although this experiment was designed to test Pierrehumbert and Beckman's theory against the traditional (two-level) theory, the results for accent type, both for the first measurement point and for the difference measure, also have implications for some other theories of Japanese pitch accent. The higher f_0 for final-accented noun phrases at the first measurement point is predicted by the three-level theory, because it assigns high tones from the second mora of a final-accented phrase, but mid tones from the second mora of an unaccented phrase. This same result is also predicted by Fujisaki's model, since it posits a larger accent component for accented phrases of any type than for unaccented phrases (Fujisaki, 1993:4, Hirai *et al.,* 1996:10). (Fujisaki's model will not be described in detail here. For details, please see Fujisaki & Hirose, 1982, 1984, and Fujisaki, 1993.)

The effect of accent type on the difference in f_0 between the two measurement points, however, is not predicted by the three-level theory, even though the hallmark of the three-level theory was its ability to distinguish final-accented and unaccented words: the string LHHH for final accent and LMMM for no accent cannot result in a slope difference for the latter three moras of the word. Fujisaki's model does not appear to predict this result either, since the two accent types would have the same phrase command, and the accent commands, while of differing size, would both parallel the decay of the phrase command, producing no difference in slope. However, a fitting of phrase and accent components to these data with Fujisaki's model has not been attempted, so the exact results within this model are not known. In the noun phrases which consist of /ano/ plus a noun, Fujisaki's model would posit separate accent commands for /ano/ and the noun, and if the noun is accented, its accent command would be larger than that for /ano/. This would account for the lesser fall between the two measurement points for final-accented noun phrases (Fujisaki, 1996, personal communication). However, an inspection of the data for individual sentences reveals that the same result holds true for sentences which consist of a single word, which could not have two separate accent commands.

I conclude that the results for change between measurement points provide support for Pierrehumbert and Beckman's theory as applied to the possible difference between final-accented and unaccented words or phrases. The unaccented tokens do indeed fall more than the accented ones between the phrasal peak and the end of the word, and the two accent types are different within the noun phrase itself, without reference to the following syllable. None of the other theories predicts this difference in the change between measurement points. However, the greater f_0 for final-accented noun phrases at the phrasal peak cannot be predicted by a strictly left-to-right theory such as Pierrehumbert and Beckman's, and is predicted by some of the other theories. The experiment also confirms the reliability of reiterant speech for such studies.

I am grateful to John Ohala, Yoko Hasegawa, Charles Fillmore, Nick Campbell, and Norio Higuchi for discussions of this paper, and to Patrice Speeter Beddor, Timothy Vance, and an anonymous reviewer for their helpful comments. I would also like to thank Yoko Hasegawa for help with the materials for the experiment and Mike Ward and Shawn Ying for technical

assistance. Part of this work was carried out while at ATR and supported by the NSF Summer Institute in Japan.

References

Fujisaki, H. (1993) From information to intonation, *Keynote Lecture at the International Conference on Signal Processing, Beijing,* 1–12

Fujisaki, H. & Hirose, K. (1982) Modeling the dynamic characteristics of voice fundamental frequency with applications to analysis and synthesis of intonation, *Preprints of Papers for the Working Group on Intonation, The XIIIth International Congress of Linguists, Tokyo,* 57–70

Fujisaki, H. & Hirose, K. (1984) Analysis of voice fundamental frequency contours for declarative sentences of Japanese, *Journal of the Acoustical Society of Japan,* (E)**5**, 233–242

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

Hasegawa, Y. & Hata, K. (1992) Fundamental frequency as an acoustic cue to accent perception, *Language and Speech,* **35**, 87–98

Higurashi, Y. (1983) *The accent of extended word structures in Tokyo standard Japanese.* Tokyo: Educa

Hirai, T., Iwahashi, N., Higuchi, N. & Sagisaka, Y. (1996) Automatic extraction of f_0 control rules using statistical analysis. Unpublished manuscript, ATR

Kawakami, S. (1995) *Nihongo akusento ronshuu* [Collected papers on Japanese accent]. Tokyo: Kyuko Shoin

Kubozono, H. (1993) *The organization of Japanese prosody.* Tokyo: Kurosio

Larkey, L. S. (1983) Reiterant speech: an acoustic and perceptual validation, *Journal of the Acoustical Society of America,* **73**(4), 1337–1345

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

McCawley, J. D. (1977) Accent in Japanese. In *Studies in stress and accent. Southern California occasional papers in linguistics 4* (L. Hyman, editor), pp. 261–302. Los Angeles: Department of Linguistics, University of Southern California

Nihon Hoosoo Kyookai (1985) *Nihongo hatsuon akusento jiten.* [Japanese pronunciation and accent dictionary]. Tokyo: Nihon Hoosoo Shuppan Kyookai

Pierrehumbert, J. B. & Beckman, M. E. (1988) *Japanese tone structure. Linguistic Inquiry, Monograph 15.* Cambridge, MA: MIT Press

Poser, W. J. (1984) *The phonetics and phonology of tone and intonation in Japanese.* MIT dissertation

Sakuma, K. (1929) *Nihon onseigaku* [Japanese phonetics]. Tokyo: Kyoobunkan

Sugito, M. (1968) Tokyo nihaku-go odaka to heiban akusento-koo [Comparison between "high-final" and "level" tone in the Tokyo accent], *Onsei Gakkai Kaihoo,* **129,** 1–4

Sugito, M. (1982) *Nihongo akusento no kenkyuu* [Research on Japanese accent]. Tokyo: Sanseido

Uwano, Z. (1977) Nihongo no akusento [Japanese accent]. In *Iwanami kooza nihongo 5: On'in* [Iwanami Japanese course 5: phonology] (S. Ono & T. Sibata, editors), pp. 281–321. Tokyo: Iwanami Shoten

Vance, T. J. (1987) *An introduction to Japanese phonology.* Albany: State University of New York Press

Vance, T. J. (1995) Final accent vs. no accent: utterance-final neutralization in Tokyo Japanese, *Journal of Phonetics,* **23**, 487–499

Weitzman, R. S. (1970) *Word accent in Japanese. Studies in the phonology of Asian languages 9.* Acoustic phonetics research laboratory, University of Southern California