Title
Fundamental Frequency as an Acoustic Cue to Accent Perception

Journal
Language and Speech, 35(1-2)

ISSN
0023-8309

Authors
Hasegawa, Yoko
Hata, Kazue

Publication Date
1992

DOI
10.1177/002383099203500208

Peer reviewed

FUNDAMENTAL FREQUENCY AS AN ACOUSTIC CUE TO ACCENT PERCEPTION*

YOKO HASEGAWA
University of California, Berkeley
and
KAZUE HATA
Speech Technology Laboratory, Santa Barbara, CA

It is generally recognized that the most significant cue to accent location is fundamental frequency (F_0) in both Japanese and English. Furthermore, it is widely believed that a syllable is perceived as accented if the syllable contains an F_0 peak. However, Sugito (1972) found that, in Japanese, if an F_0 peak is followed by a steep F_0 fall, the syllable preceding the F_0 peak may be perceived as accented. In this article we present two experiments which investigate the relationship between F_0 peak and F_0 fall rate in accent perception for Japanese and English. The first experiment confirms that, for Japanese, both F_0 peak location and F_0 fall rate affect listeners' judgments of accent location. Specifically, the later the F_0 peak occurs in a given syllable, relative to the syllable boundary, the greater the F_0 fall rate necessary for listeners to perceive the preceding syllable as accented. The second experiment shows that this phenomenon is not unique to Japanese: Perception of accent location in English is also influenced by both F_0 peak location and post-peak F_0 fall rate.

Key words: accent, pitch accent, stress, Japanese.

INTRODUCTION

The term 'accent' refers to a phenomenon which is fundamental to linguistic description and theory. Onishi (1942) claimed that the function of accent is to either differentiate meanings of words, or to make a portion of words or phrases prominent for discourse reasons. He analyzed 'accent' as being an impressionistic sum of any features that could serve these purposes. Fry (1955, 1958) discovered that potential perceptual cues to accent are pitch, length, loudness, and sound quality. In neutral declarative

* Parts of this research were previously reported in Hata and Hasegawa (1988a, b) and Hasegawa and Hata (1990). We are grateful to Professor John Ohala for his guidance and encouragement in conducting research in experimental phonology. We also wish to thank the following individuals for their comments and suggestions at various stages of this project: Ted Applebaum, Gregory De Haan, Glen Grosjean, Carlos Gussenhoven, Brian Hanson, Hector Javkin, John Kingston, Hideo Komatsu, Steve Pearson, Hisashi Wakita, Raymond Weitzman, and Helen Wheeler.

Address correspondence to Kazue Hata, Speech Technology Lab, 3888 State Street #202, Santa Barbara, CA 93105.

intonation, the accented syllables typically carry, relative to non-accented syllables, higher F_0 and amplitude, longer duration, and various differences in spectral patterns (e.g., differing energy distribution among vowel formants) in the acoustic domain. While such acoustic cues to accent have been identified, their relative perceptual significance has yet to be established. Commonly used terms such as *stress* and *pitch accent* presuppose the supremacy of one cue over the others (Fry, 1958, p. 128). In this article, *accent* is used to refer to prominence for linguistic purposes both in words and in sentences, regardless of the acoustic cue(s) by which such prominence is achieved.

Beckman (1986) has reported that while F_0 is the most effective cue in English, monolingual American English speakers show great variability in their use of cues in judging English accent. Although the relative importance of those cues differs from study to study,[1] researchers generally agree that F_0 is the most important cue to accent perception in English.

In Japanese, where a quantity opposition in vowel length is phonemic, duration plays virtually no role in accent production/perception (Oyakawa, 1971; Mitsuya and Sugito, 1978; Beckman, 1986). Amplitude, on the other hand, was once believed to be distinctive (Onishi, 1942; Neustupný, 1966), but later studies have shown that it has little influence (Weitzman, 1969; Sugito, 1972; Fujisaki and Sugito, 1977; Beckman, 1986; Beckman and Pierrehumbert, 1986). Thus, accent in Japanese is principally manifested by F_0 modulation.

One might naturally conclude, then, that if an F_0 peak occurs within a syllable, the syllable is perceived as accented. However, Neustupný (1966) found for Japanese that the perceived accent and the actual F_0 peak do not necessarily synchronize: The listener may perceive an accent on a syllable even when the F_0 peak does not occur on it.[2] The same phenomenon has been reported for English (Lehiste and Peterson, 1961), for Swedish (Bruce, 1977), and for German (Kohler, 1990).

Sugito (1972) found that this illusory accent placement is due to the F_0 contour falling after the peak: If the peak is followed by a steep F_0 fall, the listener perceives the preceding syllable as accented. This phenomenon explains why native Japanese

---

[1] The relative importance of cues to accent depends on the position of an accented syllable in a word (McClean and Tiffany, 1973), on the syntactic structure in which the word appears (Nakatani and Aston, 1978) and on the type of intonational focus (i.e., narrow, broad, or neutral) within which the word appears (Eady, Cooper, Klouda, Mueller, and Lotts, 1986).

[2] Fujisaki, Morikawa, and Sugito (1976) have suggested that the desynchronization of the F_0 peak and the syllable boundary in acoustic data is not psychologically real, but instead is a reflection of different processing times required to detect F_0 changes *vs.* segmental boundaries. They claim that listeners can detect F_0 changes faster than they can detect segmental boundaries, and therefore that the two phenomena synchronize in perception. Javkin (1976) and Maddieson (1976) conducted experiments to determine timing factors involved in the recognition of F_0 changes and segmental boundaries, and their results do not provide conclusive evidence for Fujisaki *et al.*'s hypothesis.

TABLE 1

F_0 fall rates of stimuli in Experiment 1

Stimulus	Rate (Hz/csec)	Stimulus	Rate (Hz/csec)	Stimulus	Rate (Hz/csec)
1	2.4	6	3.7	11	8.3
2	2.5	7	4.1	12	11.0
3	2.8	8	4.7	13	16.5
4	3.0	9	5.5	14	33.0
5	3.3	10	6.6		

listeners can perceive an accent on a syllable with a devoiced vowel. Although there is no F_0 on the devoiced vowel, a steep F_0 fall on the following syllable forces the listener to associate an accent with the devoiced vowel.

The current study investigates the influence of F_0 peak location and the rate of the immediately following F_0 fall on the perception of accent placement. In two experiments, one with Japanese and another with English, we tested the hypothesis that there is a correlation between the F_0 peak location and the F_0 fall rate immediately after the peak when the F_0 peak does not occur on a perceived accented syllable: the later the peak, the greater the F_0 fall rate.

EXPERIMENT 1: ACCENT PERCEPTION IN JAPANESE

Methods

Using STLtalk, and MITalk'79-based speech synthesis system developed at Speech Technology Laboratory (Javkin, Hata, Mendes, Pearson, Ikuta, Kaun, De Haan, Jackson, Zimmermann, Wise, Henton, Gow, Matsui, Hara, Kitano, Lin, and Lin, 1989), we synthesized 108 nonsense stimuli of the form /mamama/ with two variables: F_0 peak location and F_0 fall rate. Considering Steele's (1986) claim that F_0 peak location varies according to the vowel duration in accent placement, we prepared two sets of stimuli: one with a vowel duration of 100 msec for each of the three vowels, and the other with a 130-msec vowel duration. The duration of all /m/'s was fixed at 70 msec. The F_0 contour of the stimuli had a rise-fall shape with the starting F_0 at 125 Hz, linearly ascending to 160 Hz, and ending at 80 Hz. The magnitude of the F_0 fall within the middle syllable was always 33 Hz (four semitones), while the rate of F_0 fall varied as shown in Table 1.

The duration of F_0 fall varied from 14 csec to 1 csec in 1-csec steps, corresponding to fall rates between 2.4 Hz/csec and 33 Hz/csec. This range of fall rates was suggested

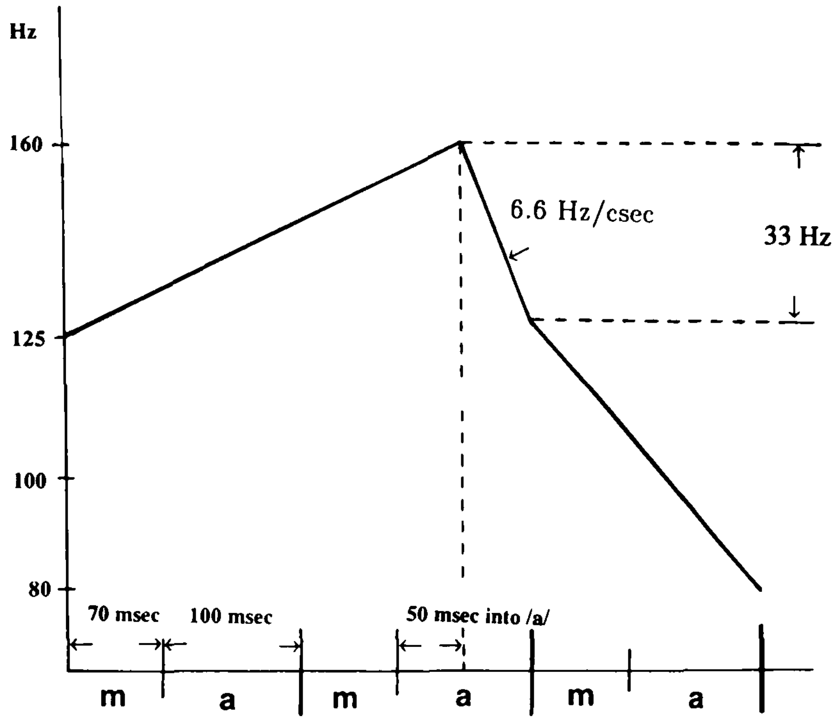

Fig. 1. F_0 contour of a sample stimulus in Experiment 1.

by a prior production study (Hasegawa and Hata, 1988). The F_0 peak occurred approximately 20, 30, 50, 60, or 70% into the second vowel of /mamama/. Figure 1 shows the F_0 contour of a sample stimulus where an F_0 peak occurs about 50% into the second vowel, and the F_0 fall rate is 6.6 Hz/csec. To reduce the length of the listening test, some F_0 peak locations had fewer than 14 fall-rate variations. (A pilot study confirmed that when the peak is at or beyond the 50%-location, lower fall rates do not influence accent perception. So we did not consider such cases in this experiment.)

The two sets of stimuli were separately randomized and presented to 15 adult native speakers of Japanese, who were asked to determine whether the accent pattern was like /námida/ 'tear' (accent on the first syllable) or like /okási/ 'snack' (accent on the second syllable) for each stimulus.

Results and discussion

Figures 2 and 3 show the subjects' responses for the vowel durations of 100 msec and 130 msec, respectively. In these figures, the abscissa represents the F_0 fall rate in Hz/csec, the ordinate represents the percentage of subjects' responses in which the first syllable was perceived as accented, /mámama/, and different curves represent the different F_0 peak locations within the second vowel (i.e., 20%, 30%, 50%, 60%, and

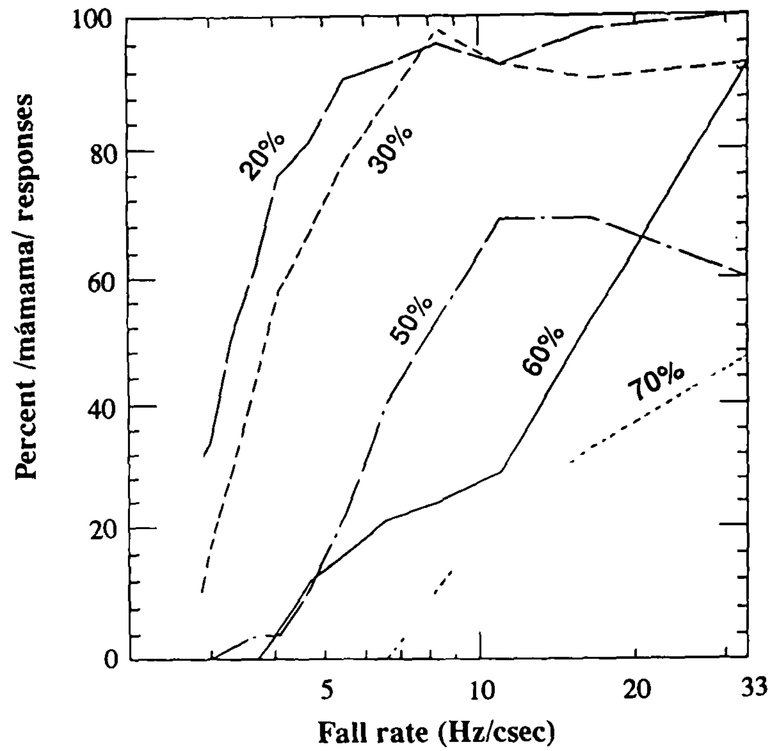

Fig. 2. Responses of /mámama/ in percent as a function of F_0 fall rate for various F_0 peak locations in Experiment 1. Each curve is labeled with its corresponding F_0 peak location, 20, 30, 50, 60, or 70 percent into the second vowel (vowel durations of 100 msec).

70% into the vowel). The curves were fitted to the raw data with the scatter plot smoothing method *LOWESS* (Becker and Chambers, 1984).

The results show the following tendency: The later the F_0 peak in the second vowel, the greater the F_0 fall rate required for a subject to associate an accent with the first syllable. For the 100-msec vowel stimuli with a 20% or 30% F_0 peak location, more than half of the subjects perceived the first syllable as accented, even when the F_0 fall was as gradual as 3–4 Hz/csec. When the F_0 peak occurred at the 50% location, an F_0 fall of approximately 8 Hz/csec was necessary for the same judgment. At the 60% location, a much steeper fall of 16 Hz/csec was required. Furthermore, at the 70% location even the steepest fall (33 Hz/csec) failed to shift the perceived accent to the first syllable, i.e., the majority of subjects judged the second syllable as accented. The same tendency for later peaks to require steeper falls was observed in the 130-msec vowel stimuli.

There was a difference in subjects' judgments between the two vowel durations, however. The location of the F_0 peak was found to have a limit beyond which the accent was never perceived on the preceding syllable even when the F_0 fall was very steep. This limit depended on vowel duration. For the 100-msec vowel stimuli, the fall rate

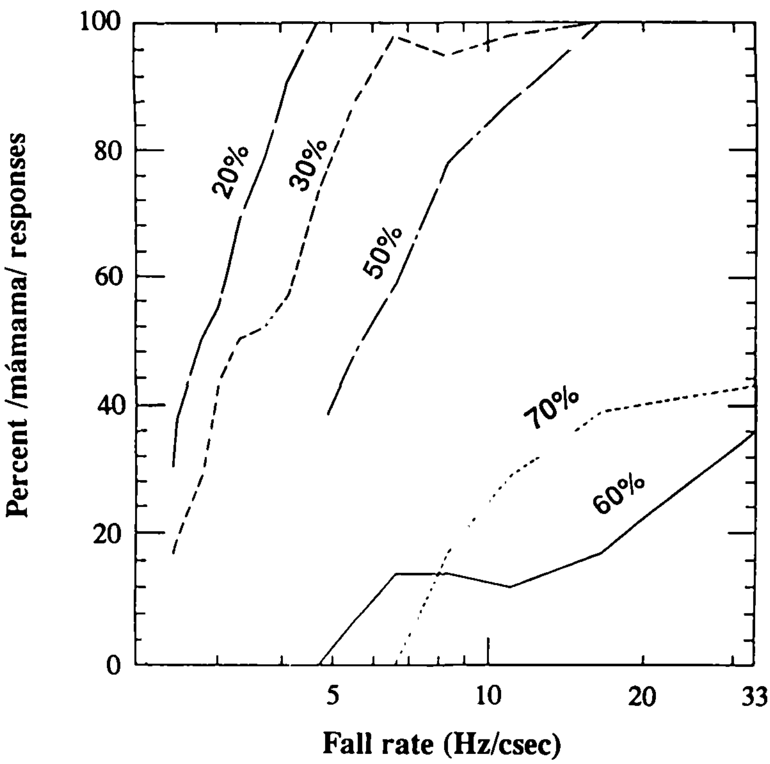

Fig. 3. Responses of /mámama/ in percent as a function of F_0 fall rate for various F_0 peak locations in Experiment 1. Each curve is labeled with its corresponding F_0 peak location, 20, 30, 50, 60, or 70 percent into the second vowel (vowel durations of 130 msec).

of 33 Hz/csec yielded 93% identification of the accent on the first syllable at the 60% F_0 peak location. But when the peak was 70% into the second vowel, the first syllable was not perceived as accented. In this case, the limit occurs somewhere between the 60% and 70% locations. For the 130-msec vowel stimuli, on the other hand, accent was never perceived on the first syllable beyond the 60% location, even when the fall was as steep as 33 Hz/csec. It seems that, as the vowel becomes shorter, a somewhat later F_0 peak (in terms of ratio to the vowel duration) can shift the perceived accent when the peak is followed by a steep fall.

We conclude that F_0 as a cue to Japanese accent must be understood as an integration of the F_0 peak and the post-peak fall rate. Since F_0 has been recognized as the prominent cue to accent in Japanese, and we manipulated only F_0 in the experiment, clearly F_0 peak location and F_0 fall rate are two acoustic correlates of accent perception.

Fry (1958) and Morton and Jassem (1965) claim that F_0 change itself is sufficient for listeners to perceive accent and that the magnitude of the F_0 change, unlike amplitude and duration, is insignificant. We believe that, although the magnitude may not matter, the rate of F_0 change does, at least in the case of Japanese listeners. In the second experiment, we explore this effect in English to see whether native speakers of English also exhibit such a correlative effect.

EXPERIMENT 2: ACCENT PERCEPTION IN ENGLISH

This experiment investigates whether or not F_0 peak location and F_0 fall rate influence the perception of accent location in declarative intonation in English. The focus of an English utterance, if any, is often expressed by placing the so-called contrastive accent on a certain syllable within the focused constituent (Bolinger, 1954, 1961; Halliday, 1967; Chafe, 1976; Lambrecht, 1986). For example, when one answers the question 'Is this her net?' the contrastive accent is on *my* in 'This is my net' On the other hand, in neutral (unmarked) intonation, e.g., responding to 'What's this?' the nuclear accent, which is the most prominent accent in an intonation phrase, occurs on the last word, *net*. Due to the coupling with utterance-final lowering of F_0, the fall rate must be greater if the nuclear accent is on the final syllable than if it is elsewhere in the utterance (Olive, 1974; Maeda, 1976). Therefore, such utterances are likely candidates for observing a correlative effect of F_0 peak location and F_0 fall rate, if this effect occurs in English.

Methods

Using the same system as in Experiment 1, we synthesized ten variations of the English utterance, 'This is my net' with different F_0 fall rates on *net*. The duration and amplitude of each syllable were kept constant across the stimuli. The F_0 peak always occurred at the onset of /ɛ/ in *net*, because we did not know the limit of the F_0 peak location in English which allows an accent to be perceived on the previous syllable. The F_0 contour of the utterances started at 121 Hz, linearly rose to 150 Hz at the onset of /ɛ/ in *net*, and ended at 102 Hz: The F_0 difference between the peak and the end was 48 Hz. The duration of the F_0 fall varied from 11 csec to 2 csec in 1-csec steps, corresponding to fall rates between 4.4 Hz and 24 Hz/csec. Figure 4 illustrates the F_0 contours of two sample stimuli (1 and 10).

Thirty-four native speakers of American English, monolingual adults, participated in the experiment. The instructions were given in synthetic speech to allow the subjects to familiarize themselves with the synthetic voice. They were then asked to judge whether each utterance was more appropriate for responding to 'What's this?' (inducing the accent on *net* in 'This is my net') or to 'Is this her net?' (inducing the accent on *my*): Hereafter, the former will be referred to as a *net-response*, and the latter as a *my-response*. Each subject listened to two sets of the ten stimuli, which were randomized in different orders.

Results

The results are summarized in Table 2. The second column indicates the fall rate; the third indicates the percentage of consistent judgments, i.e., a speaker making the same judgment in both sets, (the number of subjects appears in parentheses); the fourth

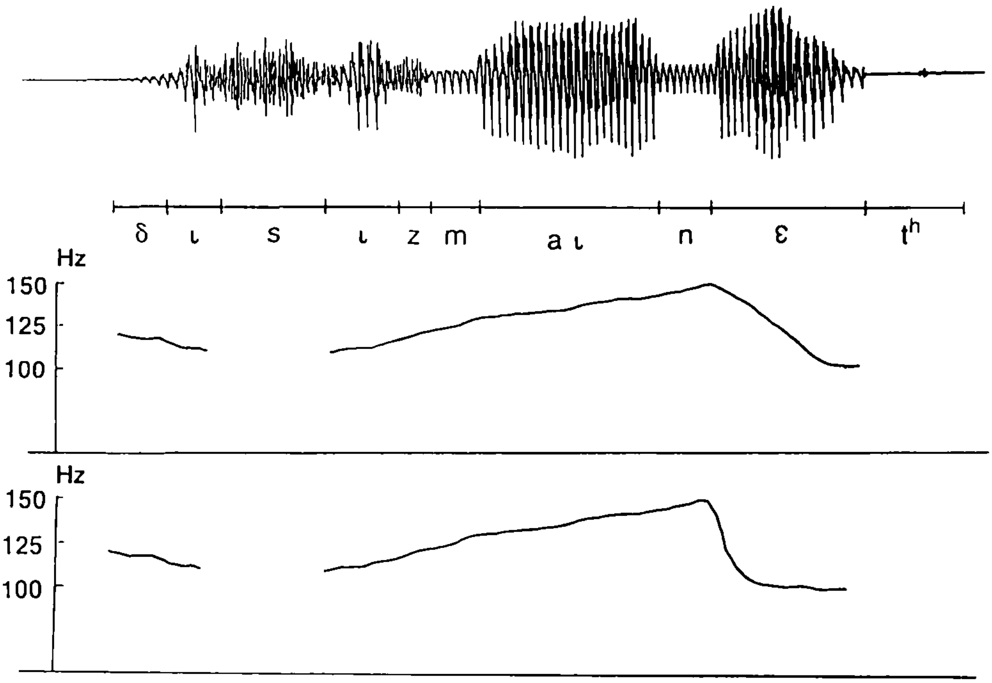

Fig. 4. Examples of F_0 contours for stimuli in Experiment 2. The top contour is with 4.4 Hz/csec fall rate and the bottom contour, with 24 Hz/csec fall rate.

and fifth columns indicate the percentage of these consistent judgments which were *net-* and *my-responses,* respectively. As shown in the third column, the subjects' judgments are most consistent at the two extreme fall rates: 91% when the rate is smallest (4.4 Hz/csec), and 79% when it is greatest (24 Hz/csec). The farther the fall rate is from these two extremes, the fewer the consistent judgments. This result implies that, given that the F_0 peak occurs at the onset of the vowel, if the F_0 fall rate is very gradual or very steep, the listener can determine the location of accent consistently from F_0 information, but if the rate is close to 8 Hz/csec, then F_0 information by itself is ambiguous as a cue to accent location.

The fourth and fifth columns of the table show that there is an asymmetry between the two judgments with greater and smaller fall rates. At 4.4 Hz/csec, 97% of the consistent judgments are *net-responses.* In contrast, at 24 Hz/csec, the subjects' judgments are split between *net-* and *my-responses.* If the rate is small, the accent is perceived on the syllable where the actual F_0 peak occurs, but if the rate is large, the accent is only somewhat more likely to be perceived on the preceding syllable.

The general tendency for the greater fall rate to shift the perceived accent in this experiment is in accord with the results of the first experiment with Japanese accent. However, the contribution of fall rate to the perception of accent location seems weaker in English than in Japanese. Given the right combinations of rate and peak location, the

TABLE 2

Percentages and raw numbers (in parentheses) of consistent responses in Experiment 2 (34 subjects). *Net-responses* and *my-responses* are for those subjects who gave consistent judgments

Stimulus	Rate (Hz/csec)	Consistent judgments	*net-responses*	*my-responses*
1	4.4	91 (31)	97 (30)	3 (1)
2	4.8	79 (27)	96 (26)	4 (1)
3	5.3	79 (27)	96 (26)	4 (1)
4	6.0	74 (25)	84 (21)	16 (4)
5	6.9	65 (22)	73 (16)	27 (6)
6	8.0	59 (19)	79 (15)	21 (4)
7	9.6	62 (21)	52 (11)	48 (10)
8	12.0	71 (24)	50 (12)	50 (12)
9	16.0	71 (24)	33 (8)	67 (16)
10	24.0	79 (27)	41 (11)	59 (16)

perceived accent shift occurred almost 100% of the time in Japanese, whereas even at the 0% peak location with a large fall rate, only 67% of the responses show a shift of accent to an earlier syllable in English. This implies that F_0 is not as prominent a cue to accent perception in English declarative intonation as it is in Japanese.

DISCUSSION AND CONCLUSIONS

Two experiments were conducted to investigate F_0 as a cue to accent placement in Japanese and English. Native listeners of both languages perceived an accent on a syllable which did not contain an F_0 peak. In Experiment 1 with Japanese, the later the F_0 peak occurs relative to the syllable boundary, the greater the fall rate necessary for the listener to perceive an accent on the preceding syllable. In Experiment 2 with English, where the F_0 peak location was fixed, the greater the F_0 fall rate, the more often the preceding syllable was perceived as accented.

However, with Japanese, given two successive syllables and a F_0 peak close to the beginning of the second vowel, almost all subjects judged the first syllable as accented even with the fall rate less than 5 Hz/csec, whereas with English, a much greater fall rate failed to achieve consistent *my-responses.* It should be pointed out that in

Experiment 2 the *net-response* is unmarked, whereas the *my-response* is highly marked. Many subjects commented after the experiment that some stimuli were not as natural as they should be for 'This is my nét' but that they nevertheless gave a *net-response* because the accent they perceived on *my* was not sufficient to carry a contrastive accent.

We conclude by considering implications of our findings for sound change and speech synthesis. Assuming that both F_0 peak location and F_0 fall rate are factors in the perception of accent, it can be argued that these factors may also play a significant role in sound change. Ohala (1981, 1983) argues that some sound change can result from errors during transmission of pronunciation from one speaker to another. In the case of Japanese accent, Hirayama (1968) and McCawley (1968, 1977) report that the accent system of two-syllable words in the Tokyo-type dialect is likely to have evolved from the proto-form by shifting accent one syllable to the right (see also Hata and Hasegawa, 1988b). One possible explanation is that if the F_0 peak occurred on the post-accent syllable, and the fall rate was not great enough, then the accent could have been perceived on the syllable where the actual peak occurs, that is, the syllable to the right of the one which the speaker intended to be accented.

An implication of our findings for speech synthesis is also noteworthy. We have sometimes encountered synthetic English utterances with competing cues to accent location, e.g., the synthetic utterance 'This is a test' The duration of the final syllable and the vowel quality of the penultimate syllable indicated that the nuclear accent is on *test*, but F_0 indicated it is on *a*. The cause is likely to be an excessive post-peak F_0 fall. Even when the F_0 peak is located within the intended syllable, if the fall rate exceeds a certain limit, the listener who relies heavily on the F_0 cues is likely to hear an accent on the preceding syllable.

In the literature, the proper location of the F_0 peak of the nuclear accent in English continues to be controversial. Ashby (1978) claims that the F_0 peak occurs at a fixed location in the vowel regardless of the vowel length in his production data, whereas Steele (1986) contends that the peak location should vary according to the vowel duration. Furthermore, Pierrehumbert (1981) and Silverman (1987) report that F_0 peak location varies between the nuclear accent and prenuclear accents. Their F_0 algorithms for synthesizing English intonation place the peak earlier in the nuclear accent than in prenuclear accents. Our experiment suggests that the proper F_0 peak location should be determined taking the post-peak F_0 fall rate into consideration.

In this paper, we have shown that there is a correlative effect between F_0 peak location and fall rate in Japanese. We also demonstrated that English exhibits the same tendency, although the role of F_0 contours in accent perception is not as salient as in Japanese. Experiments with different variables (e.g., duration, amplitude) remain to be conducted in the future.

REFERENCES

ASHBY, M. (1978). A study of two English nuclear tones. *Language and Speech*, **21**, 326–336.

BECKER, R.A., and CHAMBERS, J.M. (1984). *S: An Interactive Environment for Data Analysis and Graphics*. Belmont, CA: Wadsworth Advanced Book Program.

BECKMAN, M.E. (1986). *Stress and Non-stress Accent*. Dordrecht: Foris Publications.

BECKMAN, M.E., and PIERREHUMBERT, J.B. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, **3**, 255–309.

BOLINGER, D. (1954). English prosodic stress and Spanish sentence order. *Hispanica*, **37**, 152–156.

BOLINGER, D. (1961). Contrastive accent and contrastive stress. *Language*, **37**, 83–96.

BRUCE, G. (1977). *Swedish Word Accents in Sentence Perspective*. Travaux de l'Institute de Linguistique de Lund. Lund: CWK Gleerup.

CHAFE, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. Li (ed.), *Subject and Topic* (pp. 25–55). New York: Academic Press.

EADY, S.J., COOPER, W.E., KLOUDA, G.V., MUELLER, P.R., and LOTTS, D.W. (1986). Acoustical characteristics of sentential focus: Narrow *vs.* broad and single *vs.* dual focus environments. *Language and Speech*, **29**, 233–251.

FRY, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, **27**, 765–768.

FRY, D.B. (1958). Experiments in the perception of stress. *Language and Speech*, **1**, 126–152.

FUJISAKI, H., MORIKAWA, H., and SUGITO, M. (1976). Temporal organization of articulatory and phonatory controls in realization of Japanese word accent. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, **10**, 177–190. University of Tokyo.

FUJISAKI, H., and SUGITO, M. (1977). Onsei no butsuriteki seishitsu (Acoustical characteristics of speech sounds). In S. Ono and T. Shibata (eds.), *Iwanami Kooza: Nihongo 5, On'in* (pp. 63–106). Tokyo: Iwanami Shoten.

HALLIDAY, M.A.K. (1967). Notes on transitivity and theme in English, Part II. *Journal of Linguistics*, **3**, 199–244.

HASEGAWA, Y., and HATA, K. (1988). Delayed pitch fall in Japanese. *Journal of the Acoustical Society of America*, **83**, S29.

HASEGAWA, Y., and HATA, K. (1990). Perceptual shift of accent in English. *Journal of the Acoustical Society of America*, **88**, S127.

HATA, K., and HASEGAWA, Y. (1988a). Delayed pitch fall in Japanese: Perceptual experiment. *Journal of the Acoustical Society of America*, **84**, S156.

HATA, K., and HASEGAWA, Y. (1988b). Delayed pitch fall phenomenon in Japanese. *Proceedings of the Western Conference on Linguistics*, 87–100, Oct. 14–16, 1988, California State University, Fresno.

HIRAYAMA, T. (1968). *Nihon no hoogen* (Japanese dialects). Tokyo: Koodansha.

JAVKIN, H.R. (1976). Auditory basis of progressive tone spreading. *Journal of the Acoustical Society of America*, **60**, S45.

JAVKIN, H.R., HATA, K., MENDES, L., PEARSON, S., IKUTA, H., KAUN, A., DE HAAN, G., JACKSON, A., ZIMMERMANN, B., WISE, T., HENTON, C., GOW, M., MATSUI, K., HARA, N., KITANO, M., LIN, D.-H., and LIN, C.-H. (1989). A multi-lingual text-to-speech system. *Proceedings of IEEE International Conference on Acoustics, Speech & Signal Processing*, **1**, 242–245. May 23–26, 1989, Glasgow, Scotland.

KOHLER, K.J. (1990). Macro and micro F_0 in the synthesis of intonation. In J. Kingston and M.E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (pp. 115–138). Cambridge: Cambridge University Press.

LAMBRECHT, K. (1986). Topic, focus, and the grammar of spoken French. Ph.D. Dissertation, University of California, Berkeley.

LEHISTE, I., and PETERSON, G. (1961). Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America*, **33**, 419–425.

MADDIESON, I. (1976). Tone spreading and perception. *Journal of the Acoustical Society of America*, **60**, S45.

MAEDA, S. (1976). A characterization of American English intonation. Ph.D. Dissertation, MIT.

MCCAWLEY, J.D. (1968). *The Phonological Component of a Grammar of Japanese.* The Hague: Mouton.

MCCAWLEY, J.D. (1977). Accent in Japanese. *Southern California Occasional Papers in Linguistics*, **4**, 261–302.

MCCLEAN, M.D., and TIFFANY, W.R. (1973). The acoustic parameters of stress in relation to syllable position, speech loudness and rate. *Language and Speech*, **16**, 283–290.

MITSUYA, F., and SUGITO, M. (1978). A study of the accentual effect on segmental and moraic duration in Japanese. *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics*, **12**, 97–112. University of Tokyo.

MORTON, W.A., and JASSEM, W. (1965). Acoustic correlates of stress. *Language and Speech*, **8**, 159–181.

NAKATANI, L.H., and ASTON, C.H. (1978). Acoustic and linguistic factors in stress perception. Unpublished manuscript, Bell Laboratories.

NEUSTUPNÝ, J.V (1966). Is the Japanese accent a pitch accent? *Onsei-Gakkai Kaihoo*, 121. Reprinted in M. Tokugawa (ed.), *Akusento* (pp. 230–239). Tokyo: Yuuseidoo, 1980.

OHALA, J.J. (1981). The listener as a source of sound change. In C.S. Masek *et al.* (eds), *Papers from the Parasession on Language and Behavior* (pp. 178–203). Chicago: Chicago Linguistics Society.

OHALA, J.J. (1983). The direction of sound change. In A. Cohen and M.P.R. v.d. Broecke (eds.), *Abstracts of the 10th International Congress of Phonetic Science* (pp. 253–258). Dordrecht: Foris Publications.

OLIVE, J. (1974). Speech synthesis by rule. Speech Communication Seminar, Stockholm, Aug. 1–3, 1974.

ONISHI, M. (1942). Kokugo akusento-ron (On Japanese accent). In M. Togo (ed.), *Nihongo no Akusento* (Accent in Japanese) (pp. 15–26). Tokyo: Chuuoo-kooron.

OYAKAWA, T. (1971). On the duration of Japanese short vowels. Monthly Internal Memorandum, 1971/2, Linguistics Dept., University of California, Berkeley.

PIERREHUMBERT, J. (1981). Synthesizing intonation. *Journal of the Acoustical Society of America*, **70**, 985–995.

SILVERMAN, K.E.A. (1987). The structure and processing of fundamental frequency contours. Ph.D. Dissertation, University of Cambridge.

STEELE, S.A. (1986). Nuclear accent F_0 peak location: Effects of rate, vowel and number of following syllables. *Journal of the Acoustical Society of America*, **80**, S51.

SUGITO, M. (1972). Ososagari-koo: Dootai-sokutei ni yoru nihongo akusento no kenkyuu (Delayed pitch fall: An acoustic study). *Shoin Joshi Daigaku Ronshuu* 10. Reprinted in M. Tokugawa (ed.), *Akusento* (Accent) (pp. 201–229). Tokyo: Yuuseidoo, 1980.

WEITZMAN, R. (1969). Japanese accent: An analysis based on acoustic-phonetic data. Ph.D. Dissertation, University of Southern California.