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.