The Lexical Accent Effects on the Perception of Fundamental Frequency Peaks in Japanese

Takahito Shinya, University of Massachusetts Amherst (tshinya@linguist.umass.edu)

INTRODUCTION

Relative perceived prominence between two fundamental frequency (F0) peaks is known to be influenced by at least two factors:

The declination effect
The second F0 peak (P2) has to be substantially lower than the first peak (P1) for them to sound equal, because listeners compensate for the F0 declination (Pierrehumbert 1979).

The baseline adjustment effect
The extent to which P2 is lower than P1 becomes greater as P1 height increases. (Terken 1991, Gussenhoven et al. 1997). Gussenhoven et al. 1997 proposes that the baseline within which the F0 maxima and minima are scaled declines as a function of P1 height.

Lexical pitch accent in Japanese
Japanese is a language which has lexical pitch accent and its acoustic correlate of lexical pitch accent is F0:

- An accented word exhibits a higher F0 peak and a sharp F0 drop at the accented mora while an unaccented word shows a less high peak and show no F0 drop.
- An accented word affects the realization of the following word such that it is downstepped with respect to what precedes it.

QUESTION: What role does lexical pitch accent in Japanese play in the perception of intonational prominence?

METHODS

Stimuli
- Sentences consisted of three word: N1-Gen N2-Nom V.
- Accentedness of N1 and N2 were varied: *aa, au, ua* and *uu* (*a*: accented, *u*: unaccented).
- F0 height of N1(P1) and N2 (P2) were varied using the PSOLA algorithm (pitch synchronous overlap add) with Praat: P1 in 4 steps and P2 in 7 steps (step size = 1 semitone with 100 Hz base freq ≈ 10 Hz).
- Target tone is pitch accent H* for *a*, phrasal H for *u*.

Listeners: 19 native speakers of Japanese.

Task: Judging relative prominence of the two F0 peaks by saying which of them (N1 or N2) is given more "emphasis" by the speaker.

Analysis: Following Gussenhoven et al. (1997), probit analysis (Finney 1971) was used to estimate 50% crossover points, i.e. points where P1 and P2 have the same perceived prominence.

Stimulus sentences

a. *aa* *Ina'mori*-no *an'iyome*-ga *inai.*
   Inamori-Gen sister-in-law-Nom not found
   Inamori's sister-in-law is not found.

b. *au* *Ina'mori*-no *omiyage*-ga *kieta.*
   Inamori-Gen souvenir-Nom disappeared
   Inamori's souvenir disappeared.

c. *ua* *Inamura*-no *an'iyome*-ga *inai.*
   Inamura-Gen sister-in-law-Nom not found
   Inamura's sister-in-law is not found.

d. *uu* *Inamura*-no *omiyage*-ga *kieta.*
   Inamura-Gen souvenir-Nom disappeared
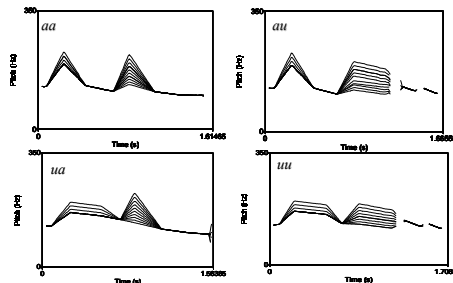   Inamura's souvenir disappeared.

Figure 1 F0 contours used in the experiment.

Predictions

- If listeners compensate for the extra F0 height in an accented word, P2 that has the same perceived prominence as P1 will be higher when it is accented than when it is unaccented.

- If listeners compensate for downstep, P2 will be perceived higher when P1 is accented than when it is unaccented.

- The declination effect will be observed in all conditions since it is assumed to always be seen when two F0 peaks occur one after the other.

- The baseline adjustment effect will be observed in all conditions since it depends only on the excursion size of P1.

References
Finney, D. J. (1971) *Probit Analysis*. New York: Cambridge Univ. Press.
Gussenhoven, C., Repp, B. H., Rietveld, A. C. M., Rump, H. H. and Terken, J. (1997) "The perceptual prominence of fundamental frequency peaks," *JASA* 102, 3009-3021.
Pierrehumbert, J. (1979) "The perception of fundamental frequency declination," *JASA* 66, 363-369.
Terken, J. (1991) "Fundamental frequency and perceived prominence of accented syllables," *JASA* 89, 1768-1776.

RESULTS

- In both graphs in Figure 2, all mean P2 values are below the $y = x$ function (except the lowest P1 in *ua*): the declination effect.

- In Figure 2a, *aa* shows higher mean values than *au* ($F_{(1,18)} = 68.57$, $p < 0.001$), indicating perceptual compensation for lexical pitch accent.

- In Figure 2b, *ua* shows higher mean values than *uu* ($F_{(1,18)} = 69.92$, $p < 0.001$), which confirms the lexical accent compensation.
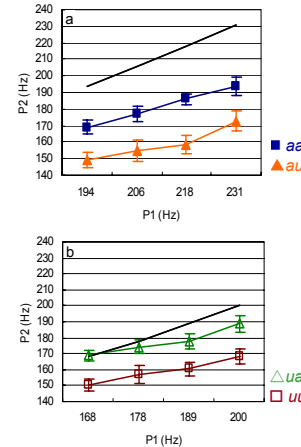

Figure 2 Mean values of P2 giving the same prominence as P1 as a function of P1 height for different accent-type conditions: (a) *aa* vs. *au*; (b) *ua* vs. *uu*. The function $y = x$ is shown for a reference. The error bars are 95 % confidence intervals.

When the accentedness of P1 and P2 differs, P2 has to be higher when P2 is accented than when it is unaccented in order for P2 to sound equal to P1 in perceived prominence. This suggests the presence of perceptual compensation for lexical pitch accent.

- In Figure 3, *aa* exhibits higher P1-P2 mean values than *uu*. A planned contrast supported the observation ($F_{(1,18)} = 4.411$, $p = 0.050$).

- Since the accent types of P1 and P2 do not differ in *aa* and *uu*, the obtained difference between these two conditions cannot be the lexical accent effect. It should be perceptual compensation for downstep.

- The P1-P2 values become greater as P1 increases: the baseline adjustment effect.
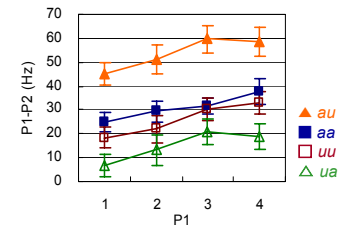

Figure 3 Mean P1-P2 difference values (= distance between mean values and the $y = x$ function) for different accent-type conditions (*aa, au, ua,* and *uu*). The error bars represent 95 % confidence intervals.

P2 is perceived more prominent when P1 is accented than when it is unaccented due to the perceptual compensation for downstep.

DISCUSSION

Accent discounting – The F0 peak of an accented word is perceptually discounted such that its perceived value is lower than its physical value.
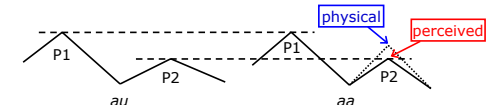
Downstep enhancement – The F0 peak that is preceded by an accented word is perceptually enhanced such that its perceived value is higher than its physical value.