THE INFLUENCE OF PHRASE BOUNDARIES ON PERCEIVED PROMINENCE IN TWO-PEAK INTONATION CONTOURS

Toni Rietveld & Carlos Gussenhoven

Department of Language and Speech
University of Nijmegen,
Postbus 9103, 6500 HD Nijmegen, The Netherlands
Tel. +31 24 361 2905, FAX +31 24 361 5939, E-mail: a.rietveld@let.kun.nl

ABSTRACT

Dutch listeners rated the perceived prominence of the second F_0 peak in a number of artificial two-peak intonation contours. When the two peaks were contained within a single intonation phrase, as signalled by the shape of the contour between the two peaks, and the pitch range of the contour was neutral, the perceived prominence of the second peak appeared to be determined by the F_0 of the first. However, when the two peaks occurred in separate intonation phrases, or when the pitch range was wide, no such effect was found. The results suggest that the prominence of peaks is judged as a function of the prominence of the contour of its intonation phrase, at least in contours with normal pitch range. These findings confirm the effect of pitch range found in the literature for English, and extend them by their demonstration that the effect is chiefly to be attributed to the phonological shape of the contour, and only secondarily to its pitch range.

1. INTRODUCTION

The perceived prominence of the second peak (P2) in two-peak F_0-contours in English is known to depend on the F_0 of the first peak (P1): the perceived P2-prominence is either negatively or positively correlated with the F_0 of P1 [1]. The positive correlation was found in contours with a P2 of 140 Hz, and a negative correlation with a P2 of 160 Hz. The former effect was also found for Dutch [2]. An explanation for these findings was suggested by [3]. Ladd argued that P2 prominence is judged as a function of the whole contour when P2 is below approximately 145 Hz, whereas P2 prominence is judged independently of the surrounding peaks when it is higher than 145 Hz ('local prominence'). The positive correlation is thus explained by the fact that P1 is perceived to be in the same contour as P2 when P2 has normal pitch range, while the negative correlation is explained by the possibility of throwing a high-range P2 into relief by lowering P1.

1.1 An alternative explanation

The explanation offered in [3] relies crucially on a presumed effect of pitch range on the perception of F_0 peaks. However, if the two opposite correlations are due to the coherence or non-coherence of P1 and P2, it is reasonable to suppose that the effects are really a function of the phrase structure of the two peaks, and that pitch range variation only indirectly causes the difference, in that wide pitch range might predispose the listener to perceive the peaks as occurring in separate intonation phrases, while average pitch range might rather suggest a single intonation phrase. An alternative explanation therefore is that the positive correlation between perceived P2-prominence and the F_0 of P1 is an indication that P1 and P2 are perceived as occurring in the same intonation phrase (or 'association domain'), while the negative correlation is an indication that the two peaks are perceived as belonging to different intonation phrases.

1.2 Testing for phrasing and pitch range

In order to test this alternative hypothesis, an experiment was carried out with Dutch listeners in which the variable pitch range was crossed with a phrasing variable. The latter variable consisted of two types of intonation pattern. One type represented a sequence of two intonation phrases, each containing a H*L pitch accent, while the other represents a single intonation phrase, with a prenuclear and a nuclear H*L pitch accent (see [4,5] for relevant intonational analyses). The factor pitch range was included in the experiment by pairing a high P1 (6 F_0-values) with a fixed P2 of 156 Hz in one half of the stimuli, and a low P1 (6 F_0-values) with a fixed P2 of 136 Hz). The hypothesis we intended to test was that in the utterances with a single intonat-ion phrase the perceived prominence of P2 is positively correlated with the F_0 of P1, regardless of pitch range, whereas in the utterances with two phrases this positive correlation is absent.

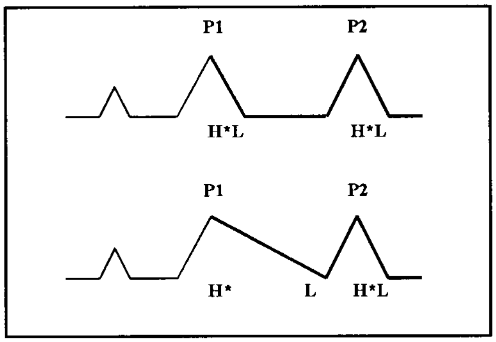

Figure 1. Utterances with two peaks in different intonation phrases (top) and contained in a single intonation phrase (bottom).

Pilot research with Dutch listeners suggested that utterances with two-phrase contours of the type that is schematically shown in the top contour in figure 1 are not consistently perceived as having two phrases, especially when listeners are repeatedly exposed to the same sentence. The pilot experiment concerned used a single carrier sentence, and did not show an effect for phrase structure. We conjectured that the reason for the absence of such an effect was the satiation caused by the lack of variation in the language materials that the subjects were exposed to, as a result of which their unguarded, automatic perception of the phrasing contrast might have been frustrated. Accordingly, in the present experiment we used five carrier sentences, for which, moreover, differences in phrasing lead to different interpretations. The carrier sentences were simple arithmetic expressions of the type '*three plus five times a hundred and three*', which can be interpreted as '(3 + 5) × 103' (i.e. 824) or as '3 + (5 × 103)' (i.e. 518). The prediction was that a phrase break after the P1 on the second numeral (i.e. '5' in the example) will induce the former reading, and will cause judges to rate the P2 prominence independently of the prominence of P1. By contrast, absence of a phrase break after '5' will induce the perception of the latter reading, and will cause judges to rate P2 prominence as a function of P1 prominence.

The experiment consisted of three parts. In part 1, subjects were given four examples of the intonation phrasing contrast, together with written version which included the associated bracketings. In part 2 they were sensitised to the relevance of the phrase structure by presenting them with a subset of 20 stimuli and asking for them to indicate which of two bracketings was the one suggested by the pronunciation in each stimulus. In part 3 the test stimuli were presented to the subjects, who were asked to rate the magnitude of the prominence of the second peak by means of the method of magnitude scaling.

2. THE EXPERIMENT

The carrier sentences from which the stimuli were derived were spoken by a male speaker of Dutch, who realised them in a quiet, somewhat monotonous way, without pronouncing intonation breaks. The stimuli were generated by using the PSOLA-technique as included in the speech package PRAAT developed by Paul Boersma of the University of Amsterdam. There were four factors in the experiment, which were varied as follows.

1. *The F_0-peak of P2:* 136 and 156 Hz.
2. *The F_0-peak of P1:* for a P2 of 136 Hz: 116, 126, 136, 146, 156, 166 Hz and for a P2 of 156 Hz: 136, 146, 156, 166, 176 and 186 Hz.
3. *Phrasing:* 1 or 2 intonation phrases.
4. *Sentence:* 5 sentences (arithmetic expressions).

Thus 2 x 6 x 2 x 5= 120 experimental stimuli were obtained, which were presented in two different random orders to two groups of 9 subjects.

The stimuli were recorded on digital audiotape with a sampling frequency of 16 kHz. Part 3 of each tape consisted of 120 experimental stimuli, preceded by five randomly chosen stimuli and followed by one randomly chosed stimulus whose scores were ignored. The Inter-Stimulus Interval was 4 sec. The 126 stimuli were presented in 12 blocks of ten and one block of six stimuli. Each block was preceded by a bleep, followed by a six second silent interval and an anchor stimulus consisting of a single numeral (*honderdzes* '106') which contained one accent realised by means of an F_0 peak of 146 Hz on the morpheme *zes*. The prominence of the accented syllable concerned was said to correspond to a prominence level exactly half-way on the scale. The subjects were asked to rate the magnitude of the last accent of each stimulus by placing a mark on a 100 mm scale without any calibration or preset categories, a procedure similar to that known as 'line production' [6].

3. RESULTS

Scores were registered by measuring the distance in mm between the beginning of each 100 mm line and the subject's mark on each scale. The scores were pooled over the five carrier sentences, which resulted in 24 data points per subject. An analysis of variance was carried on these data. Three within-subject factors were included in the model: PHRASING (one phrase, or 'linked' vs two phrases, or 'unlinked'), HEIGHT-OF-PEAK-2 (high: 156 Hz and low: 136 Hz) and HEIGHT-OF-PEAK-1 (6 levels). As predicted there was a three-way interaction between PHRASING, HEIGHT-OF-PEAK-2 and HEIGHT-OF-PEAK-1: $F(5,85)=3.83$, $p=0.004$, Huynh-Feldt corrected. No other interactions were significant. Significant main effects were HEIGHT-OF-PEAK-2: $F(1,17)=40.86$ and HEIGHT OF PEAK 1: $F(5,85)=4.13$, $p=0.00$, Huynh-Feldt corrected. The mean scores for the low-range stimuli were 50.7 (linked) and 50.1 (unlinked), and for the wide-range stimuli 62.2 (linked) and 62.5 (unlinked).

In figure 2 we display the mean ratings of the subjects as a function of the three factors involved. Clearly, the three-way interaction is mainly due to the strong correlation that exists between the F_0 of P1 and the perceived prominence of P2 in the condition of a linked contour (one intonation phrase) and a low-range P1 and P2. The perceived prominence of P2 increases almost monotonously with the height of F_0 of P1. There are no such correlations in the other three conditions: linked contour with a

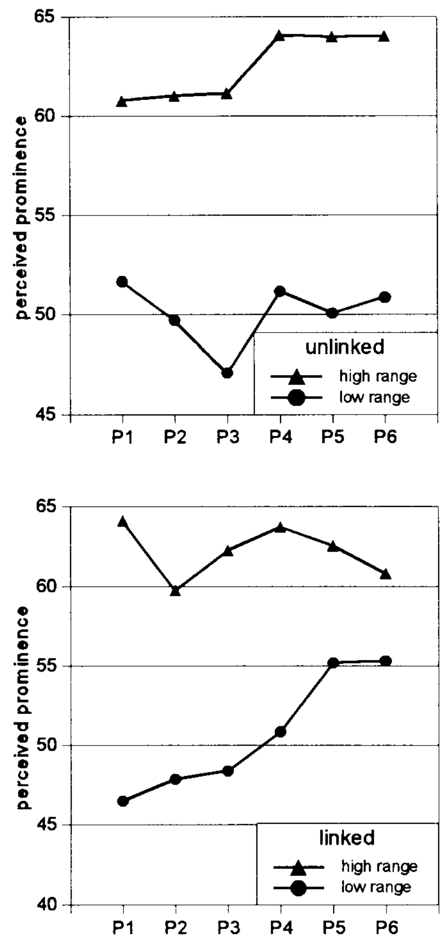

Figure 2. Mean ratings of Peak 2 as a function of phrasing, the F_0 of Peak 2 (136 Hz and 156 Hz), and the F_0 of Peak 1 (6 levels). Each mean score is averaged over 5 sentences $\times 18$ subjects $= 90$ scores.

wide-range P2, and unlinked contours with low-range and wide-range P2.

4. CONCLUSION

The postive correlation between the F_0 of P1 and the perceived prominence of P2 in two-peak intonation contours must be attributed to the way the two-peak contour is perceived: if it is perceived as a single intonation phrase, the positive correlation will occur. If it is not, the positive correlation will be absent. The perception of an intonation phrase break between the two peaks is most effectively induced by an intonation contour shape that signals such a break ('unlinked pitch accents'), while its absence is most readily perceived when the contour has a shape that, according to current theories of intonational form, is associated with a single intonation phrase ('linked pitch accents'). The conclusion that the phonological shape of the contour is a strong predictor of the occurrence of the positive correlation is supported by the results of our experiment. After all, in the low-range condition, the effect was only found when the contour had the appropriate single-phrase shape. However, in addition to the phonological shape of the contour, the pitch range would appear to have an effect, too. In the higher pitch range, the 'linking' of the two pitch accents did not have the same effect that it had in the lower-range stimuli. This confirms results obtained by [1], who found that the low-range contours, but not wide-range contours, produced the positive correlation between the pitch of P1 and the prominence of P2. We hypothesise that the reason for the effect of range is that the wide-range stimuli exceeded the kind of peak height that is normally found in linked contours. This would imply that for two pitch accents to be included in the same phrase, their prominence level must fall within a `neutral' range.

Finally, in no condition did we find the opposite effect, that of a negative correlation between the F0 of P1 and the prominence of P2. This `contrast' effect has been found for English in contours with high prominence [1].

5. REFERENCES

[1] D.R. Ladd, J. Verhoeven and K. Jacobs, "Influence of adajcent pitch accents on each other's perceived prominence: two contradictory effects", *Journal of Phonetics,* Vol. 22, pp. 87-99, 1994.

[2] C. Gussenhoven and T. Rietveld, "Fundamental frequency declination in pitch: testing three hypotheses". *Journal of Phonetics*, Vol. 16, pp.355-369, 1988.

[3] D.R. Ladd, "Constraints on the gradient variability of pitch range, or, Pitch level 4 lives!", in P. Keating (ed.) *Papers in Laboratory Phonology III,* Cambridge UP, Cambridge, pp. 43-63. 1994.

[4] C. Gussenhoven, "Adequacy in intonation analysis: The case of Dutch", in H. Van der Hulst and N. Smith (eds.) Advances in Nonlinear Phonology. Foris, Dordrecht, pp. 95-121, 1988.

[5] M.E. Beckman and J. Pierrehumbert. "Intonational structure in English and Japanese", *Phonological Yearbook*, Vol. 3, pp.255-310, 1986.

[6] M. Lodge, "Magnitude scaling". London: Sage Publications.

Acknowledgements

We are indebted to Rob van den Berg, who helped us with our experiment in many ways. We thank Bob Ladd for useful discussion of the issues that the experiment was concerned with.