The Role of Pitch in Lending Prominence to Syllables

Dik J. Hermes and H.H. Rump
Institute for Perception Research, P.O. Box 513
NL-5600 MB Eindhoven, The Netherlands

ABSTRACT

Pitch is one of the prosodic variables that play an important role in lending prominence to an accented syllable. Little is known, however, about which characteristics of a pitch contour are important in this respect. A model will be described in which it is assumed that the contribution of pitch to prominence depends linearly on the difference between the average pitch level in the nucleus of the accented syllable and the average pitch level in the nucleus of the preceding syllable. When excursion size is defined as the difference between upper and lower declination line, a consequence of this model is that, excursion sizes being equal, a fall lends more prominence to a syllable than a rise. An experiment will be described, in which this difference in prominence between a rise and a fall is measured for various declination slopes. The results are in good agreement with the predictions of the model.

INTRODUCTION

It is well known that pitch plays an important role in lending prominence to accented syllables. In this study we investigated which properties of the pitch contour contribute to the prominence of an accented syllable. Hermes and Van Gestel (1991) showed that subjects are capable of comparing the prominence of accented syllables in two utterances presented in different registers, when the prominence was lent by the same accent-lending pitch movement, a rise, a rise-fall or a fall. In another study, Hermes (1991) showed that subjects are also capable of comparing the prominence of the accented syllables in two utterances when prominence was lent by different types of pitch movements. It was found that, when lending equal prominence, the rise and the rise-fall had a larger excursion size than the fall. It could be excluded that this discrepancy between the rise and the rise-fall on the one hand and the fall on the other was due to a difference in timing. A model was proposed, in which the contribution of pitch to prominence depends linearly on the difference between the average pitch in the syllabic nucleus of the accented syllable and the average pitch in the syllabic nucleus of the preceding syllable. As shown in Fig. 1, this difference is larger for the fall than for the rise, when the excursion sizes are equal. Excursion size is defined as the distance between the upper and the lower declination line. To be more precise, a calculation shows that this difference S equals

$$S = 2dT, \tag{1}$$

in which d is the declination slope, and T is the time interval between the two points subjects use in estimating the pitch-level differences, the pitch-level estimation points. Note that S is independent of the excursion size of the pitch movements. The model predicts that this difference will increase linearly with increasing declination. In this contribution an experiment will be described to test this prediction.

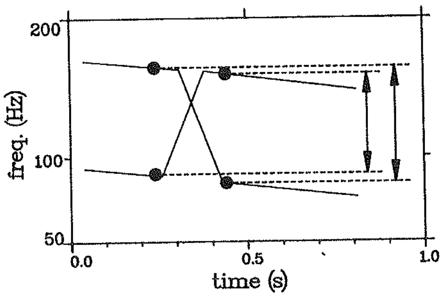

Figure 1: *Diagram showing that for equal excursion sizes the distance between pitch levels in a fall is greater than in a rise. The filled circles indicate pitch levels which might be used for the perceptual determination of prominence. Note that the difference in length between the arrows is due to the presence of declination.*

METHODS

The experiment was an adjustment experiment. Its design is shown in Fig. 2. It shows the pitch contours of two PSOLA-resynthesized utterances, /mamáma/, a Dutch nonsense word, with an accent on the second syllable. The utterances are identical, except that the second syllable in the first stimulus is accented by a fall, while the second syllable in the second stimulus is accented by a rise. The excursion size of the pitch movement in the first stimulus, the test stimulus, was fixed. The excursion size of the pitch movement of the second stimulus, the comparison stimulus, was under the subject's control. Subjects were asked to adjust the excursion size of a pitch movement in a comparison stimulus until the prominence of the accented syllable was equal to the prominence of the accented syllable in the test stimulus. When the subject had done this, the procedure was repeated for a test stimulus with another excursion size. This was done for a total of six different excursion sizes in the test stimulus, shown in the left part of the figure. When this was completed, the experiment was repeated with the rise in the test stimulus, and the fall in the comparison stimulus. In this way an accurate estimate could be obtained of the difference in excursion size between a rise and a fall, when lending equal prominence (for details see Hermes, 1991). This excursion size is expressed in E, the number of ERBs, since this psycho-acoustically defined frequency scale (Patterson, 1976) appears to be most adequate for expressing the frequency of speech intonation (Hermes and Van Gestel, 1991).

The experiment was carried out for five different declination slopes: 0.0, 0.33, 0.70, 1.09 and 1.50 E/s, which, in this register, correspond to slopes of 0 and about 2.4, 4.9, 7.3 and 9.7 semitones/s, respectively.

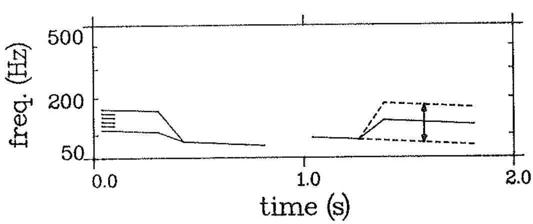

Figure 2: *Set-up of an adjustment run. The subject was first presented with one of the six test stimuli displayed on the left-hand side. Then the subject heard the comparison stimulus with an adjustable excursion size shown on the right-hand side. The run stopped when the subject indicated that the same prominence was perceived in the accented syllables of both stimuli.*

RESULTS

Twelve subjects took part in this experiment. They were students and research associates of the Institute for Perception Research. All of them reported normal hearing and most of them had experience in auditory perception experiments. In order to determine whether subjects performed consistently, the criterion for consistency described in Hermes and Van Gestel (1991) was applied. As a consequence, the results of three of the twelve

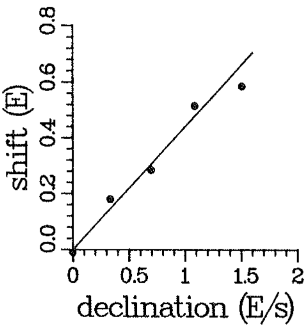

Figure 3: *The data points show the average difference S in the excursion size of the rise and the fall for the five different declination slopes. The line represents Eq. 1, the predicted relation between the difference S and declination.*

subjects were excluded from the analysis. The results of the nine consistent subjects are presented in Fig. 3. The data points show the difference in excursion size between the rise and the fall as found in the experiment. The straight line shows the predictions for S as obtained from Eq. 1. This figure shows that the predictions are quite good.

The value chosen for T in Eq. 1 was 0.22 s, the interval between the vowel onsets of the first and the second syllables. The vowel onsets were taken, since they are quite accurately defined in time, much more accurately than the nucleus which actually comprises a time interval.

DISCUSSION

We tested the effect of declination on the difference in prominence lent to a syllable by a rise and by a fall. More specifically, we tested the assumption that this prominence depended linearly on a difference in pitch level at some point before the pitch movement and at some point after the pitch movement. For the interval between these pitch-level estimation points we chose the interval between the vowel onsets, which in this utterance was equal to the distance between the syllabic nuclei.

It follows from the model that S, the difference in excursion size of the rise and the fall when lending equal prominence, is independent of the excursion size of the rise and the fall. This has a strange consequence when the data points are extrapolated to excursion size zero: a fall of excursion size zero will be adjusted to a rise with a positive excursion size. Although this was not tested experimentally, this agrees well with informal listening to stimuli of this kind. In the presence of no other pitch movement than is due to declination, a falling pitch accent is perceived on the second syllable. A similar observation can be made when only inclination is present, in which case a rising pitch accent is perceived. A prerequisite is probably that the accent on the second syllable is clear due to other cues such as loudness and vowel length. In the utterance used in this experiment, the speaker had pronounced the utterance in such a way that both loudness and vowel length in the accented syllable were larger than in the preceding syllable.

Another related consequence of this model is that, on the short time scale of these stimuli, there is no compensation for declination. One might have expected that listeners would be able to predict the discrepancy between the rise and the fall on the basis of the declination preceding and following the pitch movement, and would compensate for this. The results of the experiment described here show that listeners do not compensate.

It is not yet clear how long the time interval should be between the two pitch-level estimation points which the listeners use in comparing the difference in pitch level. In the model presented here it is implicitly assumed that the pitch-level estimation points are in the syllabic nuclei of the accented and the preceding syllables. There may, however, very well be a fixed distance which accidentally corresponds quite well with the distance between the nuclei of succeeding syllables in normal speech. In the stimuli used in these experiments the time interval between succeeding nuclei was 0.22 s. In fluent speech it has the same order of magnitude. Future experiments will have to clarify whether the distance between the pitch-level estimation points is fixed, or whether it varies with the actually realised distance between the syllabic nuclei in fluent speech.

ACKNOWLEDGEMENTS

We would like to thank Jacques Terken and Jan Roelof de Pijper for their constructive comments on the manuscript.

REFERENCES

Hermes, D.J., and Van Gestel, J.C. (1991) "The frequency scale of speech intonation", *Journal of the Acoustical Society of America,* Vol. 90, pp. 97-102.

Hermes D.J. (1991) "Prominence caused by rising and falling pitch movements with different positions in the syllable", *IPO Annual Progress Report,* Vol. 26, pp. 17-28.

Patterson R.D. (1976) "Auditory filter shapes derived with noise stimuli", *Journal of the Acoustical Society of America,* Vol. 59, pp. 640-654.