Perceptual cues to some Swedish prosodic phrase patterns - a peak shift experiment

Gårding, E. and Eriksson, L.

journal: STL-QPSR
volume: 30
number: 1
year: 1989
pages: 013-016

http://www.speech.kth.se/qpsr
PERCEPTUAL CUES TO SOME SWEDISH PROSODIC PHRASE PATTERNS - A PEAK SHIFT EXPERIMENT

Eva Gårding and Lars Eriksson
Dept. of Linguistics and Phonetics, Lund University

Abstract

Two double-peaked prosodic phrase patterns, (a) with two even accents and (b) with accented followed by deaccented, were used in an experiment in which the second peak was shifted in steps of 20 msec. In this way two series of stimuli could be generated, which, apart from (a) and (b), unexpectedly also included a third category (c), a compound phrase. The stimuli were tested, giving listeners the choice between (a),(b) and (c). Our results indicate that the pitch movements over the vowels are powerful cues for identification and also that the spectral and temporal characteristics are important. The role of the peak location is that it determines the extent to which a segment is touched by a rise, fall or combination of these movements.

INTRODUCTION

The method of shifting the time location of a pitch peak (the Fo peak of a constant triangular shape) in synthetic speech to study the effect on listeners is not unusual. Typically the aim has been to determine boundaries between distinctive prosodic categories, e.g. Accent 1 and Accent 2 in a southern Swedish dialect (Malmberg, 1955), Serbo-Croatian accents (Purcell, 1976), sentence accent in American English (Gårding & Gerstman. 1960), accents with different pragmatic values (Kohler, 1987).

We shall use the same method. Our main concern is the perceptual aspects of such a shift. More precisely we ask the question, for a certain phrase contour, what are the perceptually relevant combinations of pitch movement and segment?

MATERIAL AND METHOD

Some segmentally equivalent but prosodically and semantically different sentences were elicited in declarative intonation from a trained phonetician representing a modified Stockholm dialect (Table I). The recorded sentences were analysed digitally and manipulated by means of the ILS program. As a first step two prototypes were chosen from the several productions of each sentence (Table I) and their Fo patterns were simplified by using straight interpolation between conspicuous turning points.

The accents which are irrelevant to our study are marked ‘ ‘ for A1 and ‘ ‘ for A2 , , for deaccented, + for juncture.

The examples are borrowed from Bruce. For other speakers see Bruce 1977.

Table I. The prototypes.

<table>
<thead>
<tr>
<th>Written form</th>
<th>Syntax</th>
<th>Meaning</th>
<th>Phonetic form</th>
<th>Accent pattern</th>
<th>FO contour</th>
</tr>
</thead>
<tbody>
<tr>
<td>a = lång män</td>
<td>2-word phrase</td>
<td>tall men</td>
<td><code>lɔːga + </code>men</td>
<td>accented, unacc,</td>
<td>~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~</td>
</tr>
<tr>
<td>b = lång män</td>
<td>2-word phrase</td>
<td>tall men</td>
<td><code>lɔːga + </code>men</td>
<td>accented, unacc,</td>
<td>~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~</td>
</tr>
<tr>
<td>c = långamän</td>
<td>compound</td>
<td>men from lång</td>
<td><code>lɔːga</code>men</td>
<td>accented, unacc,</td>
<td>~ ~ ~ ~ ~ ~ ~ ~ ~ ~ ~</td>
</tr>
</tbody>
</table>

The examples are borrowed from Bruce. For other speakers see Bruce 1977.
The peak-shift experiment was conducted with two of these simplified contours, superposed on two spectral carriers, one derived from prototype (a) ending in accented /men:/, the other from prototype (b) ending in deaccented /men:/.

For the experiment a new computer program was designed. As can be seen in Fig. 1, the first peak over the syllable /lɔːr/ has been kept fixed, the second peak has been shifted in steps of 20 milliseconds towards the end of the sentence. In this way we could obtain two series of stimuli which apart from the intended categories (a) and (b) also turned out to include a third, a compound phrase (c).

In the subsequent tests there were altogether 3·40 = 120 stimuli which were judged by two groups of sixteen listeners each from Lund and Stockholm. The listeners, members of the linguistics departments, heard the stimuli over loud-speakers in the perception lab of their respective department. They were asked to place each item heard in one of the three categories (a), (b) or (c), as shown in Table I.

RESULTS AND DISCUSSION

With the carrier derived from prototype (a) (Fig. 1a) there is a sharp cross-over region between stimuli 15 and 18, i.e. at the end of /m/. Before the cross-over the compound is given about 15% of the votes and after this point about 70%. The response functions for the two-word phrases follow each other rather closely with some but much less than expected dominance for the accented /men:/ responses. The carrier from the deaccented /men:/ phrase, prototype (b), (Fig. 1b) has a similarly located cross-over with regard to segments (at the end of /m/). Here the votes for the deaccented pattern are predominant over the accented one. The compound phrase is well accepted after the cross-over also in this carrier.

![Fig. 1. Stimuli with shifted peak and response functions. a) Carrier with accented /men:/ b) Carrier with deaccented /men:/](image-url)
The difference in scores, obtained with similarly contoured stimuli in the two carriers, notably the compound, suggests that intensity, temporal and spectral characteristics play a not insignificant role in the identification of a prosodic phrase pattern.

The contours with the highest scores for both carriers are displayed in Fig. 2. They show that the three prosodic patterns can be given carrier-independent descriptions. Note, however, that the two-word contour with accented /men/ with carrier (b) which only received 22% of the votes has been disregarded.

![Fig. 2. Contours with highest scores. a) In carrier with accented /men/. b) In carrier with deaccented /men/.

Carrier-independent description:

<table>
<thead>
<tr>
<th>Two-word phrases</th>
<th>over /a/</th>
<th>over /e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>accented /men/</td>
<td>rise</td>
<td>fall</td>
</tr>
<tr>
<td>deaccented /men/</td>
<td>fall</td>
<td>low</td>
</tr>
<tr>
<td>Compound</td>
<td>low</td>
<td>fall</td>
</tr>
</tbody>
</table>

The compound seems to have a special status among listeners as compared to the two-word phrases which are not easily kept apart. (A group of fluent Swedish speaking non-native listeners have difficulties with the compound, however.) Reactions from the listeners make us speculate that the compound is separated from the other two by a simple binary choice between compound vs. non-compound, here materialized as low /a/ vs. non-low /a/. This choice precedes whatever choice has to separate the other two.

The common carrier-independent description supports the hypothesis that the pitch movements over the vowels are important perceptual cues to a prosodic pattern. (In analyses of Chinese and Swedish the tone (accent) carrying part of the syllable is the sonorant segments, the rhyme (Howie, 1976; House, Bruce, Eriksson, & Lacerda, 1988). The hypothesis may explain the low percentage of the accented pattern in the (b) carrier. Due to the narrow base of the pitch peak no stimulus in this series has the prototypical movements over the vowels.

CONCLUSION

Our results lead to the tentative conclusion that for a prosodic phrase pattern to be easily recognized it is important that the spectral, intensity and temporal pattern fit the prototype and that Fo has the prototypical movements over neighbouring vowels. A mechanical shift of the pitch peak reveals that it is when the pitch peak has produced such
patterns that the identification with a given pattern is the best. An Fo peak per se is perceptually unimportant. It is the adjoining ramps over vowels which have perceptual reality.

References


