HOW MANY RISE-FALL-RISE CONTOURS?

Janet B. Pierrehumbert and Shirley A. Steele

AT&T Bell Laboratories,
600 Mountain Ave, Murray Hill, NJ 07974
U.S.A.

ABSTRACT

This paper reports an experimental study of the rise-fall-rise intonation patterns of English. Variants of the rise-fall-rise contour differ from each other in the alignment of the F0 valley and peak with the text. The phonological analysis of this difference is controversial and is part of a wider debate on the function of category and continuum in intonation. Variants of the rise-fall-rise pattern have been viewed as either (1) falling into two categories, each with a distinct pragmatic meaning, or (2) occurring along a continuous dimension of peak delay and pragmatic meaning. This issue was addressed by examining how speakers imitate stimuli varying continuously in the alignment of the F0 rise-fall.

In the stimuli for the experiment, the alignment of the F0 rise-fall was varied in small steps by using LPC coding and resynthesis. Subjects heard the stimuli in randomized order and imitated what they heard. Peak delays in the responses were found to cluster in two groups, thus differing systematically from the peak delays in the stimuli. This result is readily explained by a model with two categories.

1. INTRODUCTION

1.1 Topic

In this paper, we report an experimental investigation of the rise-fall-rise intonation patterns of English. Variants of the rise-fall-rise contour differ from each other in the alignment of the F0 valley and peak with the text. Figure 1 exemplifies the F0 contour of this pattern with an early peak and Figure 2 shows a longer peak delay. (In both Figures a vertical line shows the location of the [m]-release.) The pattern in Figure 2 conveys speaker inbreath or uncertainty (see Ward and Hirschberg [7], [8]), while that shown in Figure 1 typically marks contrast or correction.

The phonological analysis of the rise-fall-rise intonation pattern is controversial and is part of a wider debate on the function of category and continuum in intonation. Peak delay variants of the rise-fall-rise pattern may be viewed as either (1) falling into some number of categories, each with a distinct pragmatic meaning, or (2) occurring along a continuous dimension of peak delay and pragmatic meaning. A formulation of the first view is found in Pierrehumbert [5] and a statement of the second in Gussenhoven [1].

According to Pierrehumbert, there are two different rise-fall-rise patterns, differing in how the low-high pitch accent is aligned with respect to the stressed syllable. In one, the high (H) is aligned with the stressed syllable (Figure 1). In the other, the low (L) falls on the stressed syllable (Figure 2). Using a diacritic "*" to represent alignment with the stress, the two patterns are transcribed L*+H and L*−H and are called "bitalon accents" because two tones are required to describe the accent. In both patterns, the fall-rise after the accent is explained by an L H sequence which marks the end of the phrase.

In Pierrehumbert's theory, the "*" diacritic for tones is analogous to stress for syllables. In metrical stress theory, as laid out by Liberman and Prince [4], the stress pattern of a word like "Peter" has a relatively strong syllable followed by a relatively weak syllable. In "repeat" the strength relation is reversed. Pierrehumbert's bitalon accents are treated in the same way, with a strength relation among the tones, as illustrated below:

\[ \begin{align*}
  &S \quad W \quad W \quad S \\
  &L^* \quad H \quad L \quad H^* 
\end{align*} \]

The single tone accents have a status corresponding to that of stressed monosyllables, such as "Peter".

This approach contributes to a broader picture in which tones participate in a hierarchical organization, which controls their alignment and phonetic realization. This broad picture is first proposed in Liberman [2] and is amplified and revised in Pierrehumbert and Beckman [6].

Gussenhoven asserts that English does not have a binary distinction, but a continuous dimension of peak delay. He suggests ([1], p. 218) that while there may be a preferred ("ideal") position along this continuum, a wide range is possible. His treatment of peak delay is thus analogous to the treatment of continuously variable overall pitch range in Liberman and Pierrehumbert [3]. Peak delay in pitch accents is treated as a paradigmatic feature, while word stress is viewed as a matter of metrical organization.

Our aim in the present work was to empirically investigate the category/continuum question in the rise-fall-rise intonation contour.

1.2 Method

In our experiment, subjects heard and imitated randomized rise-fall-rise contours constructed along a continuum of peak delay. If subjects perceive a continuum, the response peak delays should be approximately continuous. (A preferred peak position might cause responses to stimuli at the extremes to drift towards the center). If subjects hear categories, responses should cluster into discrete groups.

This experimental method is a variant of the paradigm familiar from studies of categorical perception of speech segments. Our study is, to our knowledge, the first application of such methods to the study of syntagmatic features. We have used an imitation task rather than the more commonly used labelling and discrimination tasks, because we were not concerned with separate analysis of production and perception systems. In most categorical perception studies, the linguistic analysis was relatively uncontroversial (e.g., English /b/ versus /f/); at issue was the status of linguistic description in the psychological system. In our study, we looked for evidence about the system of linguistic analysis; the relationship of any categories to perceptual or articulatory systems is a matter for later research.
2. EXPERIMENTAL PROCEDURES

2.1 Stimulation
The stimuli were versions of the phrase, "only a millionaire," in which the last syllable of the word "millionaire" was replaced by one of the four vowels in the order of the time of its production. The stimuli were presented at a comfortable intensity level to the right ear of the subject. In addition, the subjects' stimulus intensity. The subject was asked to judge whether the sound was perceived as a syllable or as a sequence of vowels.

2.2 Randomization
The 15 versions of the prompt were randomly assigned to blocks; then, in each of two sessions, 15 different randomized blocks were presented to the subject, for a total of 225 trials per session or 450 trials in all. Each of the 15 triads was reproduced 30 times.

2.3 Subjects
The subjects were five native speakers of American English, two females and three males. Four of the five were native speakers of the language, and one was a non-native speaker. The subjects were divided into four groups: a) English language, b) English phonetics, c) non-native English speakers, and d) non-native English speakers.

2.4 Measurement
The measurement of primary interest is peak delay, defined as the difference between the time of the FO peak and the time of the syllable. This was measured by the examiner. The measurements were made by hand. The examiner was blind to the experimental conditions. The measurements were made by hand.

3. RESULTS

3.1 Predictions about peak delays
In order to make the discussion of the data more concrete, let us first present some meaningful results for different values of the integration constant, the integration constant, the integration constant, and the integration constant.

3.2 Randomization
The 15 versions of the prompt were randomly assigned to blocks; then, in each of two sessions, 15 different randomized blocks were presented to the subject, for a total of 225 trials per session or 450 trials in all. Each of the 15 triads was reproduced 30 times.

3.3 Subjects
The subjects were five native speakers of American English, two females and three males. Four of the five were native speakers of the language, and one was a non-native speaker. The subjects were divided into four groups: a) English language, b) English phonetics, c) non-native English speakers, and d) non-native English speakers.

3.4 Measurement
The measurement of primary interest is peak delay, defined as the difference between the time of the FO peak and the time of the syllable. This was measured by the examiner. The measurements were made by hand. The examiner was blind to the experimental conditions. The measurements were made by hand.

4. CONCLUSION
This study investigated the role of peak delay in the perception of syllables in the American English language. The results showed that peak delay plays a significant role in the perception of syllables. The peak delay is a critical factor in determining the perception of syllables in the American English language. The findings suggest that peak delay can be used as a tool for improving the perception of syllables in the American English language.
3a) Peak delays for all responses, subject TWB

3b) Peak delays for all responses, subject HDT

3c) Peak delays for all responses, subject RLB

Median peak delay: subject TWB

Median peak delay: subject HDT

Median peak delay: subject RLB