A CLUSTER-SEEKING TECHNIQUE FOR PROSODIC ANALYSIS
(with special reference to Russian sentence intonation)

Leonid A. Kanter, Alexander P. Chizhov, Ksenia G. Guskova

Department of Phonetics, Faculty of Foreign Languages
(English Language Division), Lenin Pedagogical Institute
Moscow, USSR 107140

ABSTRACT

A cluster analysis algorithm proposed by Sammon is used to identify intonational zones which can be correlated with intonemes of Standard Russian.

INTRODUCTION

The need for cluster analysis arises in a natural way in many areas of phonetic research. The goal of clustering methods is to provide a means to discover structure within a complex body of data /4/. With regard to intonology the first use of a cluster-seeking technique was reported in /2/.

This paper attempts to analyse the manner in which intonation contours of five Standard Russian intonation types are located in the space of acoustic parameters.

MATERIAL

The material analysed consists of the test phrase OH SHAL [on znal] = "He knew", pronounced in dialogical contexts by sixteen male native speakers of Standard Russian. The speakers were instructed to read the phrase with context appropriate vocal modifications so that they could be identified as belonging to the following five intonation types, or communicative modes: (1) a final statement, (2) a reply statement, (3) a general question, (4) an exclamation, (5) a non-final statement.

The test phrase was read twice in each mode, whereupon 160 utterances were produced. Used as test stimuli, the utterances were then listened to and categorized by a group of subjects in terms of the set of intonation types under consideration.

The subsequent instrumental (intonographic) analysis was performed to measure fundamental frequency (F0), intensity and duration in 80 utterances selected as a result of the foregoing listening tests.

Fourteen initial parameters of each intonation contour were analysed:

(1) maximum F0 value within the first syllable;
(2) minimum F0 value within the first syllable;
(3) maximum F0 value within the second syllable;
(4) minimum F0 value within the second syllable;
(5) F at the starting point of the first syllable (F at the starting point of an utterance);
(6) F at the end point of the first syllable;
(7) F at the starting point of the second syllable;
(8) F at the end point of the second syllable (F at the end point of an utterance);
(9) F at the last turning point of an utterance;
(10) maximum F value between the starting point and the last turning point inclusive;
(11) maximum value of intensity within the first syllable;
(12) maximum value of intensity within the second syllable;
(13) duration of the first syllable;
(14) duration of the second syllable.

METHOD

To reduce variance between speakers the available acoustic parameters were subject to the following normalization procedures. The fundamental frequency parameters were normalized by the formula:

$$\gamma_{1(j)} = \frac{100 \cdot \gamma_{1(j)}}{\gamma_{max}}$$
where $y^{(2)}_j$ and $y^{(1)}_j$ are normalized (relative) and non-normalized (absolute) values of the $j$-th parameter in the $j$-th utterance respectively; $x_j$ is the maximum $x$ value in the $j$-th utterance.

The intensity and duration parameters were normalized using the formula:

$$x^{(2)}_j = \frac{y^{(2)}_j}{x_j^*},$$

where $E_{100}$ for intensity parameters and $E_{150}$ for duration parameters; $y^{(2)}_j$ and $x^{(2)}_j$ are normalized and non-normalized values of the $j$-th parameter of intensity/duration in the $j$-th utterance $x_j$.

In this study, we use in some of the algorithms of non-linear, non-parametric mapping of vectors from the multidimensional space of parameters to a plane according to Sammon's criterion [19]. This criterion makes it possible to locate the points in a plane at a minimal pairwise distance between them. The approximate distances between the corresponding vectors (intonation contours) in the multidimensional space of acoustic parameters, the criterion is formulated as follows:

$$\text{min}_x \sum_{j=1}^{N} \left( \delta_j - d_j \right)^2,$$

where $\delta_j$ is the distance between the $j$-th and $j$-th intonation contours in the multidimensional space; $d_j$ is the distance between the $j$-th and the $j$-th points in a plane; $\delta_j$ is the error of approximation, minimized through location of points in a plane.

To follow from the conditions that the location of points in a plane parameter (approximately) the location of intonation contours in the space of acoustic parameters, provided that the error of approximation (8) is negligible. Therefore, the clustering of intonation contours with respect to each intonation type can be assessed in terms of the position of the corresponding points on a plane. A cluster of points pertaining to a specific intonation type can be regarded as a natural size in the space of acoustic parameters.

The experimental data processing was computerized via ECLIPS MV/6000.

**Results**

The results of this study are displayed in Figure 1 below. It reveals the arrangement of intonation contours being analyzed on a plane. The error of approximation (9) is 0.48.

The lines in the figure delimit the clusters of points corresponding to the intonational sense of a final statement, a reply statement, a general question, an exclamation and a non-final statement. The axes in question can be correlated with intonations of standard Russian.

**Conclusions**

The cluster-seeking technique used in this study has been found to be highly effective in analyzing intonation. The technique can be regarded as a development of the algorithm method reported in [19].

Possible areas of further linguistic research involving the above technique include description of the intonational system of a language in terms of phonological oppositions, the study of a foreign accent in intonation, intonational types, phonetics etc.

**References**


Figure 1. Clustering results reflecting location of intonation contours on a plane.

---

*) The intonation contours are identified in the following manners:

- final statement
- reply statement
- general question
- exclamation
- non-final statement