ABSTRACT

The paper proposes a technique for automatic intonation type recognition for the purpose of designing an intonation training device to facilitate the acquisition of prosody by foreign language learners as well as to promote intonation improvement of those who have a speech or hearing disorder. The technique is applicable to any utterance with a fixed number of syllables and was initially tested for disyllabic speech samples of Russian and English.

1. INTRODUCTION

Broadly stated, the object of this study is to provide a means for computer-assisted phonetic instruction through which sound effects produced by a spoken voice are perceived by the learner as well as to help the learner to deal with points representing the curves on a display. A cluster analysis algorithm, proposed in [4], is used to reduce every curve to a point which is mapped on a plane. In order to make the initial cluster more convenient for teaching all extraneous data (i.e., points belonging to other clusters) are eliminated. As is shown in Fig. 1, the remaining points of each cluster are linked up in straight lines. It is assumed that the interpoint distance within the cluster does not exceed an empirically found fixed value. The procedure helps the learner to assess the cluster structure for the reference samples storage zone. When a new, prosodically distorted realization is mapped on a display, a decision as to which cluster it belongs to is made, judging by the two nearest neighboring points between which it is located. The method is known as "the nearest neighbour strategy". Some possible results are exemplified in Fig. 2. The learner can modify his intonation realization as many times as required until the point on a display reaches the right cluster, that is to say, until the learner's point is located between two reference points of the target cluster.

With implementation of the training device in question visualization of intonational zones is attained, thereby enabling the learner not only to hear intonation but also to"see" it without being tied to a specific intonation curve or a set of curves. In this case the language learner has to deal with points representing the curves on a display. A cluster analysis algorithm, proposed in [4], is used to reduce every curve to a point which is mapped on a plane. In order to make the initial cluster more convenient for teaching all extraneous data (i.e., points belonging to other clusters) are eliminated. As is shown in Fig. 1, the remaining points of each cluster are linked up in straight lines. It is assumed that the interpoint distance within the cluster does not exceed an empirically found fixed value. The procedure helps the learner to assess the cluster structure for the reference samples storage zone. When a new, prosodically distorted realization is mapped on a display, a decision as to which cluster it belongs to is made, judging by the two nearest neighboring points between which it is located. The method is known as "the nearest neighbour strategy". Some possible results are exemplified in Fig. 2. The learner can modify his intonation realization as many times as required until the point on a display reaches the right cluster, that is to say, until the learner's point is located between two reference points of the target cluster.

This sort of a technical aid for teaching intonation may be PC-based.

4. CONCLUSION

Displaying visual information for speech training purposes, as reported here, will enable the learner to make a selection from a series of options available in the reference storage zone for a particular intonation type with due account of the individual range of the learner's vocal performance.

5. REFERENCES


TOWARDS DESIGNING AN INTONATION TRAINING DEVICE BASED ON SPEECH SIGNALS CLUSTERING

Leonid A. Kanter, Alexander V. Savin and Ksenia G. Guskova

Dept. of Phonetics (English), Lenin Pedagogical State University, Moscow, USSR

ABSTRACT

The paper proposes a technique for automatic intonation type recognition for the purpose of designing an intonation training device to facilitate the acquisition of prosody by foreign language learners as well as to promote intonation improvement of those who have a speech or hearing disorder. The technique is applicable to any utterance with a fixed number of syllables and was initially tested for disyllabic speech samples of Russian and English.
Fig. 1. Internal cluster structure for the reference samples storage zones (interpoint distance ≤ 170). The test phrase OH ЗНАЛ, pronounced by 16 male native speakers of Standard Russian was considered.

Symbols for intonation types read as follows:
+ - final statement
× - reply statement
◇ - general question
↑ - exclamation
X - non-final statement

Fig. 2. Reference samples storage zones and non-native imitations of the Russian test phrase OH ЗНАЛ, produced by English learners of Russian. Non-native imitations are marked in smaller symbols within an ellipse.