

## THE TYPOLOGICAL ANALYSIS OF EMOTIONAL SPEECH PROSODY

EMMA NUSHIKYAN

English Dep. Odessa State University Odessa, Ukraine,  
USSR, 270021

## ABSTRACT

Emotion and its linguistic expression form a system: information about emotion comes by lexical cues, syntactic structures and prosodic indicators. This paper describes some further attempt to identify the acoustic parameters of emotional texts in English, Russian and Ukrainian.

## INTRODUCTION

In recent years a convergence of interest has developed among linguists /1/, psychologists /2/, physiologists /3/ and other specialists in the theory of emotion. Significant research has been conducted in the area of the verbal and non-verbal types of emotional expression /4/. Many recent writers on the manifestation of emotion have found it natural and useful to state that emotion, language and speech are related in some way as language being the principle mode of communication is also a means of expressing emotions and arousing them in others. The main aim of the present study is to reveal the system of linguistic devices of expressing emotions in English, Russian and Ukrainian.

## APPROACH

Emotion and its linguistic expression form a system. In order to characterize and understand its function we need to

consider the system properties.

Our investigation of 30000 pages of English fiction has pointed out that all levels of linguistic system may be involved in the process of emotional manifestation.

The syntactic analysis of the material under investigation proves that the structural means of expressing emotions form a special syntactic code. Language inspired by emotion undergoes a wide variety of formal changes - a breakdown of grammatical structure of the sentence, repetitions of subjects, subjects + predicates, introduction in the sentence structure of the formal elements - interjections, addresses, particles and so on.

At the lexical level we must find to what extent words alone can still yield further information in the study of emotional speech. Usually emotions are identified in the text due to the adjectives which emphasize the high degree of some quality (e.g. magnificent, awful etc.), the corresponding adverbs (e.g. terribly, exceedingly etc.), nouns (e.g. treasure, ruffian), verbs (e.g. hate, adore).

Despite the fact that in the majority of cases the emotive meaning depends on the integration of prosodic, syntactico-semantic and contextual information the role of intonation must not be ignored as there are cases in which it is the only

carrier of affective information.

## SUBJECTS AND SPEECH MATERIAL

In this research ten English, six Russian and four Ukrainian speakers participated as subjects. To preserve the identity and comparability of the experimental material 115 English texts and their Russian and Ukrainian translations were chosen. These texts expressed the 16 most frequently observed positive and negative emotions. As the emotional aspect of textual prosody implies that the effective messages can be superimposed on the neutral texts the material was recorded twice - as samples of emotional and non-emotional speech. The tape was then presented to 30 Russian, 30 English and 30 Ukrainian listeners who were instructed to listen carefully to the presentation of the texts and to decide which emotion the speaker was trying to convey. A criterion of 95% correct identification was set for including the text into the further instrumental analysis.

The original speech material was instrumentally analyzed into separate components - fundamental frequency, amplitude, duration and spectral composition. Oscillograms were obtained with the help of the Visi-Pitch and M-4030-1 computer which has been programmed to yield the desired speech parameters. Spectrograms were made on Sona-Graph of the Kay Elemetrics Corporation, using the wide-band filter (300 Hz). The synthesis of emotional speech was done on formant synthesizer.

## DATA ANALYSIS AND RESULTS

The typological study of emotional and neutral texts in English, Russian and Ukrainian disclosed common ways of expressing emotion. The detailed contrastive analysis of their acoustic structure has revealed that the frequency range, frequ-

ency interval of the terminal tone, of the semantic centre are the most informative parameters differentiating emotional and non-emotional speech in all the languages under study. In this way the similarity of emotion expression is manifested. The application of methods of mathematical statistics (t ratio, Student's t) proved that the difference between the above mentioned acoustic parameters of emotional and non-emotional speech was statistically significant as they belong to two different population variencies.

Our study presents experimental evidence that it is the movement of the fundamental frequency (its configuration) which bears the meaning of emotional intonation in a peculiar language. For instance, the utterances expressing anger are pronounced by the majority of English speakers with a sliding scale, while Russian and Ukrainian speakers use the broken descending scale with two peaks of the fundamental frequency. The corresponding neutral utterances are pronounced with the gradually descending scale with the peak of the fundamental frequency on the first stressed syllable. On the whole in the majority of emotional texts (with the exception of the texts expressing sorrow, tenderness, offence) a more complicated character of the fundamental frequency was observed.

It appears probable that both configuration and pitch levels have to be specified for certain emotions in the languages under study. The pitch level is higher in the texts expressing anger, delight, joy, amazement and lower in those expressing sorrow, guilt, tenderness, offence as compared with the neutral texts.

The textual aspect of prosody enables us to state the degree of communicative text dynamism. Each text contained sen-

tences in which semantic and emotive information was concentrated. They were called communicatively strong, while the other sentences were called communicatively weak. It is the range of the fundamental frequency, the frequency interval of the first and the last stressed syllable which is greater in communicatively strong sentences as compared with the communicatively weak ones in all the languages under study.

The study of the hierarchical structure of the temporal composition of emotional and non-emotional texts shows that the difference can be observed at all the levels - separate sound, syllable, utterance and text.

The increase in duration of utterances expressing irony, joy, contempt comes from the increase of vowel duration. The duration properties of consonants influence are also of great importance for emotional speech. The analysis of the spectrograms of emotional speech shows a longer stop gap and a more intense burst at the release of the English p, t, k and a lengthened interval of vocal tract contraction for  $\int$ , s, tʃ, dz. For Russian the increase in duration of fricatives s, ʃ, is of greater importance, for Ukrainian -h, x, š.

The mean duration of the syllable and the duration of the first and last stressed syllable are also characteristic differentiating parameters for emotional and non-emotional speech. The research shows that duration is the most variable acoustic parameter.

The temporal organization at the textual level reveals specific segmentation of emotional texts. Emotional state of the speaker has manifested in the change of duration of phonation parts of the texts and pauses. To measure this influence the relative informative duration parameter

$\Theta$  has been introduced.

$$\Theta = \frac{\Theta_{ph} + \Theta_p}{\Theta_p}$$

where  $\Theta_{ph}$  - phonation duration,

$\Theta_p$  - pause duration.

The quantitative analysis of the intensity of emotional texts proves that in most cases greater total energy of the text is observed due to the increase of energy of the first and last stressed syllable. The decrease of total energy occurs in the texts expressing sorrow, tenderness, offence, worry.

Spectrographic measurements of formant frequencies of emotional and neutral speech show a constant increase of the total energy of the stressed vowel of the nucleus at the expense of  $F_1$  and  $F_2$  in emotional speech. We also found the frequency range enlargement as well as the greater importance of  $F_3$  and  $F_4$ . The shift of the intensity of formant frequencies of the stressed vowels of the nucleus into higher regions was noticed in the utterances expressing strong emotions (e.g. rage, amazement, threat, delight and so on) and into the lower regions in the utterances expressing sorrow, offence, tenderness. The spectrographic analysis also reveals the greater role of the high frequency noise regions of consonants in emotional speech.

Statistical analysis of the main acoustic characteristics shows that each of them can distinguish emotional and neutral texts, but none has a differentiating function. The integral use of these characteristics enables us to create the algorithm of recognition of emotional and neutral utterances. For creating this algorithm A.Wald's sequential analysis based on the likelihood ratio has been applied.

The final stage of instrumental analysis is presented in a series of experiments on synthesis of emotional speech the aim of which is to check the validity of the acoustic parameters of natural emotional speech and to form the rules for high quality formant synthesis.

The synthesis programme was as follows: first the so called "neutral" sentence was modelled, then the emotional one with the identical lexico-grammatical structure. Altogether 30 Russian sentences expressing delight, surprise, anger, fear were obtained. The constructed model comprised the normalized contours ( $F_{0n}$ ) of an initial, mid and final accent groups (A.G.), each containing prenuclear (1), nuclear (2) and postnuclear parts (3). Fig. 1 represents the normalized fundamental frequency contours for sentences, expressing surprise (a), delight (b), fear (c) and the same neutral ones.

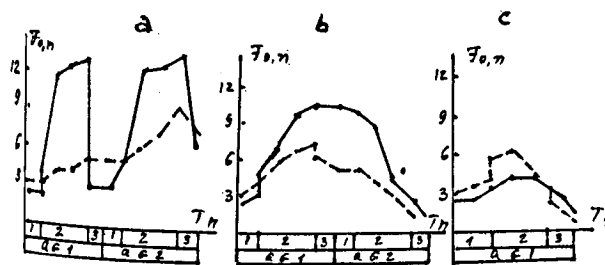


Fig. 1. Normalized fundamental frequency contours of sentences expressing (a) surprise, (b) delight, (c) fear are plotted on the contours of the same neutral ones. The solid line curve displays the emotional fundamental frequency contour, while the broken line displays the corresponding neutral one.

The high degree of perceptual acceptability of the synthesized emotional speech was proved by the auditive analysis. The tape was presented to 45 students of Russian department of Odessa University instructed to identify the type of emotion. Their answers presented in table 1

demonstrated the correct application of the synthesis rules.

Table 1. Identification of the type of emotion in synthesized speech.

Type of emotion	: Correct answers (%)
surprise	: 95
delight	: 90
fear	: 73

Our experiments on synthesis show that for expressing some emotions (e.g. surprise) fundamental frequency is the most important prosodic variable, and others need only minimal changes. For expressing other emotions some changes of formant structure must be added to a neutral sentence.

#### CONCLUSIONS

The systematical level approach to emotion realization in a text made it possible to single out some regularities in the interaction of components in this system and to disclose the isomorphism in its expressive means.

The results of this typological study permit us to suppose that prosodic structure of emotional texts in English, Russian and Ukrainian displays universal as well as particular properties in manifestation of emotion.

#### REFERENCES

- 1/ Bryzgunova E.A. Emotsionalno-stilisticheskiye razlichija ruskoj zvuchaschej rechi. - Moscow: MGU, 1984.
- 2/ Vitt N.V. Emotsionalnaja regulatsija rechevogo povedeniya // Vopr. psichologii, 1981. - N 4. - S. 60-69.
- 3/ Simonov P.V. Emotsionalny mozg. - Moscow: Nauka, 1981.
- 4/ Williams C.H., Stevens K.N. Emotion and speech: some acoustic correlates // J. Acoust. Soc. Am., 1972. - v. 55. - P. 1238-1250.