PROBLEMATIC DIFFERENTIATION OF EMOTIONAL STATES VIA PROSODY IN ENGLISH PHONETICS

Kateryna Lysenko
PhD in Philology, associate professor, Department of foreign languages for math faculties Institute of Philology Taras Shevchenko National University of Kyiv, Ukraine

Perception and differentiation of the emotional state of the speaker through the analysis of prosodic models is one of the least studied areas of modern phonetics. Because rapid assessment of emotions is important for identifying and prioritizing information, both linguistic and paralinguistic analysis tools need to be used to recognize them.

Emotions are conveyed in spoken words through verbal and nonverbal mutually informative channels, which are revealed in parallel. To understand the functional organization of emotions, the basic rules of perception of emotional facial expression, voice and words are important.

Spoken words, of course, contain speech and paralinguistic elements that are important and mutually informative for communication. Speech information consists of the literal, symbolic meaning of a word, while paralinguistic information also includes contextual meanings of the word. For example, the meaning of the word "crazy", whether pronounced "mentally disturbed", "fierce" or "wildly agitated", can be distinguished based on the assessment of contextual paralinguistic information, such as the current emotional state of the speaker, revealed by his tone of voice and facial expression [1]. Speech and paralinguistic fragments of information are revealed in parallel, while the spoken word unfolds gradually.

In the auditory system, the voice is perceived in parallel with facial expressions, because it collectively conveys both a person’s personality and his emotional state. It is believed that some aspects of vocal emotions (including such an emotional category as anger, disgust, fear, sadness, joy) are perceived quickly based on the tone and intensity of analysis of short segments of familiar nonverbal vocalizations (e.g. shouting, crying, laughing, etc.), and this information is transmitted immediately and directly. However, the analysis of other aspects of vocal emotions through emotional prosody to recognize the identity of emotions may require sampling of longer vocal segments and their more detailed spectral analysis.

The voice is also a natural carrier of speech. Voice paralinguistic and linguistic signals are separated in such a way that the low-frequency band primarily carries prosodic signals important for the transmission of emotions, while the high-frequency band primarily carries phonemic signals important for verbal
communication [2]. Speech voice in general is a socio-cultural phenomenon, it is an indicator of social index information, the sound of which gives an idea of education, social status, profession of the speaker and the national features of his voice. Speech voice is a basic sound component of rhetorical discourse and contributes to the realization of the discursive plan of the speaker, which depends on the genre and its individual preferences. Thus, the manifestation of the components of the speech voice in speech depends on the type of rhetorical strategies of the speaker. Noting the specifics of the manifestation of individual components of speech voice in the implementation of rhetorical strategies of the speaker, it should be emphasized that primarily use different voice modulations, as well as shades of tone, accompanied by variation of intonation parameters. At the same time, the natural quality of the voice and the timbre of the speaker further influence the choice of voice modulations. An important feature of this trend is that voice variation depends on the type of rhetorical strategy of the speaker. In this case, voice modulations are multifunctional in nature, thus, the same voice modification contributes to the realization of different communicative intentions of the speaker, and vice versa, the same rhetorical strategy can be implemented by several modulations. It depends primarily on the type of speech, as well as on the acting skills of the speaker.

Voice is a particularly important means of conveying an emotional state, as it is relatively independent of the listener and the speaker and his ability to see (as opposed to facial signals). Acoustic signals that transmit voice emotion consist of pitch (fundamental frequency), volume (intensity), rhythm (duration of segments and pauses) and timbre (distribution of spectral energy) [3] and are modulated by physiological factors (e.g. heart rate, blood flow, muscle tension), which vary depending on the emotional state of the person. It is believed that the two main aspects of the voice convey the emotional state on different time scales. The process of speech, consisting of pitch, volume contour and rhythm of speech articulation, evolves relatively slowly during suprasegmental intonations of speech. The quality of nonverbal vocalization, which consists of timbre and abrupt, aperiodic spectral changes, manifests itself more quickly [4], and clearly conveys certain emotional categories (e.g. fear, disgust) [5]. Like emotional faces, emotional voices help with the perception of emotions, as evidenced by better recognition of emotional rather than neutral nonverbal vocalizations [6], [7].

Voice quality includes not only various sound modes, but also all changes, intentional or unintentional, that the speaker is capable of, given the physiological limitations of the vocal organs, taking into account longitudinal articulatory parameters (raised and lowered laryngeal voice, labial protrusion and labiodental voice), latitudinal parameters (labial, lingual, focal and pharyngeal), velopharyngeal and phonatory parameters (including combined phonation methods).

Additional voice qualities are non-linguistic voice effects that pass through speech or interrupt it, and include laughter, giggles, trembling, sobbing and crying.

In general, researchers of emotional prosody have a standard approach to methods and tools of investigation. The main emotions are usually read both by intonation and facial expression. Minor fluctuations in the coefficients of emotion determination through prosody can be caused by factors such as the gender of the
speaker and sometimes the influence of the translator’s personality, as well as a certain category of emotions and style and semantic load in the utterance.

**References:**


