

## Abstract

The goal of this research work is to find the answer to the question about what role non-linguistic information plays in the perception of expressive speech categories, and specifically, in the case with people who are from different culture/native-language backgrounds. The solid support of the two assumptions made for of expressive speech perception is found.

1. The first assumption states that before listeners can decide to which expressive speech category a speech sound belongs, they will qualify a voice according to different descriptors, where each descriptor is an adjective for voice description.
2. The second assumption states that people who are from different cultures/native-language background have some common characteristics for perception of expressive speech categories as well as some differences.

The two assumptions are firstly represented as a layer-structure model. which is based on the fact that there is a vagueness nature in human cognition. This model illustrates that people perceive expressive speech categories (i.e. the topmost layer) not directly from a change of acoustic features (i.e. the bottommost layer), but rather from a composite of different types of “smaller” perceptions that are expressed by semantic primitives (i.e. the middle layer). By this three-layer model, the focal points of the research work become the building and evaluation of these two relationships. After this, the model is applied to find the common features as well as differences of people who are or are not acquainted with the language of the voice they heard. To achieve the research goal, this model will be built by perceptual experiments, verified by rule-based speech morphing, and applied to the analysis of non-linguistic verbal information.

In the building process, in order to find those descriptors (i.e., adjectives) that can be used to describe the perception of expressive vocalizations, a thorough selection of semantic primitives should be conducted. After the semantic primitives are selected, analysis of a large number of acoustic features will be necessary to support the relationship between semantic primitives and acoustic features. Moreover, in order to understand the fuzzy relationship between the linguistic description of acoustic perception and expressive speech, a fuzzy inference system will be built. After the two relationships are built, the effectiveness of relationships should be evaluated.

The resulting model by the building process suggests that before listeners decide to which expressive speech category a speech sound belongs, they qualify a voice according to different descriptors, where each descriptor is an adjective for voice description. Different combinations of such descriptors by which a listener qualifies a voice can lead to the decision about to which different expressive speech categories that voice belongs.

To verify the resulting model, a rule-based speech morphing approach is adopted. The built model is transformed to rules, which are used to create morphed utterances that are perceived as different semantic primitives and expressive speech categories with different intensity levels. The morphing rules, which represent the changes of acoustic features in expressive speech, are created from the analyzed acoustic features and the built fuzzy inference system. STRAIGHT provides the tool for morphing neural utterances by the created rules. The verification process goes a step further to show the validity of the finding in the building process.

Finally, the same model is applied with the same stimuli but different listeners with different culture/native-language background to show that the role non-linguistic information plays in perception of expressive speech categories is common to people who are from different culture/native-language background.

The results of the building and the evaluation of the model provide solid support for the first assumption. The results of the model application show an example that suggests the possibility of the second assumption.