A basic study on emotional perception related to F0 contour by using fMRI

Yasuhiro Hamada (0810050)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 9, 2010

Keywords: fMRI, Emotional perception, Fundamental frequency contour (F0 contour).

Speech contains linguistic information (What are said) and non-linguistic information (emotion, individuality and gender etc.). Many researches have focused on linguistic information, although non-linguistic information is also important for speech communication. Recently, non-linguistic information is attractively attentioned for importance.

Many researches reported that the prosody is strongly related to emotions. Acoustic features forming prosody are mainly F0, power envelope and duration. Hayashi reported that F0 contour conveys much emotional information, from acoustic feature analyses and listening tests using interjectory word /eh/.

Brain activity can be measured through recently-developed instruments (e.g., fMRI). Many psychologists and neurologists reported results of brain activity measurements elicited by emotional voices. Wiethoff et al. reported that emotional voices affect right mid superior temporal gyrus rather than natural voice. Bach et al. suggested that the left inferior frontal gyrus plays a specific roll in explicit evaluation of emotional prosody. However, these reports did not consider what acoustic features affect to the brain activity. To understand speech communication by emotion, it is needed to investigate the brain activity related to emotional speech perception influenced by acoustic features. This paper investigates brain activi-

ties using synthesized voices controlled acoustic features based on results of Hayashi's research.

In this paper, as the experiment paradigm, we measured brain activities using synthesized stimuli controlled acoustic features. In our experiments, we presented six stimuli synthesized from interjectory word /eh/ with different acoustic features (F0, power envelope and spectral frequency). One is an original voice(S0) synthesized without arranging a natural voice, and the others(S1 ∼S5) are synthesized by modifying acoustic features using STRAIGHT. For the measurement, it is necessary to use high-natural stimuli that can be perceived as different emotions. F0 contour, power envelope and spectral frequency were modified carefully by paying attention to naturalness.

We conducted a psychoacoustic experiment. About the five synthesized voices and the original one, subjects were asked to answer what emotions are included. From the evaluated scores, dominant emotional words for each stimulus were *Affirmation* and *Sympathy*, *Affirmation* and *Calm* (S1), *Disappointment* and *Sadness* (S2), *Asking again* and *Surprise* (S3), *Doubt* and *Negative* (S4) and *Doubt* and *Surprise* (S5).

In the brain activity measurement, eighteen Japanese subjects with normal hearing were participated. Brain activities were measured by using fMRI. We presented six stimuli and noise to perform an oddball task. Subjects were instructed to press button when they hear the noise. We analyzed the differences of brain activities to be listened an original voice with the five synthesized voices.

Results show that each stimulus elicit superior temporal gyrus, middle temporal gyrus, supramarginal gyrus and middle frontal gyrus belonging to auditory area. These areas are said to process the difference of sound stimuli in previous reports. The difference of the activity on the superior frontal gyrus and left inferior frontal gyrus included on orbital area was shown in S1 - S0. In S2 - S0, superior parietal lobule related to sensory area and parietal association area was more activated. In S3 - S0, S4 - S0 and S5 - S0, cerebellum, caudate nucleus and putamen in basal ganglia were activated. These regions are related to speech perception and production.

By referring to the hierarchical hypothesis of feeling, we explain relationships between results of listening tests and those of brain activity measure-

ments. The hierarchical hypothesis of feeling is the hypothesis that feelings in humans are composed of primitive emotion(around a hypothalamus), basic emotion(limbic system), social feeling and intellectual feeling(cerebral cortex) with respect to evolution. This hypothesis also propose attentional system was evoluted another system of feeling.

We discuss relationships between the results of the listening test and those of the brain activity measurements by referring to the hierarchical hypothesis of feeling. Results show the stimulus S1(*Affirmation* and *Calm*) minus original voice S0(*Affirmation* and *Sympathy*) elicited different activities on superior frontal gyrus belonging to cerebral cortex. The cerebral cortex is considered to be evaluated with processing of social feeling and intellectual feeling. These results consist with hierarchical hypothesis of feeling. The stimulus S2(*Disappointment*, *Sadness*) minus original voice S0(*Affirmation* and *Sympathy*) elicited different activities on superior parietal lobule belonging to cerebral cortex. These results indicate that *Disappointment* and *Sadness* are related in social feeling and intellectual feeling. The stimulus S3, S4 and S5(*Surprise*, *Doubt*, *Negative* and *Asking again* minus original voice S0(*Affirmation* and *Sympathy*) elicited mainly caudate nucleus or putamen belonging to basal ganglia. The activity on basal ganglia is thought in processing of primitive emotion. Because basal ganglia assumes the adjustment systems of action affected by body homeostasis, emotion of attention, and tension such as *Surprise*, *Doubt*, *Negative* and *Asking again* is supposed to relating these adjustments system though still need more information of emotions.