Study on pitch perception of noise-vocoded harmonic complex tones

1810168 Yukina HOSAKA

The sound has a temporal fine structure indicating the source signal and an amplitude envelope is extracted by the gradual change of the source signal. In the amplitude modulation, the temporal fine structure refers to the carrier, and the amplitude envelope refers to the modulated wave. The temporal fine structure is important for pitch perception which perceives pitch, and sound image localization which perceives the position of sound. And the amplitude envelope is important for perception of linguistic information.

There are two theories in the theories of pitch perception. One is the temporal theory and the other is the rate-place theory. The temporal theory is the theory that pitch is related to the time patterns of auditory nerve firing triggered by the sound. There is a phenomenon called phase-locking as a property of the auditory nerve. This is a phenomenon in which the auditory nerve firing occurs at a specific phase of the sound waveform, and the interval of the auditory nerve firing approaches an integral multiple of the sound waveform period. Therefore, pitch is perceived based on the temporal pattern of the auditory nerve firing. On the other hand, the rate-place theory is a theory that sounds of different frequencies perceive the pitch of a sound by firing an auditory nerve having different characteristic frequencies in different places of the cochlea. Even now, no conclusion has been reached as to which of these theories prevails.

Cues derived from temporal fine structure play an important role for pitch perception. However, the noise-vocoded sounds have only a temporal amplitude envelope and does not have the temporal fine structure. Therefore, it is known that pitch perception of the noise-vocode sounds is too difficult.

Cochlear implant users have some problems in the perception of music, the prosody of speech, and the danger alerts in everyday life.

So far, studies using noise-vocoded sounds, which is one of the sounds simulating the hearing of cochlear implant users, have been conducted. It has been clarified that the amplitude envelope information of speech contains non-linguistic information such as personality and emotion as well as linguistic information, and sense of urgency (para-linguistic information). The perception of emotions and urgency in speech requires the prosodic features such as pitch. In addition, pitch perception using amplitude envelope has been studied. Shamma et al. have studied biological models of pitch perception combining spectral and temporal information, and Shofner et al. have examined the role of harmonic structure in pitch perception. From these facts, it is considered that the amplitude envelope information contains the cues related to pitch.

The amplitude envelopes of sounds can be perceived by both normal hearing and cochlear implant users. Therefore, if it is possible to clarify the possibility of pitch perception based on the amplitude envelope information of sounds, it will be possible to clarify the mechanism of pitch perception and to improve the quality of life (QoL) of cochlear implant users.

This aims of this study is to investigate the possibility of pitch perception using the amplitude envelope information. The first experiment was conducted by Thurston's paired comparison to investigate whether or not the pitch perceptual scales of harmonic complex tones mimicking musical instruments and its noise-vocoded sounds were placed comparable, respectively. Here, to use the musical pitch as a measure of the pitch of the sound, the harmonic complex tones mimicking musical instruments as the original sound. As results, it was found that the correct rates of pitch perception of noise-vocoded sounds are almost the same as those of the harmonic complex tones. And the pitch perceptual scales of the noise-vocoded sounds and harmonic complex tones were almost identical by conducting a Thurston's paired comparison to investigate whether the pitch perceptual scales of harmonic complex tones mimicking musical instruments. However, we have not yet revealed what kind of cues can play an important role in pitch perception. This point will be analyzed in the next experiments.

Next, the second experiment was conducted by the same ways in the first experiment to investigate whether or not the results of the first experiments are affected under spectral-tilt conditions by using both stimuli of harmonic complex tones and noise-vocoded sounds with three spectral tilts of growing down, flat, and growing up. As results, it was found that the pitch perceptual scales of noise-vocode sounds are relatively affected by spectral-tilt conditions while those of harmonic complex tones are not affected by spectral-tilt conditions. Hence, these suggest that cues derived from the temporal amplitude envelope play an important role for pitch perception although these are related to spectral-title conditions.

In addition, the third experiment was conducted by the same ways in the first and second experiment using conducted pitch discrimination studies of noise-vocoded sounds with an increased cutoff frequency when extracting the amplitude envelope to investigate cues for pitch discrimination of noise-vocoded sounds. As results, it was found that the increase in the information of the amplitude envelope caused by increasing the cutoff frequency does not affect the pitch discrimination of the noise-vocoded sounds.

Furthermore, to investigate cues for pitch discrimination of noise-vocoded sounds, the relationship between the excitation pattern and modulation spectrum and the pitch discrimination results was investigated. As a result, the pitch scales from the analysis of the excitation patterns and the modulation

spectrum did not match the pitch scales of the noise-vocoded sounds by the Thurston's paired comparison. Therefore, it was suggested that there was no relation between the excitation patterns and modulation spectrum and the pitch discrimination results.

These results suggest that the pitch perception of the noise-vocoded sounds is affected by the low-pass filter when extracting the spectral gradient and the amplitude envelope, but not by the increase in the cutoff frequency.

From the results of this experiment and analysis, it cannot be said that important cues for pitch perception is contained the amplitude envelope information. However, in Chapters 4 and 5, it was suggested that pitch discrimination was possible under the right-down condition. In this study, it was not possible to clarify the mechanism of pitch perception of amplitude envelope information and discuss the strategy of acquiring pitch perception for cochlear implant users.

In this study, the pitch discrimination of noise-vocoded sounds was investigated, but it is not clear what pitch the noise-vocoded sounds have. Therefore, to investigate what pitch is perceived from the noise-vocoded sounds using the adjustment method, etc., it is possible to further examine the cues of pitch discrimination of the noise-vocoded sounds. In addition, if it becomes clear that cues for pitch discrimination of noise-vocoded sounds, the investigation of pitch discrimination by cochlear implant users using the sound emphasizing the cues will help to obtain the pitch perception of the cochlear implant users. It is possible to the strategy can be discussed.