Effects of Frequency Range, Timbre and Intonation on Melodic Contour Identification with Acoustic Simulation of Cochlear Implant

Lichuan Ping[1, 2], Meng Yuan[1], Qinglin Meng[1, 2] and Haihong Feng[1]
1 Shanghai Acoustics Laboratory, Institute of Acoustics, Chinese Academy of Sciences, Shanghai, 200032, China
2 Graduate University of Chinese Academy of Sciences, Beijing, 100190, China
pinglichuan@gmail.com

Abstract

The aim of this study was to determine how melodic contour identification with cochlear implant was affected by the frequency range, timbre and intonation. Noise-excited vocoder was used to simulate the signal processing for cochlear implant system. Melodic contour identification tests were carried out with eight normal-hearing subjects. Stimuli from four instruments (piano, trumpet, clarinet and violin) were utilized with different frequency ranges and intonations. Our experimental results showed that (i) temporal envelope information was important for melodic contour identification when the spectral information was limited; (ii) frequency range, timbre and intonation could influence melodic contour identification performance of the normal-hearing subjects when listening to the acoustic simulation of cochlear implant.

1. Introduction

Cochlear implant (CI) is an electronic device used to restore hearing of the patients with severe to profound cochlear hearing loss. Although good speech recognition has been achieved by many patients with the state-of-the-art CI devices in quiet environment, the speech understanding in noise, the perception of tonal languages and music are still reported to be exceptionally difficult for most of the CI users [1]. Many of them complain that music sounds unpleasant and noisy. There are many difficulties in the music perception of CI users [2, 3]. Some studies suggested that the poor music perception performance of CI users was probably due to the limited spectral and temporal information delivered through the CI system [3].

Melody is considered to be one of the most important features of music. The perception of melody requires fine discrimination of changes in pitch, including both the direction and the degree of pitch change. Familiar melody recognition (with or without rhythm) has been used to evaluate CI user's melody perception abilities [4-6]. These studies showed that CI users' familiar melody recognition performance was much poorer compared with normal-hearing (NH) subjects. An interesting study by Galvin *et al* tested the identification of melodic contour with CI users [7]. The melodic contour identification (MCI) reflects the ability to detect the direction and interval in change of pitch. In their study, the subjects were asked to identify one of nine 5-note melodic contours. The duration of the 5 notes were identical. The interval between successive notes was varied. Their results showed that the NH subjects achieved a mean score of about 95% but the CI user's performance varied from 14% to 90%. The score was increased as the number of semitones between successive notes was increased. They also found that the performance was slightly worse in the lower frequency range. In a follow-up study, Galvin *et al* further investigated the contribution of the timbre cues to the MCI performance [8]. Stimuli from six instruments (organ, glockenspiel, trumpet, clarinet, violin and piano) were tested with a similar paradigm of their previous MCI study. The result showed that the MCI performance with the organ was the best and the worst with the piano.

The above MCI studies have shown that frequency range, timbre and intonation all affected the MCI performance of CI users. However, to our knowledge, there was no study that has fully investigated these circumstances. In the present study, MCI test was carried out with acoustic simulation of CIs. Four instruments (piano trumpet, clarinet and violin), two frequency ranges and four intonations were tested. NH subjects were tested with both the unprocessed stimuli (as a control session) and the stimuli processed with the acoustic simulation. The simulation was based on a noise-excited vocoder to simulate the continuous Interleaved Sampling (CIS) strategy [9].

2. Experiment design

2.1. Subjects and test materials

Eight NH subjects participated in the experiment aging from 22 to 28 (Mean = 25.875 years; SD = 2.5 years). None of them had hearing disease before. Their audiometric thresholds were better than 20 dB HL at octave frequencies from 125 and 8000 Hz in both ears.

Figure 1 shows the nine melodic contours (Rising, Flat, Falling, Flat-Rising, Falling-Rising, Rising-Flat, Falling-Flat, Rising-Falling, and Flat-Falling). These were similar to those used in Galvin et al [7]. Each contour included five notes which were presented in sequence. The interval between successive notes varied among 1, 2, 4 and 6 semitones. The interval between adjacent notes was 50 ms in sequence.

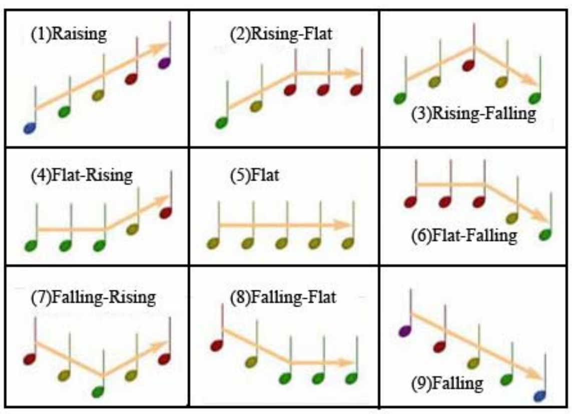

Figure 1. Nine melodic contours used in this study: Rising, Rising-Flat, Rising-Falling, Flat-Rising, Flat, Flat-Falling, Falling-Rising, Falling-Flat and Falling.

The melodic contours were generated according to a "reference note" (i.e., the mid note or the third note in the contour). There were two reference notes, G2 (98Hz) and G4 (392Hz). The frequency ranges for the two reference notes, G2 and G4, were 43.65 Hz (G1) - 196Hz (G3) and 196Hz (G3)-784 Hz (G5), respectively. These two frequency ranges were defined as the lower and higher frequency range respectively in this study. Four instruments representing different instrumental families were: piano, trumpet, clarinet and violin. All the notes were synthesized on the use of a professional MIDI-Keyboard. The duration of all the stimuli was 350 ms with 25 ms on- and off-ramp, respectively. Before signal processing, the intensity of all the stimuli were balanced with RMS (Root-Mean-Square) equalization.

2.2. Signal processing

All the test materials were processed with a 4-channel noise-excited vocoder [10, 11]. The original stimuli were first pre-emphasized on the use of a 1st-order high-pass Butterworth filter with cutoff frequency at 1 kHz. The signal was then band-pass filtered (4th-order Butterworth filters) into 4 contiguous frequency channels. The filters were designed using Greenwood function. The frequency range was from 125 to 4000 Hz. The signal envelope in each frequency band was extracted using a 4th-order Butterworth low-pass filter with a cutoff frequency of 500 Hz, followed by a half-wave rectification process. For each channel, the envelope was used to amplitude-modulate a white-noise carrier. Finally, the acoustic stimulus was generated by combining the modulated signals from each band.

2.3. Procedure

The test conditions included frequency range (lower and higher), timbre (piano, trumpet, clarinet and violin) and intonation (1, 2, 4 and 6 semitones interval). There were totally 32 different test conditions as shown in Table 1.

MCI performance was measured independently for each condition; during each test block, a stimulus was randomly selected (without replacement) from the 36 stimuli (9 contours*4 intonations). The subjects were asked to click on one of the nine response choices shown on the laptop screen. Subjects were allowed to repeat each stimulus up to three times; the test order of all the conditions was randomized for each subject. The final scores were reported as a percentage of correct responses on the melodic contours. The experiment was carried out in a sound-proof booth. A graphical user interface (GUI) in the laptop was used. All the stimuli were played to the listeners through a Sennheiser HD 650 headphone at 65 dB SPL.

3. Results

For each condition shown in Table 1, the percentage correct MCI performance was evaluated over all subjects. Figure 2 show the mean MCI performance of NH subjects listening to unprocessed stimuli (NH_US) in panel A and stimuli processed via acoustic simulation (NH_AS) in Panel B, respectively.

Table 1. Test conditions

Conditions	Intonation interval (semitones between two successive notes)				
Timbre and frequency range	**1**	**2**	**4**	**6**	**all**
Piano + Lower	P_L_1	P_ L _2	P_ L _4	P_ L_6	P_L
Piano + Higher	P_H_1	P_H_2	P_H_4	P_H_6	P_H
Trumpet + Lower	T_L_1	T_ L _2	T_ L _4	T_ L_6	T_L
Trumpet + Higher	T_H_1	T_H_2	T_H_4	T_H_6	T_H
Clarinet + Lower	C_L_1	C_ L _2	C_ L _4	C_ L_6	C_L
Clarinet + Higher	C_H_1	C_H_2	C_H_4	C_H_6	C_H
Violin + Lower	V_L_1	V_ L _2	V_ L _4	V_ L_6	V_L
Violin + Higher	V_H_1	V_H_2	V_H_4	V_H_6	V_H

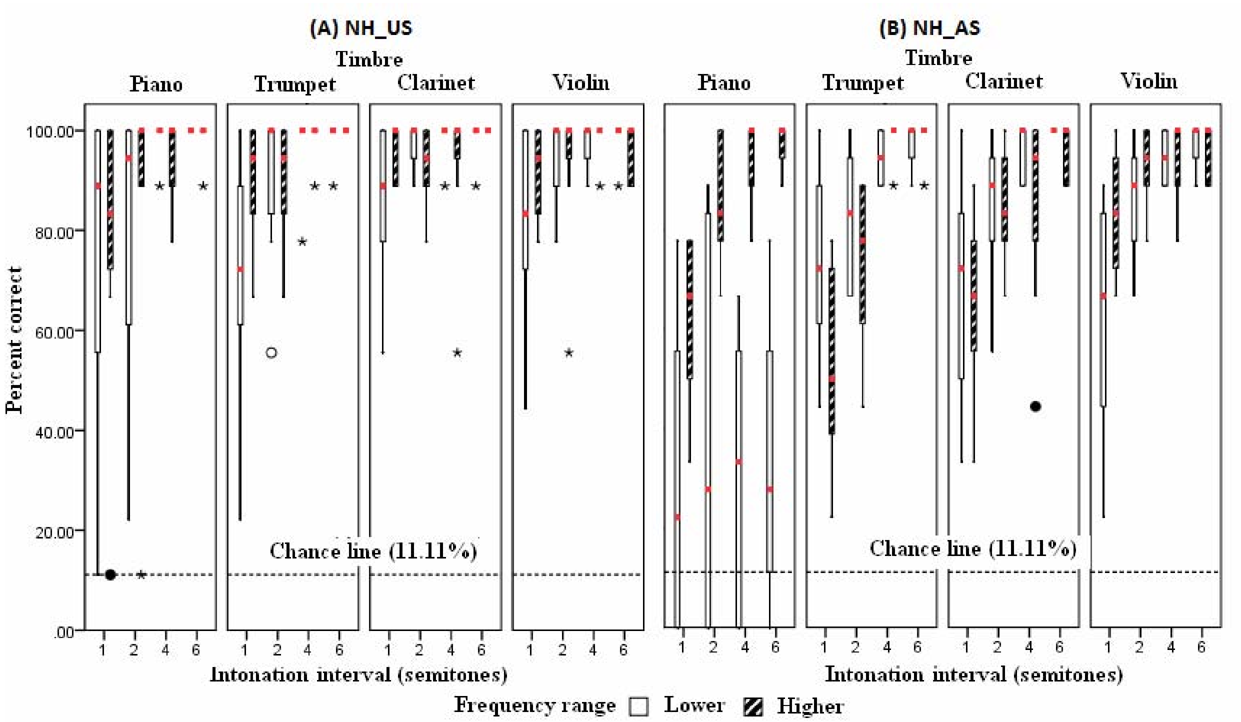

Figure 2. The MCI performance of NH subjects listening to unprocessed stimuli (left, Panel (A)) and listening to processed stimuli with acoustic simulation (Right, Panel (B))

The mean performance was shown for different frequency ranges, timbres and intonations (the bold line in each box represents the median MCI performance). The results for different timbres were presented in each The mean performance of NH_US across all conditions was 92.49% compared to 78.99% of NH_AS.

For NH_US, overall mean performance was good (>85% correct) in most conditions except the T_L_1 condition (68.01% correct). The main effects of both frequency range [$F(1, 10)= 7.0805$, $p= 0.0239$] and timbre factors [$F(3, 30)= 6.0384$, $p= 0.0024$] were significant. No significant main effect [$F(3, 30)= 1.3547$, $p=0.2755$] was found for the intonation factor. Post-hoc tests showed that the MCI performance under P_L_1, P_L_2, and T_L_1 were significantly worse than the other conditions ($p<0.05$). For NH_AS, the performance under the condition of P_L (32.29%) was worse than all the other conditions. Significant main effects were found for all factors and all the interactions ($p<0.05$). Post-hoc tests confirmed that, the MCI performance under P_L was significantly worse than other conditions. A trend was shown that (i) the MCI performance became better when the interval increased; (ii) the performance of the higher frequency range was significantly better than that of the lower frequency range in most cases ($p<0.05$).

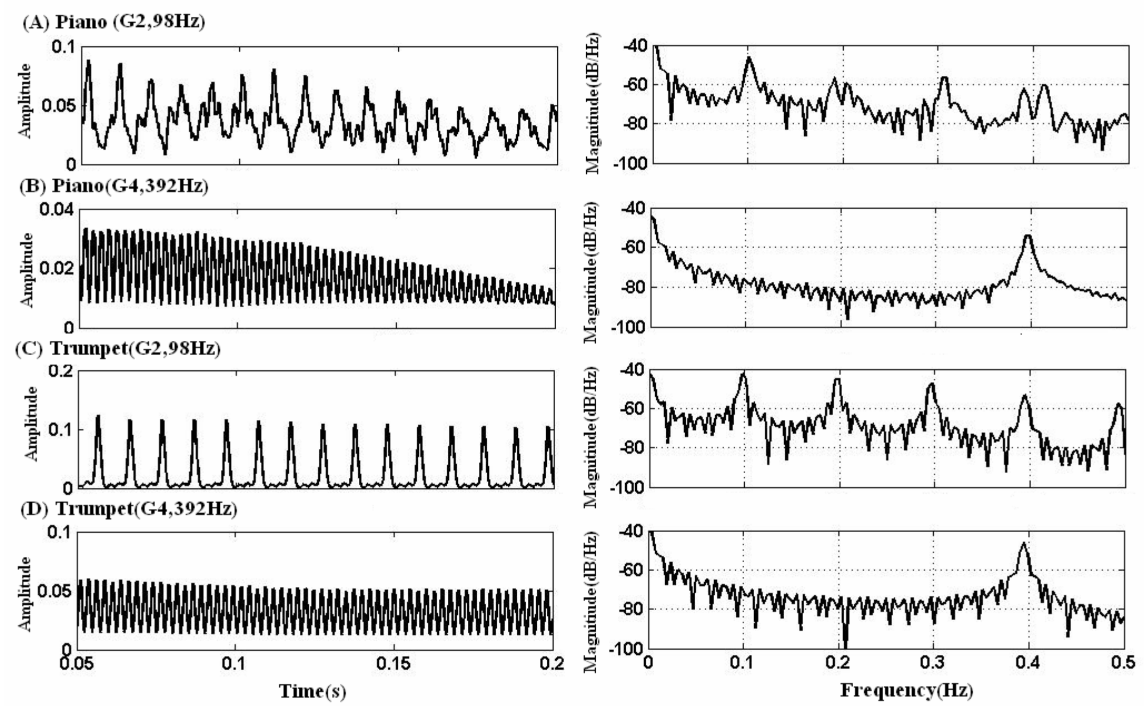

Figure 3. The envelopes (left) from the third channel (934Hz- 1975Hz) and their corresponding spectrums (right) of piano and trumpet with G2 (98Hz) and G4 (392Hz)

4. Discussion

The present study showed that frequency range, timbre and intonation influenced MCI performance of NH_AS. The MCI performance improved as the interval (semitones between two successive notes) increased for all the instruments in most case as expected. The results were worse in lower frequency range (reference root G2), particularly for piano stimuli.

Many studies have been carried out to investigate the importance of temporal envelope to speech recognition [12-14]. It was shown that temporal envelope was important to tonal language perception when the spectral information was limited [10]. In the present study, the noise-excited vocoder was used to simulate the signal processing of the CIs. Limited spectral resolution was given by the four frequency bands. For the temporal resolution, the temporal envelope was extracted with a cutoff frequency of 500 Hz, in order to preserve F0-related information (F0 and some low-order harmonics).

Figure 3 shows the temporal envelopes and their corresponding spectrums from the third channel (934.4Hz-1975Hz). Two instruments, piano and trumpet with lower F0 (G2 at 98Hz) and higher F0 (G4 at 392Hz) were presented in Panel (A), (B), (C) and (D). For G2, the F0 and its low-order harmonics could be easily distinguished in the spectrums of trumpet. Regular F0-related information was not found in the piano spectrum compared to trumpet. For G4, clear F0 information was found in both piano and trumpet.

It was shown that the temporal envelope of piano with lower F0 was irregular compared to the others. Therefore, clear F0-related information was not found in the G2 piano spectrum. Our experimental results showed that temporal envelope could provide F0-related information for relatively good MCI performance.

5. Conclusion

The experimental results of the present study showed that: (i) temporal envelope information (<500Hz) could provide relatively good MCI performance when the spectral cues were limited; (ii) frequency range, timbre and intonation influenced MCI performance of the NH subjects listening to acoustic simulation of cochlear implant.

Acknowledgements

We are grateful to all the NH listeners for participating in this study. This work has been supported in part by the Chinese Academy of Sciences Pilot Project of the Knowledge Innovation Program (KGCX2-YX-607) and Key Projects in the National

Science & Technology Pillar Program in the Eleventh Five-year Plan Period (2008BAI50B08).

References

[1] Zeng, F. G, Rebscher, S., Harrison, W., Sun, X., Feng, H., "Cochlear Implants: System Design, Integration, and Evaluation", IEEE Reviews in Biomedical Engineering, vol.1, 2008, pp. 115-142.

[2] McDermott, H. J., "Music perception with cochlear implants: a review", Trends Amplif, vol. 8, 2004, pp. 49–82.

[3] Ward R. Drennan and Jay T. Rubinstein, "Music perception in cochlear implant users and its relationship with psychophysical capabilities", J Rehabil Res Dev. vol. 45(5), 2008, pp. 779–789.

[4] Kang, R. S., Nimmons, G. L., Drennan, W. R., et al., "Development and validation of the University of Washington Clinical Assessment Of Music Perception test", Ear Hear., vol. 30(4), 2009, pp.411-418.

[5] Gfeller K, Turner C, Mehr M, Woodworth G, Fearn R, Knutson JF, Witt S, Stordahl J. "Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults", Cochlear Implants Int, vol.3 (1), 2002,pp. 29–53.

[6] Kong, Y.-Y., Cruz, R., Jones, J. A., et al., "Music perception with temporal cues in acoustic and electric hearing", Ear Hear., vol. 25, 2004, pp. 173–185.

[7] Galvin, J., Fu, Q.-J., and Nogaki, G., "Melodic contour identification by cochlear implant listeners," Ear Hear., vol.28, 2007, pp.302–319.

[8] Galvin, J., Fu, Q.-J., and Oba, S., "Effect of instrument timbre on melodic contour identification by cochlear implant users," J. Acoust. Soc. Am., vol.124, 2008, pp.EL189–EL195.

[9] Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM, "Better speech recognition with cochlear implants", Nature, vol.352, 1991, pp.236-238.

[10] Meng Yuan, et at., "Cantonese tone recognition with enhanced temporal periodicity cues" J. Acoust. Soc. Am., vol.126 (1), 2009, pp. 327-337.

[11] Green, T., Faulkner, A., and Rosen, S., "Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved sampling cochlear implants," J. Acoust. Soc. Am., vol.112, 2002, pp. 2155–2164.

[12] Kong, Y.-Y., and Zeng, F.-G., "Temporal and spectral cues in Mandarin tone recognition", J. Acoust. Soc. Am., vol.120, 2006, pp. 2830–2840.

[13] Shannon, R.V., Zeng, F.-G., Kamath, V., Wygonski, J., Ekelid, M., "Speech recognition with primarily temporal cues", Science, vol.270, 1995, pp. 303–304.

[14] Xu, L., et al., "Features of stimulation affecting tonal-speech perception: Implications for cochlear prostheses," J. Acoust. Soc. Am., vol.112, 2002, pp. 247–258.