Simulation Study of Channel Number of Cochlear Implant in Quiet State

J. Wang and Y. S. Chen(✉)

School of Electronics and Communication, Shenzhen Institute of Information Technology, Shenzhen, Guangdong 518172, People's Republic of China
chenyoushengtsinghua@aliyun.com

Abstract. This paper studied the effect of changing channel number of cochlear implant in low, medium, and high frequency band on Mandarin Chinese tone recognition in quiet state, respectively. The input speech was filtered by a bandpass filter bank with the frequency range of 60–8000 Hz. The number of channels in low or medium frequency band was set as 2, 4, 8, 16, and 32, respectively. The number of channels in high frequency band was set as 1, 2, 4, 8, 16, and 32, respectively. In each pass band, temporal envelope was extracted by full wave rectification and low-pass filtering. The envelope was modulated with a band-limited noise of the corresponding channel. To avoid signal leakage, the modulated signal was band-pass filtered again. Then Root Mean Square equalization was performed. Then, the signals of each channel were summed to obtain a simulated composite signal for each frame. Finally, the signals of each frame were superimposed and averaged to obtain a speech signal. Tone recognition experiments found that, the number of channel in cochlear implant had a certain effect on tone recognition rate of Mandarin Chinese in quiet state. As the number of channel was increased, the tone recognition rate rose slowly. For the recognition of tone 3, the channel type (low, medium, and high) had a significant impact.

Keywords: Cochlear implant · Mandarin chinese tone recognition · Channel number

1 Introduction

Cochlear implant is a hearing recovery device. It directly stimulates auditory nerve fibers of the severe hearing impaired patients with a weak current, excites auditory nerve fibers to mimic the physiological functions of the peripheral auditory system, generates a neural release pattern similar to normal persons, and thereby partially restores the hearing of deaf patients [1]. At present, the vast majority of cochlear implant products adopt speech processing strategies based on temporal envelopes, which can effectively recognize non-tonal language. However, these strategies do not encode the tonal information of speech, making it difficult for users speaking Mandarin Chinese to perceive tonal language.

In order to enhance the tonal information of the cochlear implant, researchers studied the effect of channel number on tone recognition of Mandarin Chinese. Some

M. Guan and Z. Na (Eds.): MLICOM 2020, LNICST 342, pp. 242–249, 2021.
https://doi.org/10.1007/978-3-030-66785-6_28

researchers found that in quiet state, when the total number of channels of the cochlear implant was increased from 1 or 2 to 4, the tone recognition rate was not significantly improved [2, 3]. However, Xu *et al*. used Chinese single syllable words of the same duration to conduct simulation experiments in quiet state [4]. Xu *et al*. found that, when the total number of channels was increased from 1 to 4, the tone recognition rate increased significantly, and that the recognition rate of 3 or 4 channels was significantly higher than that of 1 channel. This phenomenon contradicts Fu's previous results. In addition, in the experiment of Xu *et al*., when the number of channels continued to increase from 4, the tone recognition rate of 6, 8, 10, and 12 channels no longer increased significantly, but reached a platform. Besides, the recognition rates of these channels were all significantly higher than the recognition rate of only 1 channel. In addition, the recognition rates of 8 and 12 channels were significantly higher than those of 2 channels.

It is worth noting that the variable in the above studies is the total number of channels. The current paper changed the number of channels in the low, medium, and high frequency bands of the cochlear implant, respectively, and tried to explore the impact of channel number in different frequency bands on tone recognition of Mandarin Chinese.

2 Methods and Materials

2.1 Theoretical Methods

A frequency range $[f_l, f_h]$ is divided into n channels by a multiplier factor x. Then Eqs. (1) should be satisfied.

$$x = (f_h/f_l)^{1/n} \tag{1}$$

Among the n channels, the higher cutoff frequency of the i-th channel f_{ih} can be calculated by Eq. (2).

$$f_{ih} = f_l * x^i \qquad (1 \le i \le n) \tag{2}$$

The lower cutoff frequency of channel $i + 1$ equals the higher cutoff frequency of channel i. Thus frequency range of each channel and the multiplication factor x can be obtained from Eqs. (1) and (2).

In this experiment, the frequency range [60, 8000] Hz was divided into three parts, including the low frequency band ([60, 1000] Hz), the medium frequency band ([1000, 4000] Hz), and the high frequency band ([4000, 8000] Hz).

Change Number of Channels in the Low Frequency Band. Keeping 2 channels for the medium frequency band (1000–2000, 2000–4000 Hz) and 1 channel for the high frequency band (4000–8000 Hz) unchanged, increase the number of channels in the low frequency band from 2 to 32 (i.e. 2, 4, 8, 16, and 32) to explore the impact of the number of the low frequency band on tone recognition rate. This situation is represented as “Low” case. The total number of channels in this case is from 5 to 35 (i.e. 5, 7, 11, 19, and 35).

Change Number of Channels in the Medium Frequency Band. Keeping 2 channels for the low frequency band (60–244.9, 244.9–1000 Hz) and 1 channel for the high frequency band (4000–8000 Hz) unchanged, increase the number of channels in the medium frequency band from 2 to 32 (i.e. 2, 4, 8, 16, and 32) to explore the impact of the number of the medium frequency band on tone recognition rate. This situation is represented as “Medium” case. The total number of channels in this case is from 5 to 35 (i.e. 5, 7, 11, 19, and 35).

Change Number of Channels in the High Frequency Band. Keeping 2 channels for the low frequency band (60–244.9, 244.9–1000 Hz) and 2 channels for the medium frequency band (1000–2000, 2000–4000 Hz) unchanged, increase the number of channels in the high frequency band from 1 to 32 (i.e. 1, 2, 4, 8, 16, and 32) to explore the impact of the number of the high frequency band on tone recognition rate. This situation is represented as “High” case. The total number of channels in this case is from 5 to 36 (i.e. 5, 6, 8, 12, 20, and 36).

2.2 Speech Processing Strategy

This paper proposed an improved continuous interleaved sampling (CIS) strategy based on the channels designed in Sect. 2.1. The voice was first enframed, added with a Hamming window, and passed through an FIR linear phase band-pass filter bank. The channels of the filter bank were designed by the method introduced in Sect. 2.1. Then, full-wave rectification and low-pass filtering was applied on the signal of each channel to obtain an envelope. The envelope was modulated with a band-limited noise of the corresponding channel. To avoid signal leakage, the modulated signal was band-pass filtered again. Then Root Mean Square (RMS) equalization was performed. Then, the signals of each channel were summed to obtain a simulated composite signal for each frame. Finally, the signals of each frame were superimposed and averaged to obtain a speech signal.

2.3 Materials

The experiment used a phonetic vocabulary of Mandarin Chinese recorded by Fu *et al.* [5], with a total of more than 1,500 female pronunciation words. The voice was processed by Matlab software, passed through the sound card (Echo Indigo IOx, 24-bit resolution and 44100 Hz sampling rate), and played through Sennheiser HD 650 headphones. The subjects were tested in double-layer soundproof rooms.

2.4 Subjects

Twelve college students (3 male and 9 female) aged between 19 and 23 years participated in the experiments. All listeners were native Mandarin speakers with normal hearing (thresholds below 15 dB HL) at octave frequencies between 125 and 8000 Hz in both ears. Each listener was trained for at least 1 h until the results became stable.

2.5 Test Content

Test contents include total tone recognition rate (Total), one tone recognition rate (T1), two tone recognition rate (T2), three tone recognition rate (T3), and four tone recognition rate (T4).

3 Results

3.1 The Result of Changing Number of Channels in Low Frequency Band

When the number of low-frequency channels was 2, 4, 8, 16, and 32, the total number of channels over the entire frequency range was 5, 7, 11, 19, and 35, respectively.

Increasing the number of low-frequency channels, Total, T1, T3, and T4 all rose slowly, but there was no significant difference. T2 showed fluctuations.

Each tone recognition rate was ranked from high to low as: T3 > T4 > T2 > T1. Total remained at 63%–70%, which is consistent with previous experimental results [2, 3].

In addition, LSD pairwise comparison found that, for T4, recognition rate for the 2 low-frequency channels was significantly lower than that for the 32 low-frequency channels (p = 0.041 < 0.05) (Fig. 1).

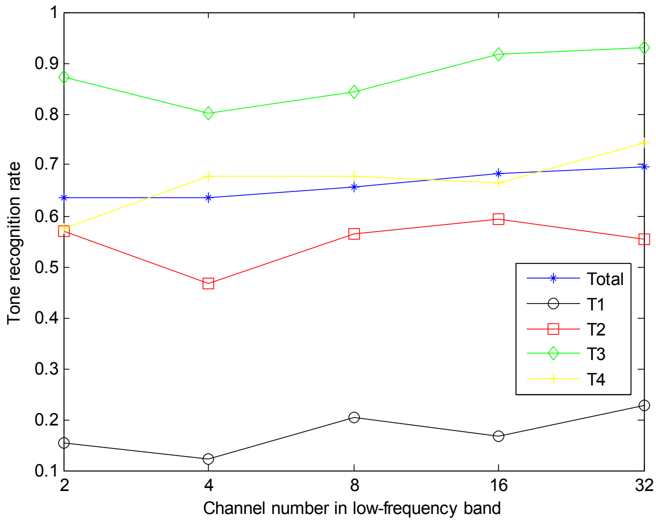

Fig. 1. The effect of changing number of channels in low frequency range on tone recognition rate in quiet state

3.2 The Result of Changing Number of Channels in Medium Frequency Band

When the number of medium-frequency channels was 2, 4, 8, 16, 32, the total number of channels over the entire frequency range was 5, 7, 11, 19, and 35, respectively.

Increasing the number of medium-frequency channels, Total, T1, and T3 all rose slowly, but there was no significant difference. T2 and T4 fluctuated.

Each tone recognition rate was ranked from high to low as: T3 > T4 > T2 > T1. The overall tone recognition rate was maintained at 60%–70%, which is consistent with the previous experimental results [2, 3] (Fig. 2).

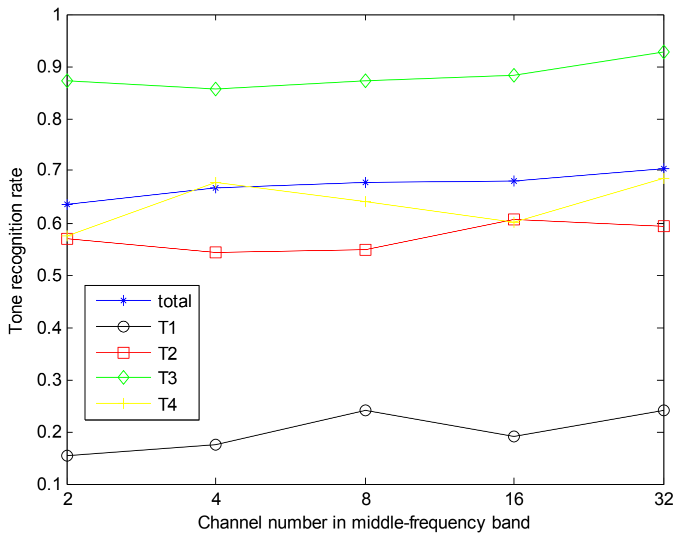

Fig. 2. The effect of changing number of channels in medium frequency range on tone recognition rate in quiet state

3.3 The Result of Changing Number of Channels in High Frequency Band

When the number of high-frequency channels was 1, 2, 4, 8, 16, and 32, the total number of channels over the entire frequency range was 5, 6, 8, 12, 20, and 36, respectively.

Increasing the number of high-frequency channels, Total, T1, T2, T3, and T4 rose slowly or fluctuated.

Each tone recognition rate was ranked from high to low as: T3 > T4 > T2 > T1. Total remained at 60%–72%, which is consistent with the previous experimental results [2, 3].

For T3, the recognition rate for 8 high-frequency channels was significantly lower than that for 32 high-frequency channels ($p = 0.042 < 0.05$).

For T1, the recognition rate for 1 high-frequency channel was significantly lower than that for 16 high-frequency channels ($p = 0.032 < 0.05$) (Fig. 3).

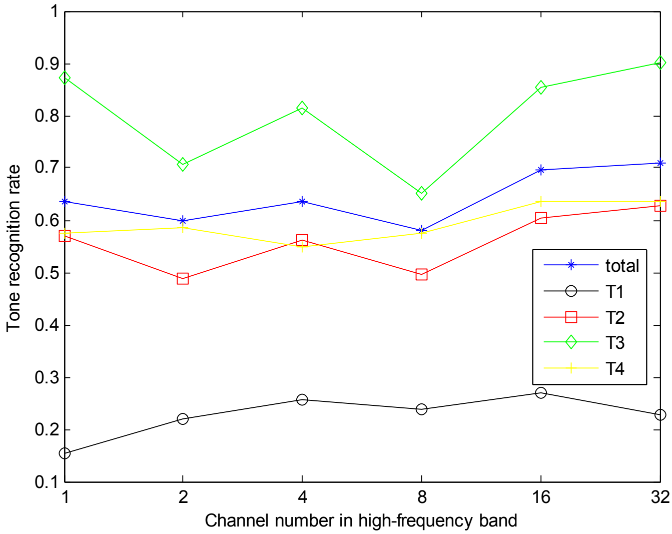

Fig. 3. The effect of changing number of channels in high frequency range on tone recognition rate in quiet state

4 Discussion

When the number of channels in the low, medium, and high frequency bands was changed respectively, Total for the three cases was maintained between 60% and 72%, which is consistent with previous results [2, 3]. Luo xin and Fu Qian-jie used noise band carrier to carry out simulation experiments, and found that the 2-channel and 4-channel tone recognition rates were about 62% in quiet state [2]. The research of Fu *et al*. found that, when the number of channels was 1 to 4, the score of Mandarin Chinese tone recognition was stable at about 60%–80% in the quiet state [3].

For the low, medium, and high cases, the recognition rates of each tone had a common phenomenon. T3 was the highest, followed by the T4 and T2 in sequence, and the lowest was T1. This indicates that tone 3 and tone 4 are easier to recognize, and the tone 1 and tone 2 are not easy to recognize. This is consistent with previous results [3]. The reason why tone 3 and tone 4 are easier to identify may be related to duration. Tone 3 has the longest length and tone 4 has the shortest length. However, the contribution of duration cue might be small [6].

When the number of channels in the low, medium, and high frequency bands was increased respectively, the tone recognition rates fluctuated or rose slowly. However, there was no significant difference. This is similar with previous results [2]. Luo xin and Fu Qian-jie used noise band carrier to carry out simulation experiments, and found that, when the total number of channels increased from 2 to 4, Mandarin Chinese tone recognition in quiet state was not affected by the total number of channels in the spectrum [2, 3]. Fu *et al*. had similar conclusions [3]. This explains from another point of view that for the CIS simulation, 4–6 channels are enough to achieve a good recognition rate of Mandarin Chinese tones in quiet state.

However, LSD pairwise comparison found that, for T4 in the low frequency band, 2-channel tone recognition rate was significantly lower than the 32-channel tone recognition rate ($p = 0.041 < 0.05$). This may be related to the easier identification of tone 4. The curve of tone 4 changes from high to low significantly with the shortest length [3, 6]. Another reason is that when the number of channels is 32, the tone recognition rate is very similar to that of the original sound. Kong and Zeng found that, when the total number of channels was increased from 1 to 2, the tone recognition rate was basically the same based on temporal envelope. When the total number of channels was increased from 4 to 12, the tone recognition rate increased. When the number of channels was 32, the tone recognition rate was similar to that of the original sound [7]. Therefore, for tone 4 recognition in the low frequency band, recognition rate with 32 channels was significantly higher than that with 2 channels.

In addition, for tone 1 recognition in the high frequency band, LSD pairwise comparison found that, recognition rate of 1 channel was significantly lower than that of 16 channels ($p = 0.032 < 0.05$). This can also be explained by the fact that recognition rate can be improved significantly when the number of channels is big enough.

Moreover, non-parametric test found that type of channel (low, medium, and high) had a significant effect on T3 ($p = 0.011 < 0.05$). LSD pairwise comparison revealed that there were significant differences between the rate for 32 channels in low frequency band and that for 8 channels in the high-frequency band ($p = 0.042 < 0.05$).

Variance analysis and non-parametric test found that type of channel had no significant effect on Total, T1, T2, or T4. However, LSD pairwise comparison found three significant differences. First, for T1, recognition rates for 4 channels in low frequency band was significantly ($p < 0.05$) lower than those for 4, 8, 16, and 32 channels in high frequency band. Second, the former was significantly ($p < 0.05$) lower than those for 8 and 32 channels in medium frequency band. Finally, for T4, recognition rates for 4 and 8 channels in the high frequency band were significantly ($p < 0.05$) lower than those for 32 channels in the low frequency band. This can also be explained by the fact that recognition rate can be improved significantly when the number of channels is big enough.

5 Conclusion

In quiet state, the number of channels has a certain effect on the tone recognition rate of Mandarin Chinese. As the number of channel increases, the tone recognition rate rises. For the recognition of tone 3, the channel type (low, medium, and high) has a significant impact. However, channel type has no significant impact on other tone recognition, except for a few local significant difference found by LSD pairwise comparison.

Acknowledgements. This paper was granted by the characteristic innovation project of Guangdong University in 2019 (Grant No. 2019GKTSCX094) and Guangdong Province higher vocational colleges & schools Pearl River scholar funded scheme (2016).

References

1. Wei, C.G., Cao, K., Zeng, F.G.: Mandarin tone recognition in cochlear-implant subjects. Hear. Res. **197**(1–2), 1–95 (2004)
2. Luo, X., Fu, Q.J.: Importance of pitch and periodicity to Chinese-speaking cochlear implant patients. In: IEEE International Conference on Acoustics. IEEE (2004)
3. Fu, Q.J., Zeng, F.G., Shannon, R.V., et al.: Importance of tonal envelope cues in Chinese speech recognition. J. Acoust. Soc. Am. **97**(1), 505–510 (1998)
4. Xu, L., Tsai, Y., Pfingst, B.E.: Features of stimulation affecting tonal speech perception: implications for cochlear prostheses. J. Acoust. Soc. Am. **112**(1), 247–258 (2002)
5. Fu, Q.J., Zhu, M., Wang, X.: Development and validation of the Mandarin speech perception test. J. Acoust. Soc. Am. **129**(6), 267–273 (2011)
6. Whalen, D.H., Xu, Y.: Information for mandarin tones in the amplitude contour and in brief segments. Phonetica **49**(1), 25–47 (1992)
7. Kong, Y.Y., Zeng, F.G.: Temporal and spectral cues in Mandarin tone recognition. J. Acoust. Soc. Am. **120**(5), 2830 (2006)