A COMPARATIVE STUDY ON PERCEPTUAL CHARACTERISTICS OF MANDARIN CONSONANTS IN DIFFERENT ACOUSTIC TRANSMISSION CONDITIONS

Siyu Zhang & Zihou Meng

Communication Acoustics Laboratory, Communication University of China, Beijing, China
zsy608@gmail.com

ABSTRACT

In order to compare the effects of different acoustic transmission conditions on Mandarin consonants perception, two subjective speech intelligibility tests were separately carried out in reverberant and noisy condition. The results showed that the affricates and fricatives were the most intelligible in reverberation. Friction might be the most important cue for consonant perception in reverberation. Those consonants articulated frictionally were more intelligible than others in low frequency noise, while the voiced were more intelligible in high frequency noise. The primate cues for consonant perception in noise were related to the frequency characteristic of noise. For both stops and affricates, the unaspirated sounds were more intelligible than the aspirated sounds in both reverberant and noisy conditions.

Keywords: consonant perception, transmission condition, reverberant effect, noise masking

1. INTRODUCTION

Speaking and listening are the most convenient communication type, but it will become difficult if there is some interference in surrounding. The effects of environmental factors on phonemes are not uniform. Due to the shorter length and weaker energy, the consonants are more easily masked by noise than vowels. Speech intelligibility would deteriorate sharply because of consonants loss. Perceiving consonants correctly, therefore, is primate for the overall speech intelligibility level. A Chinese Mandarin syllable is the most readily distinguished speech segment for hearing and the initial refers to the first consonant in a syllable. All the Mandarin consonants can serve as initials except the nasal /ŋ/. The Mandarin consonants mentioned in this paper refer to the 21 initials in Mandarin.

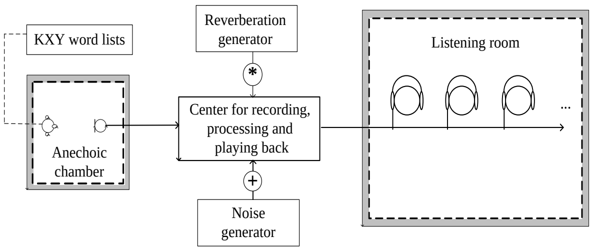
In the research field of consonant perception, some focused on the special subjects' speech perception such as children, hearing impaired people or non-native speakers [1]. For most native speakers with normal hearing, how to perceive speech correctly and steadily even in a poor listening condition has been a hot topic, some suggested that there were some important features contained in speech segments and be responsible for perception. Miller and Nicely [4] indicated five features to distinguish Sixteen English consonants by carrying out a series of listening experiments: voicing, nasality, affrication, duration and place. Singh and Black [5] found that the nasals were more intelligible and the fricatives and aspirated were less intelligible for both native and non-native speakers. Singh, et al. [6] compared the contribution weight of different articulation features to English consonants perception. Zhang, et al. [7-9] indicated the perceptual features of Mandarin consonants were: voicing, aspiration, friction and place, and then build distinctive feature systems of Mandarin. Zhang and Meng [10] indicated that the perceptual characteristic of Mandarin consonants in reverberation was closely related to the place of articulation.

It is noticeable that the conclusions mentioned here were usually made by carrying out listening experiments or speech intelligibility tests in noisy or frequency distorted transmission conditions. In fact, besides noise and frequency distortion, speech intelligibility is also affected by room acoustical condition [3]. Generally, background noise and room reverberation are the two typical acoustic interferences on speech perception. Study on the relationship between acoustic transmission condition and speech perception needs to be explored more. In this paper, we tried to compare the effects of reverberation and noise on Mandarin consonants perception.

2. EXPERIMENTAL METHOD

It is ideal to carry out speech intelligibility tests in real places with various acoustic conditions, but it is actually difficult to realize because of many uncontrollable factors. We, therefore, used a simulated method as shown in figure 1.The speech materials were recorded in an anechoic chamber, and then separately mixed with various simulated acoustic conditions and played back through headphones in a listening room. The experiments were designed according to a national standard GB 15508-1995 [2].

Figure 1: The block diagram of experimental method.



2.1. Speech materials

There were 36 different Chinese Mandarin speech intelligibility test word lists (KXY word lists) [2] used in two tests. Every list consists of 25 three syllables rows, total 75 syllables and keeps the balance of the level of difficulty and phonemic characteristic [2]. The three syllables in each row are randomly arranged and nonsense. There is a guiding phrase ‘No. n’ before each row ‘×××’. The ‘n’ stands for row number. All word lists were recorded by a female and a male broadcaster with a rate of four syllables per second in an anechoic chamber.

2.2. Acoustic transmission conditions

2.2.1. Reverberation

A simulated room impulse response with 6 seconds reverberation time and uniformly distributed later reflection was produced by the reverberant effect processor TC Reverber4000 and then processed by simulated sound absorption uniformly to generate 7 room impulse responses. All impulses located in the same receiving position but differed in reverberation time ranged from 0.5s to 3.6s. The early 10ms of each impulse was cut to avoid influence of direct sound and early reflections. The test signals in reverberant condition were made by convolving recorded word lists with room impulse responses.

2.2.2. Noise

There were 8 types of band-pass noise used in test, including 80-750Hz, 80-1600Hz, 80-2500Hz, 350-5000Hz, 1600-5000 Hz, 2500-5000 Hz and 5000-20000 Hz. 3 or 4 kinds of SNRs were set up in each noise type. The test signals in noisy condition were made by adding noise signals to recorded word lists according to given SNRs.

2.3. Subjects

Twelve subjects aged from 22 to 26 years old were chosen from graduate students in college with balanced gender ratio. All subjects can speak standard Mandarin and have no known hearing problems. They were familiar with Mandarin Pinyin spelling rules and were trained before tests start.

2.4. Speech intelligibility test

The tests were carried out in a listening room with short reverberation and low background noise. All test signals were played through headphones at a level of 70dBA. To avoid memorizing, no word list was reused in any other transmission conditions. After playing a three syllables row, the subjects were asked to write down what they thought they heard in Pinyin format. For example, writing down ‘fā nǐ huī’. Only if the written initial is the same as the played initial, the response is regarded as true. The proportion of initials perceived correctly is the so called Mandarin consonant intelligibility.

3. EXPERIMENTAL RESULTS

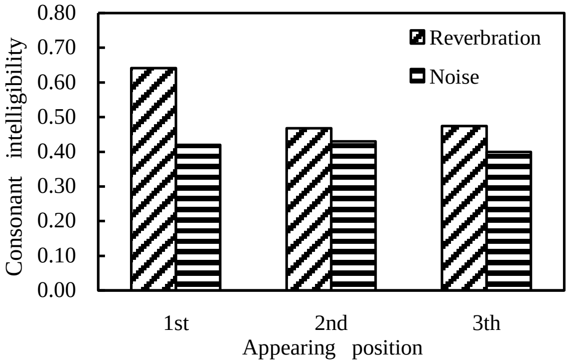
3.1. Consonant intelligibility on different appearing position

Due to the listening sequence of a three syllables row, there are 3 positions where initials would appear. The initial would appear in the beginning of the first, second or third syllable, the effect of transmission conditions especially the reverberant effect might be different on the 3 positions. We called the initial of the first syllable in a row ‘1st’, the initial of the second syllable ‘2nd’,the initial of the third syllable ‘3th’. After the validity inspection, the consonant intelligibility on each position in reverberation and noise were averaged across all valid subjects.

As shown in figure 2, consonant intelligibility of the first initial significantly differed from that of the second and the third initials in reverberation.

The first initial was less affected by reverberation so as to have a higher intelligibility, the second and the third initials were more affected so as to be less intelligible. While in noisy condition, the diversity of consonant intelligibility on 3 positions was little which reflects that masking effect on 3 positions by noise is uniform. The results showed that the interference strategies of noise and reverberation on speech perception might be different.

Figure 2: Consonant intelligibility on different appearing position.



In the next sections, the initials appeared on the second and third syllables were regarded as analyzing objects in reverberation, while initials appeared on all 3 positions were regarded as analyzing objects in noise. There were 16800 valid observations in reverberation and 25200 valid observations in noise.

3.2. Characteristics of consonant perception in reverberation

The reverberant effect on different types of Mandarin consonant will be discussed in this section. Mandarin consonants can be divided into stops, affricates, fricatives, nasals and voiced sounds et al. in terms of manner of articulation. Mandarin consonants can also be divided into front palatals, blade-alveolars, and retroflexes et al. in terms of place of articulation. The intelligibility of different types of consonant is closely related to the manner or place of articulation.

As shown in figure 3 and figure 4, the intelligibility of all types of consonant decreased with increasing of reverberation time. From the point of articulation manner, the intelligibility descending order was: fricatives and affricates, nasals, stops. It seemed that the reverberant effect was less on those articulated frictionally but stronger on stops. From the point of articulation place, the front palatals, blade-alveolars and retroflexes were less affected by reverberation than others. In fact, the palatals, blade-alveolars and retroflexes were exactly those frictionally articulated. The results suggested that friction might be the primate cue to perceive consonant in reverberation.

Figure 3: Consonant intelligibility of different manners of articulation in reverberation.

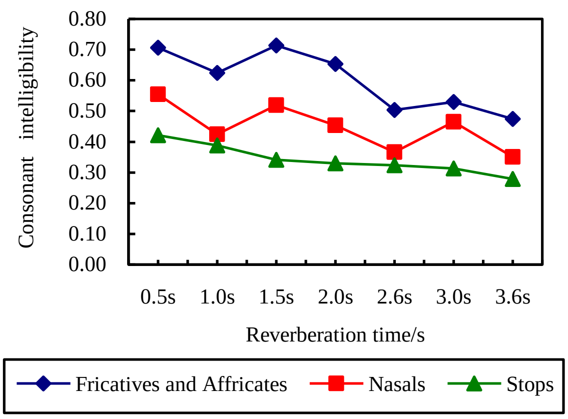
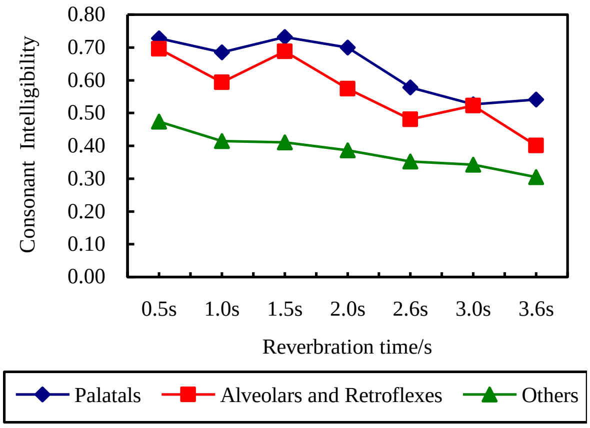


Figure 4: Consonant intelligibility of different places of articulation in reverberation.



3.3. Characteristics of consonant perception in noise

The noise masking effect on different types of Mandarin consonant will be discussed in this section. It was found that consonant intelligibility in noise was closely related to the manner or place of articulation, and the frequency characteristic of noise.

As shown in figure 5 and figure 6, from the view of articulation manner, the intelligibility descending order in low frequency noise was: affricates, fricatives, voiced, stops, and the order in high frequency noise was: voiced, fricatives, stops, affricates. From the view of articulation place, the masking effect of low noise on front palatals, retroflexes and blade-alveolars was less than that of high noise. It was suggested that the primate cue for consonant perception in noise might be various. Friction was more advantageous in low noise while voicing performed better in high noise.

Figure 5: Consonant intelligibility of different manners of articulation in noise.

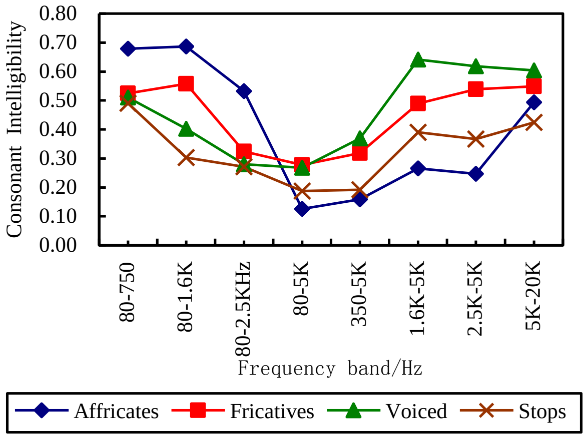
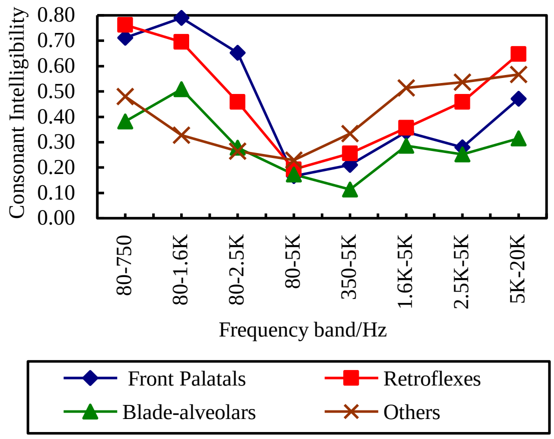


Figure 6: Consonant intelligibility of different places of articulation in noise.



4. DISCUSSION

This paper was aimed to compare the effects of different acoustic transmission conditions on Mandarin consonants perception. Two subjective evaluations of Chinese Mandarin intelligibility were separately carried out in reverberation and noise. The results showed that the effect of reverberation on consonant perception was different from that of noise. The affricates and fricatives were most intelligible in reverberation. Friction might be the most important cue for Mandarin consonant perception in reverberation. In noisy condition, consonant perception was closely related to the frequency characteristic of noise. Those consonants articulated frictionally were more intelligible than others in low frequency noise, while the voiced sounds were more intelligible in high frequency noise. The primate cue for consonant perception in different noisy conditions might be different. Friction was more advantageous in low noise while voicing performed better in high noise. We also found that for both stops and affricates, the unaspirated sounds were more intelligible than the aspirated sounds in both reverberant and noisy conditions, and this would be discussed in future work.

5. ACKNOWLEDGEMENTS

The authors are very grateful to students who participated in subjective evaluations of Mandarin intelligibility. This work was supported by Special Fund of Doctoral discipline points of China (Grant No. 20100018110001).

6. REFERENCES

[1] Carroll, D.W. 2004. *Psychology of Language, Fourth Edition*. Shanghai: East China Normal University Press.
[2] GB/T 15508-1995. 1995. Acoustic-speech articulation testing method. *Standard of PR China.* Beijing: Standards Press of China.
[3] Meng, Z.H., Dai L. 2009. A survey of the relationship of Mandarin speech articulation and sti-pa. *Audio Engineering* 33, 4-8.
[4] Miller, G.A., Nicely P.E. 1955. An analysis of perceptual confusions among some English consonants. *J. Acoust. Soc. Am.* 27, 338-252.
[5] Singh, S., Black, J.W. 1966. Study of twenty-six intervocalic consonants as spoken and recognized by four language groups. *J. Acoust. Soc. Am.* 39, 372-387.
[6] Singh, S., Woods, D.R., Becker, G.M. 1972. Perceptual structure of 22 prevocalic English consonants. *J. Acoust. Soc. Am.* 52, 1698-1713.
[7] Zhang, J.L. 2005. Distinctive feature system of Chinese Mandarin. *Acta Acustica.* 30, 506-514
[8] Zhang, J.L. 2006. The binary tree diagram of distinctive feature system of Chinese Mandarin. *Acta Acustica.* 31, 193-198.
[9] Zhang, J,L., Qi, S.Q., Lu, S.N. 1981. The primary investigation on perceptual structure of Chinese Mandarin consonants. *Acta Psychologica Sinica* 13, 76-85.
[10] Zhang, S.Y., Meng, Z.H. 2009. A cluster analysis of Mandarin initials based on perceptual similarity in reverberation. *Tsinghua Science and Technology* 49, 1262-1265.