

Research Article

Perceptual Characteristics of Chinese Speech Intelligibility in Noise Environment

Hui Song ¹ and Siyu Zhang ²

¹Art College, Northeast Electric Power University, Jilin, China

²Communication Acoustics Laboratory, Communication University of China, Beijing, China

Correspondence should be addressed to Hui Song; sandysong@neepu.edu.cn

Received 8 April 2020; Revised 3 June 2020; Accepted 11 June 2020; Published 24 July 2020

Academic Editor: Chenxi Huang

Copyright © 2020 Hui Song and Siyu Zhang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Speech intelligibility is affected by various interfering factors in a speech transmission system. Noise is one of the most common affecting factors. Subjective listening experiments were, respectively, carried out in pink noise, speech noise, and white noise-interfering environment. The perceptual characteristics of the initials, finals, tones, and syllable intelligibility were analyzed, and the function relationships between Chinese speech intelligibility and SNR in noise environment were concluded, which could be used to evaluate or predict the Chinese speech intelligibility under noise transmission conditions.

1. Introduction

Speech is a main way of human communication, and speech intelligibility is an important parameter to assess the acoustic quality of a communication system. ISO defines speech intelligibility as a professional term of *the percentage of correctly received speech units in all transmitted speech units* [1]. Speech could be interfered by various factors in the communication system, like noise, reverberation, distortion, and so on. Noise is one of the most common affecting factors. Noise unavoidably exists in our daily life and masks speech signals with different degrees, while speech intelligibility decreases with the noise level increasing. It is very necessary to investigate speech intelligibility in noise environment for the speech transmission system [2].

There were some studies on the influence of noise on speech intelligibility, which include both objective and subjective ones. For objective experiments, Houtgast tested the speech intelligibility in the environment affected both by ambient noise and reverberation [3]. Bradley tested the speech intelligibility in ten classrooms full with students [4]. Ma et al. tested the speech intelligibility in 72 noisy conditions with four different noises including car, babble, train, and street noise [5]. Lavandier and Culling conducted two

experiments to investigate the speech reception thresholds in noise condition with speech-shape in different angles [6]. Rennie et al. tested the speech intelligibility in the condition with noise interfering and quiet [7]. For subjective experiments, Rhebergen et al. designed a subjective experiment to investigate the influence of fluctuating noise on speech intelligibility [8]. Ishikawa et al. conducted a subjective experiment to observe the influence of background noise on speech intelligibility by dysphonia [9]. Bradley et al. investigated the combined influences of SNR and room acoustics parameters on speech intelligibility and concluded that the effect of SNR is much more important [10]. Van Wijngaarden et al. compared the speech intelligibility in noise environment for nonnative and native listeners [11]. Rhebergen et al. predicted the speech intelligibility in real-life background noises including animals, machines, and vehicles [12]. Luts et al. tested the French speech intelligibility in noise conditions [13]. Boon used CRM corpus to measure the speech intelligibility masked by speech-spectrum-shaped noise and concluded the functional relationship between speech intelligibility and SNR [14]. Pollack and Pickett measured the speech intelligibility in noisy environment with high levels and summarized the relationship of speech intelligibility, speech SPL, and SNR [15]. Elliott

measured the perceptual characteristics of noise-masked speech of children aged 9~17 years [16]. Payton et al. compared dry signals and speech signal intelligibility in noise, reverberant, and combined conditions and obtained the differences between the results and objective methods results [17]. Van Wijngaarden compared noise-masked speech intelligibility between native and nonnative Dutch speakers [18]. Zeng et al. investigated the influence of noise and reverberation on the Chinese speech intelligibility of elderly and young subjects [19]. Visentin et al. investigated the effects of different types of noise on speech intelligibility in university classrooms [20]. Duquesnoy measured the speech perceptual threshold for sentences in quiet and noisy environment of elderly and young subjects with normal hearing [21]. Kostić performed a subjective experiment with the MOS method and objective experiment with STOI to obtain the influence of music noise and SNR on Serbian speech intelligibility [22]. Most of the objective experiments aimed to get or improve the objective measuring methods of intelligibility, and the subjective experiments investigated the speech perception in noisy environments with different types and SNRs for different purposes. The objective results should be in accordance with subjective perceptual results, and the relationship between speech intelligibility and SNR should be studied further.

Every Chinese word is a monosyllable, which consists of initials, finals, and tones. Take “zhōng” for example, where “zh” is the initials and “ong” is the finals, while “-” is the tone which represents high flat, and “zhōng” is the syllable which is also the Pinyin of Chinese word. Thus, Chinese speech intelligibility includes initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility which is influenced by initials, finals, and tones intelligibility at the same time.

The present study would introduce a series of experiments of Chinese speech intelligibility in noise transmission systems and analyze the perceptual characteristics of speech intelligibility disturbed by different noises. The changing rules of Chinese speech intelligibility with SNR were analyzed, and the perceptual model of Chinese intelligibility was also concluded.

2. Listening Experiments of Intelligibility

The test speech signals were recorded in an anechoic chamber, and the noise would be added in postprocessing stage. Subjective experiments of speech intelligibility were performed in a listening room according to the National Standard of the People’s Republic of China Acoustics-Speech articulation testing method [23], which introduces the method of measuring and evaluating the quality of a speech transmission system quantitatively and directly. The standard applies to measuring the speech intelligibility of a speech communication system, like concert hall and human-computer interaction.

KXY lists were used for the speech listening test [23]. Every KXY list has 75 words, and the 75 words were divided into 25 groups including 3 syllables without any actual meaning when reading successively. Each group of 3

syllables were recorded after a prompt number. For example, in speech signal “No.R qiè fán yīng,” “R” represents the group number and “qiè fán yīng” are the 3 syllables needed to play back to the subjects. Each list was recorded by two professional announcers.

All the lists were recorded in an anechoic chamber, in which the background noise is less than 20 dB and RT is less than 0.1 s [24]. The speech signals were recorded by two professional announcers (one male and one female) with standard Chinese Mandarin at a standard rate (4 words per second). The speakers’ mouth was kept 10 cm away from the microphone. The SPL was 70 dB near the microphone, and thus the SNR of original speech signals was 50 dB.

In the postprocessing stage, the prerecorded speech signals were added with different types of noises and SNRs according the experiments needed in order to simulate the speech signals masked by noise. The subjective experiments were carried out in a quiet room, and the speech was played back to the listener through a headphone. The experiment procedure is shown in Figure 1.

2.1. Noise Types and SNR. Three types of noises were chosen in the experiments including pink noise, speech noise, and white noise. Pink noise commonly exists in natural environment, where the energy decreases from low frequency to high frequency with 3 dB per octave in logarithmic coordinates. Speech noise as a background noise distributes in the public places like concert, hall, and stadium, where the energy mainly focuses on low frequency. White noise is a broadband noise existing in the communication system, and its energy distributes uniformly in linear coordinate. Their time-domain and frequency-domain waveforms are shown in Figure 2. The three noises introduced above are typical noises in daily life.

Before the experimental parameters were selected, a pilot experiment was performed. In the pilot experiment, white noise was chosen as the background noise, and SNR focused on two extreme ranges of $[-20, 12]$ and $[18, 24]$. The results of the pilot experiment showed that the listeners cannot hear the speech when SNR was below -14 dB while the noise did not affect the listeners in distinguishing the speech when SNR exceeded 20 dB. Therefore, the SNR of speech intelligibility experiments in noise environment was chosen from -14 dB to 20 dB with 2 dB or 4 dB increasing and also including the initials speech recorded in studio without extra noise added with SNR of 50 dB. There were 15 transmission conditions. The information of SNR chosen in the experiments is shown in Table 1.

In order to compare and analyze the results, the SNRs were identical in the three types of noise-interfering environments, and each SNR condition included two speech signals with one male speaker and one female speaker.

2.2. Methods. Listening room and experiment procedure were all the same as those in the experiment described in [25]. If a subject writes down the initials completely same as what he heard, his record was regarded as a correct result; otherwise, considered wrong one, and the ratio of correct

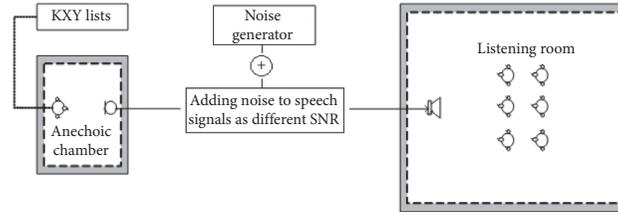


FIGURE 1: Experiment procedure of the subjective listening experiment.

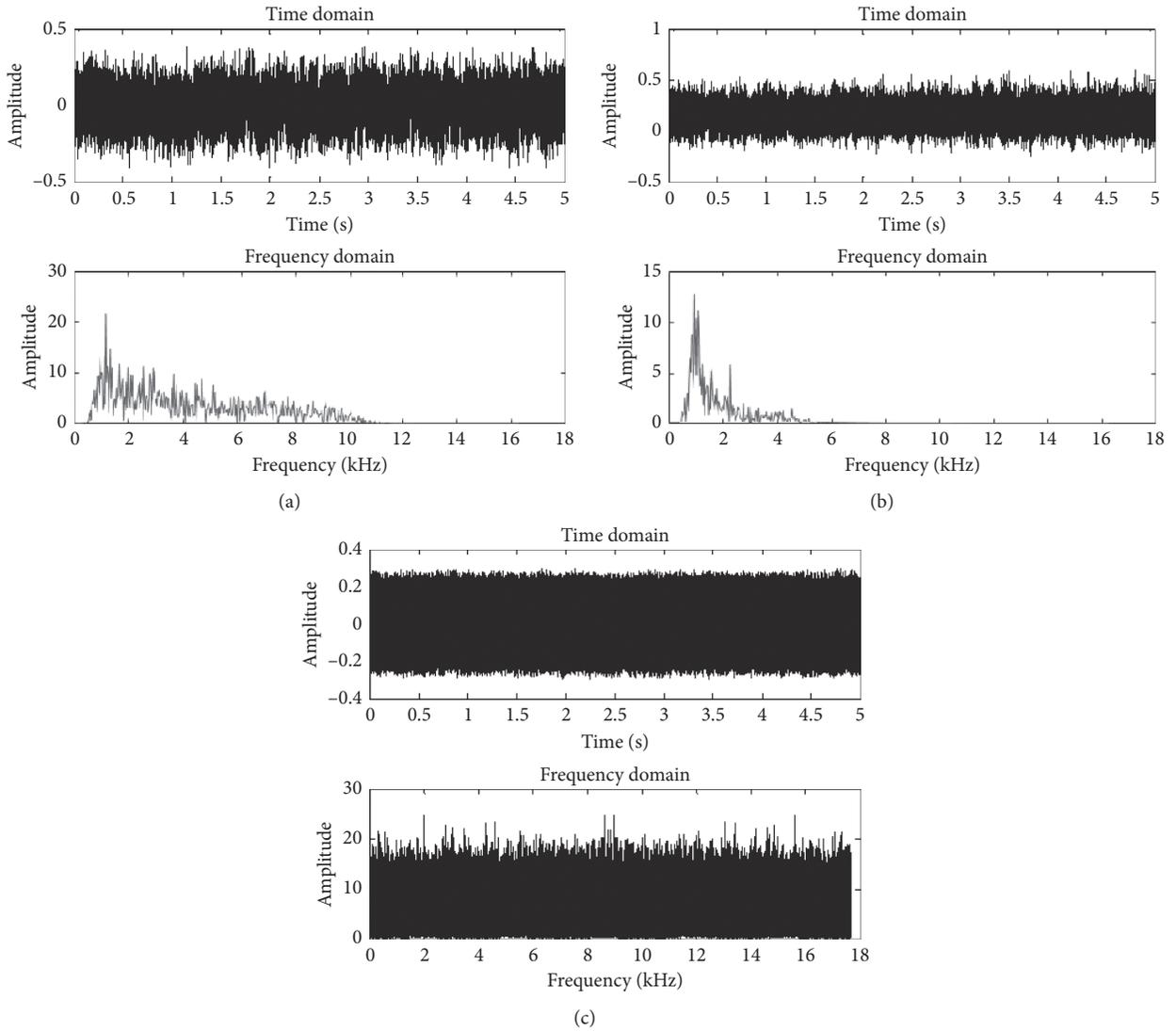


FIGURE 2: Time-domain and frequency-domain properties of noise. (a) Pink noise. (b) Speech noise. (c) White noise.

TABLE 1: SNR of speech intelligibility experiments in noise environment.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
SNR (dB)	-14	-12	-10	-8	-6	-4	-2	0	4	8	12	16	18	20	50

answer was called initials intelligibility score. In the same way, finals intelligibility, tones intelligibility, and syllable intelligibility were also calculated as described in following equations [23]:

$$Q = \sum_{i=1}^n \frac{(N_i/75) \times 100}{n}, \quad (1)$$

$$Q_{sm} = \sum_{i=1}^n \frac{(N_{ism}/75) \times 100}{n}, \quad (2)$$

$$Q_{ym} = \sum_{i=1}^n \frac{(N_{iym}/75) \times 100}{n}, \quad (3)$$

$$Q_{sd} = \sum_{i=1}^n \frac{(N_{isd}/75) \times 100}{n}, \quad (4)$$

where Q , Q_{sm} , Q_{ym} , and Q_{sd} represent the syllable intelligibility, initials intelligibility, finals intelligibility, and tones intelligibility, respectively, N_i is the quantity of corrected perceptual syllable intelligibility by number i listener, while N_{ism} is the quantity of corrected perceptual initials intelligibility by number i listener, while N_{iym} is the quantity of corrected perceptual finals intelligibility by number i listener, while N_{isd} is the quantity of corrected perceptual tones intelligibility by number i listener, and n is the total number of listeners involved in the experiment.

To guarantee the reliability of the experiments, when one listener's score was three times standard deviation different with the average score of all the listeners, his results were wiped out as invalid data. Then, the intelligibility score was recomputed.

2.3. Subjects. There were three types of noises with SNR 15 in the experiments which means 45 acoustic transmission conditions. A total of 90 KXY lists were involved, while every transmission condition needs two KXY lists (a male speaker and a female speaker).

The series of intelligibility experiments were divided to 3 groups according to the noise type:

- (a) There were 14 subjects (7 males and 7 females) who participated in the pink noise-interfering experiment, while the results of 13 subjects (6 males and 7 females) were valid after the standard deviation was tested three times.
- (b) There were 13 subjects (6 males and 7 females) who participated in the speech noise-interfering experiment, while the results of 13 subjects (6 males and 6 females) were valid after the standard deviation was tested three times.
- (c) There were 14 subjects (7 males and 7 females) who participated in white noise-interfering experiment, while all of the 13 subjects' results (6 males and 7 females) were valid after the standard deviation was tested three times.

All of the subjects were undergraduates, aged from 19 to 23 years, and were all native speakers of Chinese Mandarin without any known hearing problems. The subjects were familiar with Chinese Pinyin spelling rules and are experienced in relevant listening experiments. A simple listening training was taken to the subjects before the formal subjective experiments. A total of 92250 ((14 + 14 + 13) × 75 × 30 = 92250) stimulus-response events happened.

3. Results

3.1. Intelligibility of Mandarin in Noise Environment. After data processing, syllable intelligibility, initials intelligibility, finals intelligibility, and tones intelligibility of Mandarin in pink noise, speech noise and white noise-interfering environments were calculated, respectively.

3.1.1. Influence of Gender on Perceptual Intelligibility. Because of the differences of speeches between male and female in fundamental frequency, the bandwidth of frequency, spectrum structure, and intelligibility perception of speech may be different for listeners. Comparisons of intelligibility scores tested by the male speaker and female speaker in pink noise, speech noise, and white noise environments are shown in Figure 3.

Figure 3 shows that the variational trend of intelligibility curves between male and female speakers is almost the same. No significant differences ($F(2,90) = 0.616$, $p > 0.05$) can be found in the speech intelligibility perceptual results between male and female announcers after the ANOVA test. Thus, the Chinese speech intelligibility scores were averaged across the results calculated from two genders under each transmission condition.

3.1.2. Intelligibility of Mandarin in Different Noise Environments. Initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility tested in pink noise environment are shown in Figure 4.

It is shown that the speech intelligibility increases as the SNR increases for all the four curves, and the tones intelligibility scores are the highest, then the finals intelligibility, thirdly the initials intelligibility, and the syllable intelligibility the lowest. The perceptual curves of tones and finals intelligibility have the same trend. Tones intelligibility increases significantly as SNR increases and when SNR is less than -8 dB. Tones intelligibility arrives at 0.9, and the curve is relative stability when SNR exceeds -8 dB. Finals intelligibility has a large increase as SNR increases and when SNR is less than -4 dB, while the relative stability of the finals intelligibility is 0.9 when SNR exceeds -4 dB. The perceptual curves of initials and syllable intelligibility have the same trend. Both two curves increase as SNR increases and when SNR is less than 16 dB, while initials intelligibility stabilizes at 0.9 and syllable intelligibility stabilizes at 0.88 and when SNR exceeds 16 dB.

Initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility tested in speech noise environment are shown in Figure 5.

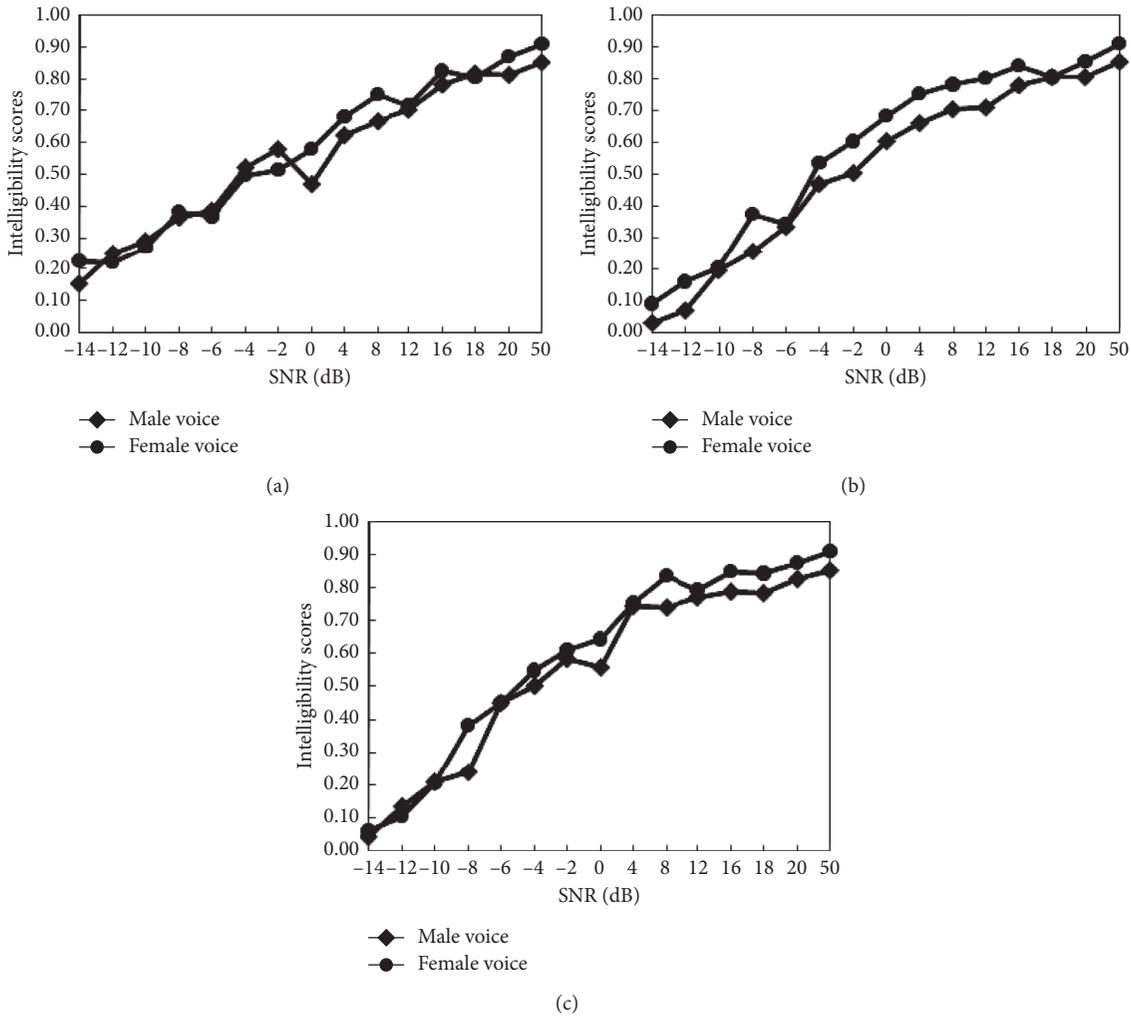


FIGURE 3: The influences of announcers' gender to Chinese speech intelligibility in noise environments. (a) White noise. (b) Pink noise. (c) Speech noise.

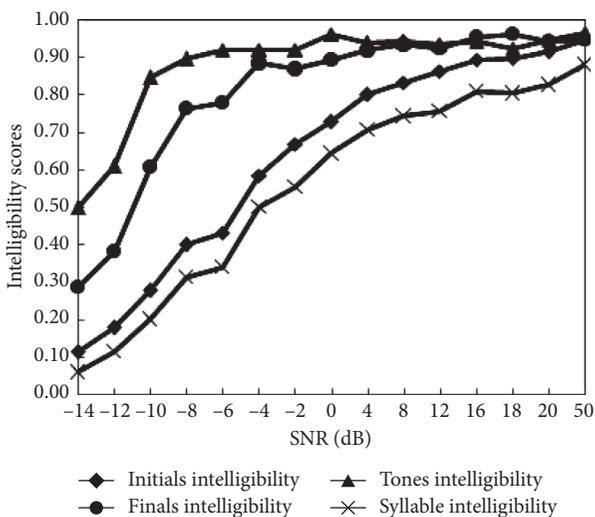


FIGURE 4: Chinese speech intelligibility in pink noise environment.

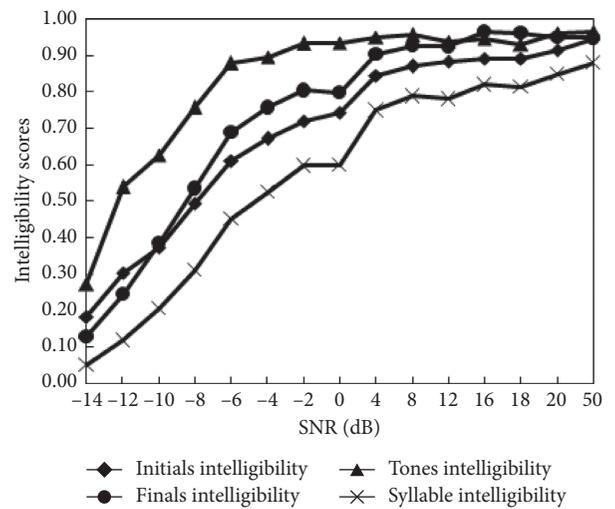


FIGURE 5: Chinese speech intelligibility in speech noise environment.

The perceptual curves of tones intelligibility, finals intelligibility, initials intelligibility, and syllable intelligibility test in speech noise environment have the same trend with the results tested in pink noise environment. Tones intelligibility increases as SNR increases. When SNR reaches -6 dB, tones intelligibility arrives at 0.9 and the curve is relatively stable. Finals intelligibility, initials intelligibility, and syllable intelligibility have a large increase as SNR increases, when SNR is less than 4 dB. When SNR greater than 4 dB, those curves increase slowly and tend towards stability.

Initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility tested in white noise environment are shown in Figure 6.

It is shown that the speech intelligibility increases as the SNR increases for all the four curves, and the results of tones intelligibility is the highest, then the finals intelligibility, thirdly the initials intelligibility, the syllable intelligibility the lowest. The perceptual curves of tones intelligibility remain in a high score (about 0.9) which means white noise has little influence on tones perception. Finals intelligibility increases significantly as SNR increases, when SNR is less than -8 dB, and increases slowly until the curve tends towards stability when SNR exceeds -8 dB. The measuring curves of initials and syllable intelligibility have the same trend. Both of the two curves increase as SNR increases, when SNR is less than 16 dB, and increase slowly when SNR is more than 16 dB. Initials intelligibility is affected most by noise, and syllable intelligibility is mainly determined by initials.

3.1.3. Influence of Noise Type on the Perceptual Intelligibility.

Both temporal and spectrum characteristics of pink noise, speech noise, and white noise are different which lead to different influences on intelligibility perception. The comparison of the influence on initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility by pink noise, speech noise and white noise are shown in Figure 7.

It is shown that the curves of initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility have the same trend when speech signals are interfered by pink noise, speech noise, and white noise. Finals intelligibility and tones intelligibility have a little difference above the three noises especially when SNR is less than -6 dB, because white noise has little influence on finals and tones. Initials intelligibility and syllable intelligibility has no significant difference, and the three curves of syllable almost overlap.

ANOVA on initials intelligibility, finals intelligibility, tones intelligibility, and syllable shows that no significant differences ($F(2,45)=0.211$, $p>0.05$) were observed in initials intelligibility scores above the three noises. The same results also exist in the ANOVA on finals intelligibility ($F(2,45)=1.071$, $p>0.05$), tones intelligibility ($F(2,45)=1.742$, $p>0.05$), and syllable intelligibility ($F(2,45)=0.029$, $p>0.05$). Therefore, Chinese speech intelligibility scores were averaged across all of the data tested in the three noise environments for each SNR.

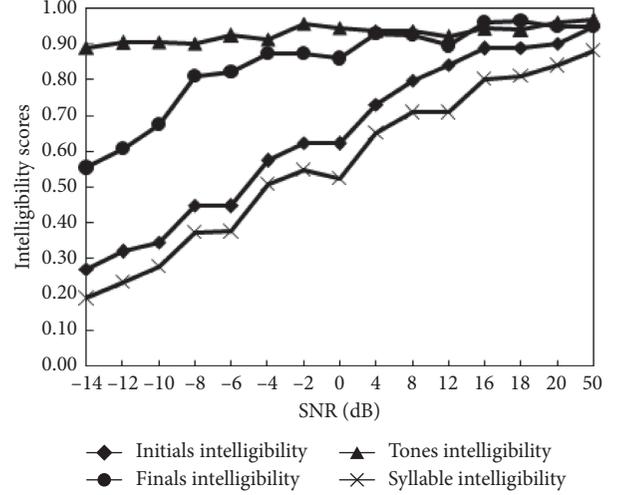


FIGURE 6: Chinese speech intelligibility in white noise environment.

Figure 7(d) shows that Chinese syllable intelligibility scores seem to be stable at a saturating value when SNR exceeds a certain value and the scores cannot reach 1. Even when no extra noise is interfered, perceptual intelligibility is also influenced by some unavoidable factors like vocal and receiving aspects which lead the systematic error in the experiments. The saturation of perceptual intelligibility was calculated after repeating tests on the original speech signals recorded in the anechoic chamber, and the results are shown in Table 2.

3.2. Perceptual Characteristics of Chinese Intelligibility in Noise Environment. Initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility of Mandarin are calculated, respectively, by averaging the results tested in pink noise, speech noise, and white noise conditions, and their scatter diagram is shown in Figure 8.

It can be observed that the trends of speech intelligibility increasing with the SNR obey the exponential law. Speech intelligibility scores were fitted according to the least squares fitting method, and the fitted curves are also plotted in Figure 8. Fitting formulas are shown in equations (5) to (8), and fitting precision is shown in Figure 9.

$$Q_N = -0.33 \times e^{-0.07 \times \text{SNR}} + 0.92, \quad R^2 = 0.99, \quad (5)$$

$$Q_{Nsm} = -0.29 \times e^{-0.08 \times \text{SNR}} + 0.98, \quad R^2 = 0.99, \quad (6)$$

$$Q_{Nym} = -0.09 \times e^{-0.15 \times \text{SNR}} + 0.96, \quad R^2 = 0.99, \quad (7)$$

$$Q_{Nsd} = -0.01 \times e^{-0.25 \times \text{SNR}} + 0.95, \quad R^2 = 0.99, \quad (8)$$

where Q_N represents the syllable intelligibility in noise environment, Q_{Nsm} represents the initials intelligibility, Q_{Nym} represents the finals intelligibility, Q_{Nsd} represents the tones intelligibility, SNR represents the signal to noise ratio, and R^2 represents the fitting precision.

Equations (5) to (8) also provided mathematical models to predict the speech intelligibility of Mandarin according to

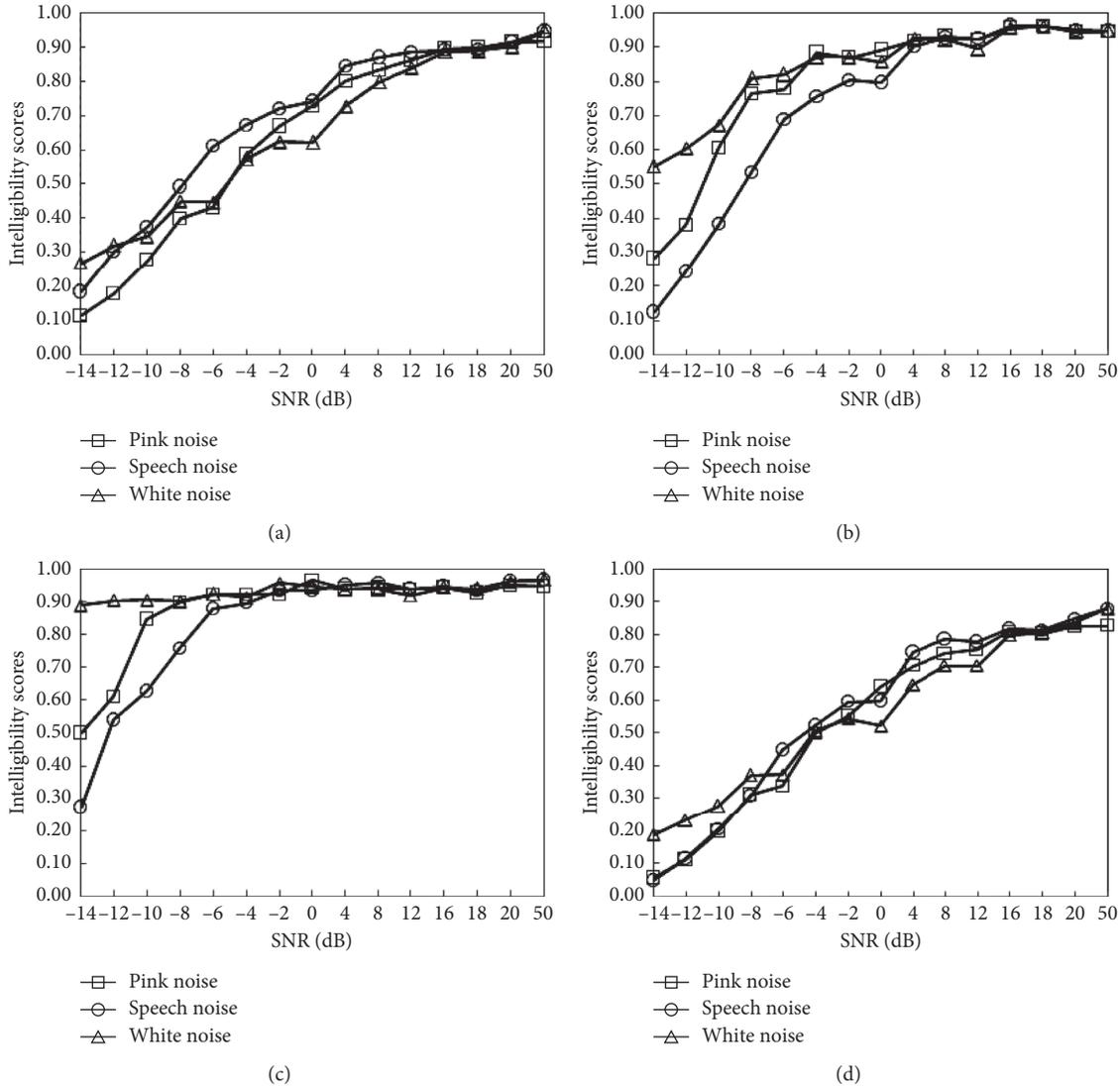


FIGURE 7: Influence of noise type on Chinese speech intelligibility. (a) Initials intelligibility. (b) Finals intelligibility. (c) Tones intelligibility. (d) Syllable intelligibility.

TABLE 2: Saturation of the perceptual intelligibility.

	Perceptual intelligibility
Initials intelligibility	0.93
Finals intelligibility	0.96
Tones intelligibility	0.96
Syllable intelligibility	0.86

SNR in the case of noise interference only and it had high precision with R^2 reaching 0.99. To observe the precision of the predicting models intuitively, a scatter diagram of measuring and predicting values of speech intelligibility was plotted in Figure 9.

Figure 9 shows that all of the points distribute around the central line, and it verifies that the predicting models had high precision on predicting initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility of Chinese Mandarin in noise-interfering environment

which can be named CIPMNE (Chinese Intelligibility Predicting Model in Noise Environment). In the actual application, syllable intelligibility is only thing used to represent the quality of the transmission system, and the CIPMEN can be only defined as a method of predicting the syllable intelligibility.

4. Discussion

There is already a mature objective method of speech intelligibility called STI which is described in IEC 60268 [26]. As the calculation procedure of STI is too complicated, a simplified method of STI called STIPA is widely used in the actual measurement. STIPA can evaluate the transmission quality of speech with respect to intelligibility and has been verified to be effective for western languages. But, the language system of Chinese is different from western languages.

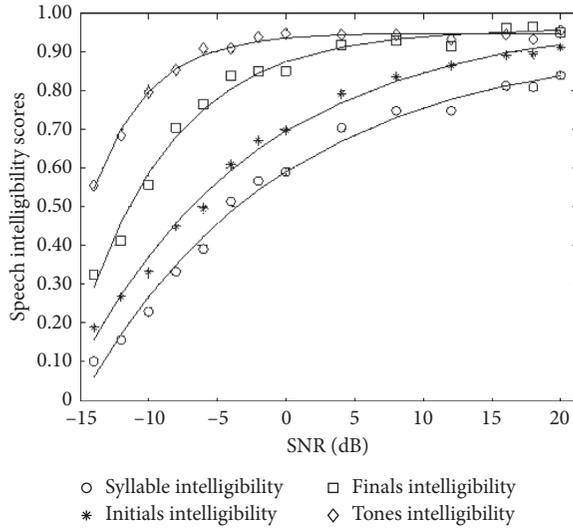


FIGURE 8: Perceptual Characteristics of Chinese intelligibility in noise environment.

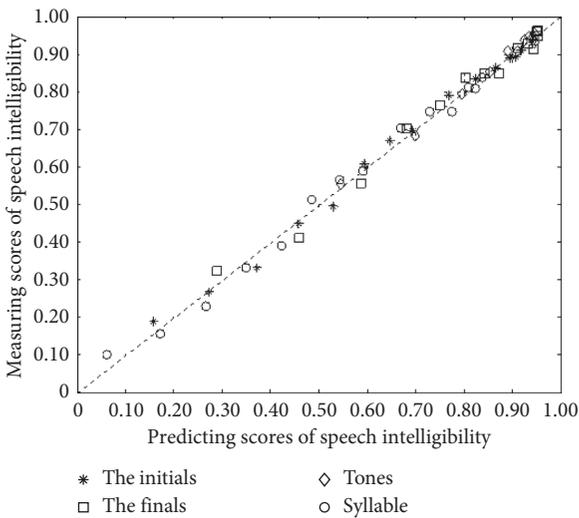


FIGURE 9: Scatter diagram of measuring and predicting values of speech intelligibility in noise environment.

Which method is fit for Chinese, STIPA or CIPMNE, needed to be tested.

Series data were obtained by STIPA and CIPMNE methods, respectively, under the same SNR, and the comparison of the results obtained by the two methods is shown in Figure 10. The differences between the results obtained from the two methods and subjective syllable intelligibility scores are shown in Figure 11.

Figure 10 shows that the results of CIPMNE and STIPA under the same SNR are obviously different. CIPMNE is greater than STIPA when SNR is less than 16 dB and less than STIPA when SNR is more than 16 dB. Especially when SNR is around 0 dB, the difference almost reached 0.3.

It should be noticed in Figure 11 that the curves of CIPMNE and syllable intelligibility are almost have the same trends but different from STIPA. Because CIPMNE is based

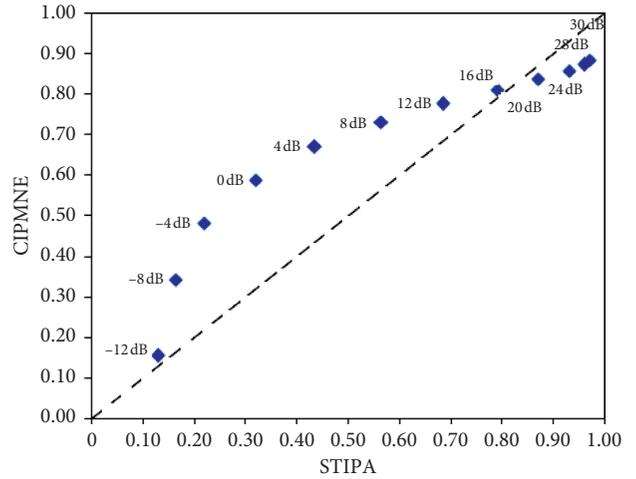


FIGURE 10: Scatter diagram of CIPMNE and STIPA.

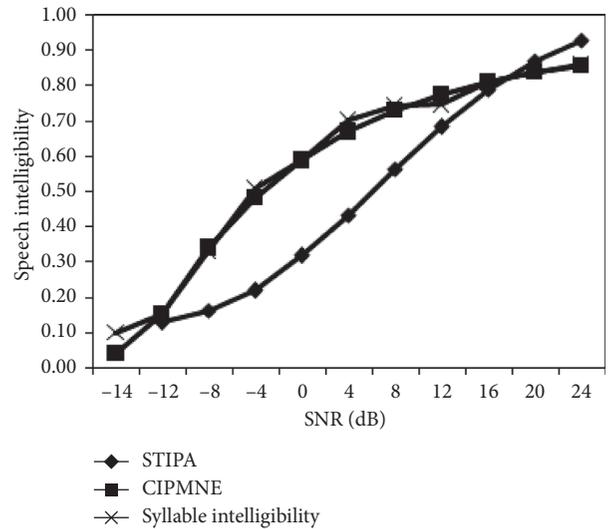


FIGURE 11: Speech intelligibility measured by STIPA, CIPMNE, and subjective listening experiments.

on the perception of syllable intelligibility. The results of speech intelligibility varying with SNR are different when measured by CIPMNE and STIPA, and the results measured by CIPMNE are closer to subjective results. The aim of the researches on the objective measurement method of the speech intelligibility is to approach the subjective perception results as much as possible, and instead of subjective measurement, maybe CIPMNE is fit for Chinese speech intelligibility measurement in the transmission condition of noise-interfering only, or STIPA needs to be revised according to the Chinese speech intelligibility obtained by the subjective listening experiments.

5. Conclusions

The present work introduced a series of Chinese speech intelligibility subjective experiments in the environment deteriorated by white noise, pink noise, and speech noise. Initials intelligibility, finals intelligibility, tones intelligibility,

and syllable intelligibility were obtained, and the changing rules of Chinese speech intelligibility with SNR were analyzed. The results showed that the influences of the three types of noises on perceptual intelligibility have no significant difference. Tones intelligibility scores are the highest, then the finals intelligibility, thirdly the initials intelligibility, and the syllable intelligibility the lowest, which means that the tones have the strongest antinoise property and the property of syllable is the weakest. Speech intelligibility increases with SNR obeying exponential low values and tends to get saturated gradually. The best-fitting exponential function relationships between initials intelligibility, finals intelligibility, tones intelligibility, and syllable intelligibility and SNR, respectively, were established on the basis of the least square method and a model called CIPMNE was proposed based on these functions to evaluate and predict the Chinese speech intelligibility under noise-interfering transmission conditions. A comparison of CIPMEN and STIPA was also made to analyze the differences between the two methods on the Chinese speech intelligibility measurement.

Data Availability

The data used in this research are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by National Nature Science Foundation of China(11204278), Education Department of Jilin Province Research Funds (JJKH20170155KJ), and Doctoral Research Funds of Northeast Electric Power University (BXJXM-2017219).

References

- [1] H. J. M. Steeneken and T. Houtgast, "Basics of the STI measuring method," in *Proceedings of the Past, Present, and Future of the Speech Transmission Index, International Symposium on STI*, pp. 13–44, Soesterberg, Netherlands, October, 2002.
- [2] S. Zhang, H. Song, and Z. Meng, "Relationship between Chinese Mandarin intelligibility and speech transmission index STIPA under simulated transmission conditions," in *IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP)*, IEEE, Chengdu, China, July 2015.
- [3] T. Houtgast, "The effect of ambient noise on speech intelligibility in classrooms," *Applied Acoustics*, vol. 14, no. 1, pp. 15–25, 1981.
- [4] J. S. Bradley, "Speech intelligibility studies in classrooms," *The Journal of the Acoustical Society of America*, vol. 80, no. 3, pp. 846–854, 1986.
- [5] J. Ma, Y. Hu, and P. C. Loizou, "Objective measures for predicting speech intelligibility in noisy conditions based on new band-importance functions," *The Journal of the Acoustical Society of America*, vol. 125, no. 5, pp. 3387–3405, 2009.
- [6] M. Lavandier and J. F. Culling, "Prediction of binaural speech intelligibility against noise in rooms," *The Journal of the Acoustical Society of America*, vol. 127, no. 1, pp. 387–399, 2010.
- [7] J. RENNIES, T. Brand, and B. Kollmeier, "Prediction of the influence of reverberation on binaural speech intelligibility in noise and in quiet," *The Journal of the Acoustical Society of America*, vol. 130, no. 5, pp. 2999–3012, 2011.
- [8] K. S. Rhebergen, N. J. Versfeld, and W. A. Dreschler, "Extended speech intelligibility index for the prediction of the speech reception threshold in fluctuating noise," *The Journal of the Acoustical Society of America*, vol. 120, no. 6, pp. 3988–3997, 2006.
- [9] K. Ishikawa, S. Boyce, L. Kelchner et al., "The effect of background noise on intelligibility of dysphonic speech," *Journal of Speech, Language, and Hearing Research*, vol. 60, no. 7, pp. 1919–1929, 2017.
- [10] J. S. Bradley, R. D. Reich, and S. G. Norcross, "On the combined effects of signal-to-noise ratio and room acoustics on speech intelligibility," *The Journal of the Acoustical Society of America*, vol. 106, no. 4, pp. 1820–1828, 1999.
- [11] S. J. Van Wijngaarden, H. J. M. Steeneken, T. Houtgast, and T. Houtgast, "Quantifying the intelligibility of speech in noise for non-native listeners," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1906–1916, 2002.
- [12] K. S. Rhebergen, N. J. Versfeld, and W. A. Dreschler, "Prediction of the intelligibility for speech in real-life background noises for subjects with normal hearing," *Ear and Hearing*, vol. 29, no. 2, pp. 169–175, 2008.
- [13] H. Luts, E. Boon, J. Wable, and J. Wouters, "FIST: a French sentence test for speech intelligibility in noise," *International Journal of Audiology*, vol. 47, no. 6, pp. 373–374, 2008.
- [14] D. S. Boon, "Evaluation of speech intelligibility with the coordinate response measure," *The Journal of the Acoustical Society of America*, vol. 109, no. 5, pp. 2276–2279, 2001.
- [15] I. Pollack and J. M. Pickett, "Masking of speech by noise at high sound levels," *The Journal of the Acoustical Society of America*, vol. 30, no. 2, pp. 127–130, 1958.
- [16] L. L. Elliott, "Performance of children aged 9 to 17 years on a test of speech intelligibility in noise using sentence material with controlled word predictability," *The Journal of the Acoustical Society of America*, vol. 66, no. 3, pp. 651–653, 1979.
- [17] K. L. Payton, R. M. Uchanski, and L. D. Braida, "Intelligibility of conversational and clear speech in noise and reverberation for listeners with normal and impaired hearing," *The Journal of the Acoustical Society of America*, vol. 95, no. 3, pp. 1581–1592, 1994.
- [18] Van Wijngaarden and J. Sander, "Intelligibility of native and non-native Dutch speech," *Speech Communication*, vol. 35, no. 1–2, pp. 103–113, 2001.
- [19] J. Zeng, J. Peng, and Y. Zhao, "Comparison of speech intelligibility of elderly aged 60–69 years and young adults in the noisy and reverberant environment," *Applied Acoustics*, vol. 159, Article ID 107096, 2020.
- [20] C. Visentin, N. Prodi, F. Cappelletti, S. Torresin, and A. Gasparella, "Speech intelligibility and listening effort in university classrooms for native and non-native Italian listeners," *Building Acoustics*, vol. 26, no. 4, pp. 275–291, 2019.
- [21] A. J. Duquesnoy, "Effect of a single interfering noise or speech source upon the binaural sentence intelligibility of aged persons," *The Journal of the Acoustical Society of America*, vol. 74, no. 3, pp. 739–743, 1983.

- [22] D. Kostić, “The influence of musical noise, type major and minor chord, to the intelligibility of speech in Serbian language,” in *Proceedings of the UNITECH*, Gabrovo, Bulgaria, December 2017.
- [23] GB/T. 15508-1995, *Speech Articulation Testing Method (In Chinese)*, Standards Press, Beijing, China, 1995.
- [24] F. Zhao and B. Shi, “Acoustic processing of anechoic chamber with short reverberation,” *Audio Technique*, vol. 5, pp. 22–24, 2004, in Chinese.
- [25] S. Zhang and H. Song, “Chinese Mandarin speech intelligibility under band-limited transmission conditions,” in *2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, IEEE, Chengdu, China, December 2017.
- [26] IEC 60268-16-2011, “Sound system equipment—part 16: objective rating of speech intelligibility by speech transmission index,” 2011.