Journal of Linguistics and Communication Studies
ISSN 2958-0412
www.paradigmpress.org/JLCS
Volume 5 Number 1 March 2026

The Effects of Noise on English Vowel Perception by Chinese EFL Learners of Different Proficiency Levels

Bingyi Wu[1]

[1] Ocean University of China, Shandong, China
Correspondence: Bingyi Wu, Ocean University of China, Shandong, China.

doi:10.63593/JLCS.2026.03.03

Abstract

Daily communication usually occurs in various kinds of noise backgrounds. Noise interferes with the process of recognizing the target sounds, thus affecting the accuracy of speech perception. In the present study, three speech perception environments were set up: quiet, English babble noise and Chinese babble noise. 60 undergraduate English majors were selected as subjects and divided equally into a higher proficiency level group and a lower proficiency level group. By measuring and calculating their correct identification rates of twenty English vowels in three different listening conditions, the study investigated the effects of second language proficiency and noise type on Chinese English learners' vowel perception. The results showed that learners at the higher level showed a significant advantage over those in the lower level group in identifying English vowels in both quiet and noisy environments, indicating a positive correlation between L2 proficiency and English vowel perception. In contrast, the results of the higher and lower level groups showed different trends in English and Chinese noise. The accuracy of the higher proficiency level group in English babble and Chinese babble are almost the same, while the lower proficiency level group performed significantly better in Chinese babble than in English babble. This suggests that both L2 proficiency and the types of noise have an effect on English vowel perception, and that there is an interactive effect between the two factors.

Keywords: vowel perception, noise, informational masking

1. Introduction

1.1 Research Background

Speech perception is of vital importance in language communication. It not only plays an important role in conveying messages between the speaker and listener, but also contributes to helping L2 learners to develop their second language listening ability and pronunciation. Speech sounds can be divided into vowels and consonants, of which vowels are the nucleus of syllables as well as the direct bearer of major suprasegmental speech flow features such as stress and intonation (Ladefoged, 2006).

Most speech communication in daily life occurs in noisy background. In real life and work environments, various noises interfere with the tracking and recognition of target speech signals at various stages of sound processing, which is called auditory masking of noise (Watson, 2005). The impact of noise on second language communication is particularly obvious. In some noisy places, the perception and understanding of a second language become extremely challenging (Shimizu *et al.*, 2002; Yang & Zhao, 2014; Xu *et al.* 2018). Generally speaking, noise can produce energetic masking and informational masking to interfere with listeners' speech perception. There has been enormous research on the speech perception of L2 learners. However, most of them concentrate on their performance in a quiet environment and ignore the influence of noise, which is much ubiquitous in real life.

1.2 Purpose and Significance of the Study

The present study aims to investigate the English vowel perception in babble noise by Chinese English learners. The first objective of this study is to consider whether there is a difference in L2 learners' vowel perception as their English proficiency improves. The second objective is to explore the effect of informational masking with different masker types on the performance of Chinese English learners of different proficiency levels. Finally, the study attempts to explore whether there is an interactive effect between the two factors, English proficiency and masker intelligibility, in influencing English vowel perception.

This study is significant because speech perception is essential in the process of language communication. First of all, the accuracy of listeners' speech perception not only influences their communication with others, but also plays an important role in L2 speakers' acquisition of the language, especially the acquisition of pronunciation. Besides, among all the speech sounds, vowel perception directly affects the accuracy of pronunciation acquisition and the effect of language communication. Moreover, noise can be everywhere in daily life, so the interference caused by noise should be paid more attention to. In an occupational environment such as airports, speech perception is vital. Workers are required to perceive speech sounds accurately in noisy environments to guarantee safety (Shimizu *et al.*, 2002; Yang & Zhao, 2014; Xu *et al.*, 2018). Finally, this study is meaningful in the comprehension of Chinese EFL learners' speech perception performance in noise. In this way, teachers can be inspired to improve their teaching methods and students can adjust their way to study in order to improve their speech perception ability and further improve their communication ability. Further, the results of this study can be applied in the area of language assessment to enhance the validity and reliability of language listening tests.

2. Literature Review

2.1 Informational Masking of Babble Noise

Informational masking is believed to be caused by the competition between the masker and the target sound in the auditory center for processing resources (Yang *et al.*, 2014). In the exploration of second language speech perception, researchers often use white noise, pink noise, brown noise, speech-shaped noise, babble noise and other noise backgrounds with different signal-to-noise ratios (SNRs) to obtain different levels of informational masking effects. Some studies also incorporate background noise into speech signals to avoid the ceiling effect, which has become a standard method (Winters & O'Brien, 2013; Yang & Zhao, 2014).

Research has been conducted to study the possible factors that may influence the amount of informational masking on speech perception. Mi *et al.* (2003) measured the perception of L2 English vowels in quiet, long-term speech-shaped noise (LTSSN) and multi-talker babble (MTB) by English-native (EN) listeners and Chinese-native listeners in the U.S. (CNU) and China (CNC). In MTB, CNU listeners performed significantly better than CNC, which not only suggested that informational masking played a different role in speech perception by listeners with different backgrounds, but also indicated that exposure to native English may reduce informational masking of MTB. Different types of noise cause a range of various degrees of informational masking, which differ in their masking effects for listeners with different language backgrounds. Jin and Liu (2012) investigated English sentence perception in quiet and two types of maskers, MTB and LTSSN, by English-, Chinese-, and Korean-native listeners. The results found that non-native listeners were more affected by background noise than native listeners. Besides, the masking effects of MTB were greater for Chinese listeners, which indicated that there might be interaction effects between the listeners' native language and competing noise.

2.2 Native and Non-Native Speakers' Speech Perception in Noise

The effect of noise on speech perception is much greater for non-native than for native listeners (Mayo *et al.*, 1997; Rogers *et al.*, 2006; Bradlow & Alexander, 2007; Cooke *et al.*, 2008). The question was first addressed in the laboratory by Black and Hast (1962), nearly half a century ago. They compared word perception scores in quiet and white noise at SNRs of +4, 0, and -4dB. A modest disadvantage for non-native relative to native listeners in quiet conditions of about 11% grew to 16%, 25% and 29% with increasing noise. In Cooke's (2008) study, English and Spanish native speakers were required to recognize keywords in English sentences presented in quiet and in noise, including stationary noise and competing utterances. The study suggested the native advantage in all the listening conditions. Even for bilinguals who started learning the second language at an early age, they were still not comparable with native speakers. The experiments of Rogers *et al.* (2006) confirmed that although the perceptual performance of bilinguals who began to learn a second language at an early age was similar to that of native speakers under quiet conditions, there was still a considerable difference in noise.

Some studies compared native and non-native speakers' performance in quiet, low noise and high noise conditions and found that the size of the native advantage varied in different conditions. While Bradlow and Bent (2002) found that the native advantage remained constant as noise level increased, Garcia Lecumberri and Cooke (2006) demonstrated that native advantage increased in the high noise condition. In the study conducted by Shimizu *et al.* (2002), the researchers found that in the task of English speech perception under quiet conditions, white noise, pink noise and aircraft noise, the speech perception ability of Japanese college students decreased with the decrease of the signal-to-noise ratio of various noise backgrounds.

Moreover, the effects also vary with the noise types (Brungart, 2001; Shimizu *et al.*, 2002; Cutler *et al.*, 2004; Lecumberri & Cooke, 2006; Cooke *et al.*, 2008; Cutler *et al.*, 2008; Broersma & Scharenborg, 2010; Jin & Liu, 2012; Calandruccio *et al.*, 2014). For example, Van Engen (2010) investigated the recognition performance of English sentences by native English speakers and Chinese English learners in the background of two-talker noise (including both English and Chinese conditions) and pointed out that both groups were more affected under English noise conditions.

2.3 Individual Differences in Speech Perception

The performance of L2 speakers in noise is not only different from that of native speakers, but also exhibits a significant individual difference. In most linguistic studies, researchers have focused on two major factors: the age at the onset of second language learning (e.g., Mayo *et al.*, 1997) and second language experience (e.g., Flege *et al.*, 1999; Bradlow & Bent, 2002).

Flege *et al.* (1999) reported that experienced non-native listeners showed better English vowel production and perception than inexperienced non-native listeners. Exposure to a second language environment improves non-native speakers' L2 experience. In Mi's (2003) study, Chinese-native speakers in the U.S. performed better than those in China when they perceived English vowels in multi-talker babble, suggesting the possible effects of second language experience. Studies suggest that non-native speakers who started learning a second language earlier performed better than those who were later, but in most cases, the age of onset and L2 experience have shared effects. Mackay *et al.* (2001) found that when perceiving the first consonant of an English word under noise, Italian native speakers who began to learn English earlier performed better than those who started to learn English later, but there was no difference in their performance when perceiving the last consonant of a word. At the same time, even if the subjects started learning English early, their perception was also affected by the use of their mother language.

Except for the age of onset and L2 experience, researchers are discovering other factors that have potential effects on speech perception in noise. For example, McGowan (2015) measured native listeners' perception and transcription of Chinese-accented English in noise and found there were social factors such as social expectation that caused individual differences in participants' performance. All the participants transcribed more accurately when they were presented with a Chinese face than a Caucasian face. The individual differences in non-native speakers' proficiency have also been considered. Zhou *et al.* (2010) studied the speech perception of twenty RP English vowels by 88 English majors. Each vowel occurred three times in different carrier words in the recording. Participants were required to transcribe all the phonetic symbols they had heard, including consonants and vowels. After the experiment, their answers were rated and only the vowel transcription was scored. Results showed that the mean identification accuracy was 75% among English majors, and their speech perception ability was significantly related to the overall ability of English, oral English proficiency level and gender.

The degree of individual differences varies with subjects and listening conditions. In the study of Shimizu *et al.* (2002), the individual difference in Japanese college students' English speech perception performance was small under quiet conditions, but it became larger as the signal-to-noise ratio increased in noisy background. The subjects with a high L2 proficiency continuum were expected to perform better in noisy conditions, whereas those with low proficiency performed worse.

2.4 Summary

Based on the above review of previous research, it is clear that compared with native listeners, non-native listeners suffer more difficulties from the informational masking effects produced by noise when perceiving English speech sounds. The decline in noise varies depending on the noise types and listeners' individual differences.

However, research concerning second language speech perception in noise focused more on sentence comprehension (Van Engen, 2010; Jin & Liu, 2012), tone recognition (Mao & Xu, 2016) and perception of individual phonetic symbols (Mi *et al.*, 2013). Although most of the previous research investigated sentence recognition and consonant identification in noise, few have studied vowel identification in noise by non-native listeners, in which there are no suprasegmental cues. Moreover, as to the individual difference, researchers mainly focused on the age of arrival and learning experience, and ignored the second language proficiency of Chinese learners. Besides, few researchers have explored the effect of different languages of noise, or combined it with the effect of L2 proficiency to analyze.

Therefore, this study adopted English vowel identification tasks for Chinese EFL learners at higher and lower proficiency levels to test their vowel perception performance in quiet, Chinese babble noise and English babble noise, aiming to investigate the relationship between noise types and listeners' L2 proficiency levels.

3. Methodology

3.1 Research Questions

The study aims to answer the following three questions:

1) What are the effects of second language proficiency in influencing Chinese learners' English vowel perception in noise?
2) What are the effects of masker types in influencing Chinese learners' English vowel perception?
3) What are the interactive effects between L2 proficiency and masker types in influencing Chinese learners' English vowel perception?

3.2 Participants

All the participants were freshmen of English major at the Ocean University of China. The learners have received 9-year compulsory education and thus were of comparable educational background. A total of 67 students (18-20 years) were recruited and asked to take the *Oxford Quick Placement Test* at first. According to their test grades, 60 of them were selected to participate in the following experiment. To ensure the L2 proficiency discrepancy, 7 students were excluded because their grades were at the medium level. Then the 60 participants were averagely divided into two groups, the higher proficiency group (HG) and the lower proficiency group (LG), according to the test results and grades of their latest English exams. The mean test scores of HG and LG were respectively 92.70 and 78.57 out of a maximum 120 points. Before the formal experiment, all the participants were required to get familiar with the 20 English vowels and make sure they were able to recognize the vowels in a short time and make a choice.

Table 1. Participant Background Data Across HP and LP Groups

	HG (N=30)		LG (N=30)	
	Mean	SD	Mean	SD
Age	18.80	0.76	18.61	0.74
Years of English learning	10.00	0.79	9.76	0.70
OQPT grades	92.70	6.09	78.57	4.94

Age is expressed in years; OQPT=*Oxford Quick Placement Test*, scores out of a maximum 120.

3.3 Materials

The test was conducted under three different situations: quiet, four-talker English babble noise (EN) and four-talker Chinese babble noise (CN). Two kinds of noise backgrounds were set. English monolingual and Chinese monolingual recordings of texts were respectively selected from *New Concept English* and HSK textbooks. They respectively included two male voices and two female voices. The speakers were all native speakers of English or Chinese. The content of recordings consisted of both daily conversations and articles concerning various subjects, such as nature, technology and art.

20 Received Pronunciation (RP) English vowels served as speech stimuli. The speech materials used for this study consisted of 60 monosyllabic CNC (Consonant - Nucleus - Consonant) words (eg., *hard*, *fit*, *love)*, in which each vowel occurred three times. British pronunciation of these sixty carrier words was downloaded from the *Longman Dictionary of Contemporary English Online* (LDOCE), produced by a male voice. There was an interval of three seconds between two neighboring words to ensure there was enough time for participants to react and answer. The 60 words were presented in isolation with different sequences to produce three pieces of speech materials. Two of them were respectively mixed with EN and CN through an audio processing software. The duration of each recording was four to five minutes.

Table 2. Twenty English Vowels and Carrier Words Used in the Tests

Vowel	Word 1	Word 2	Word 3
/iː/	leave	heed	feed
/ɪ/	fit	hid	lit
/ɔː/	fort	maul	caught
/ɒ/	toss	lot	hod
/uː/	booth	loose	fool
/ʊ/	hook	book	full
/əː/	birth	hurt	perch
/ə/	was	rubber	ago
/ɑː/	cart	bark	laugh
/ʌ/	cut	love	bus
/e/	check	head	fetch
/æ/	pad	fat	had
/eɪ/	cape	late	bake
/aɪ/	live	fight	hide
/ɔɪ/	boil	join	voice
/ɪə/	beard	dear	fear
/ɛə/	chair	fair	bear
/ʊə/	tour	lure	poor
/aʊ/	house	loud	foul
/əʊ/	quote	rose	load

3.4 Data Collection

The research first selected 60 students after they took the *Oxford Quick Placement Test* and divided them into two proficiency groups based on their test results and the latest English exam scores. Each participant took the test three times under different conditions: quiet, four-talker English babble noise and four-talker Chinese babble noise. Listeners were seated in front of a computer, which presented 20 English vowels as options on the screen. After reading the instructions, they practiced with other words in order to get familiar with the experimental process. In the formal experiment, each of them listened to the three speech materials in random order. After they heard a target word, they were expected to choose which vowel they heard in the word. Each test was expected to cost about 15 to 20 minutes. Between every two words, each of the participants was given enough time to react, make a choice and move on to prepare for the next word. The answers were scored after the tests and the accuracy of vowel identification was calculated.

A pretest was conducted to verify the feasibility of the experiment. To avoid ceiling effects, the signal-to-noise ratios were selected through pilot testing to endure similar identification scores under each condition (in the 50-75% range).

3.5 Data Analysis

After the procedure of the identification experiment, the data collected were analyzed in the following way. First, the identification score was determined by a strict correct answer count. Answers choosing more than one vowel were regarded as false except that for those words which contain two vowels such as *rubber*, participants were required to choose both vowels they had heard. However, only the target vowel was graded. Then the scores were converted to percentages to represent identification rates. Finally, all analyses were carried out using SPSS Version 25.0. A one-way repeated measures ANOVA with L2 proficiency as a between-subject factor and noise type as a within-subject factor was conducted.

4. Results and Discussion

4.1 Vowel Perception by Listeners of Different L2 Proficiency Levels

Table 3. Descriptive Statistics of Mean Identification Rates of HG and LG

	N	Minimum	Maximum	Mean	SD
HG	90	15.00	100.00	67.00	20.48
LG	90	10.00	80.00	53.00	16.73

The mean identification scores by LG and HG groups under three conditions were calculated and listed in Table 3. The results indicated that learners of higher proficiency levels generally performed better than learners of lower proficiency levels in quiet and noise. Overall, listeners with higher L2 proficiency (67% correct) outperformed listeners with lower proficiency (53% correct) significantly ($p<0.01$). Therefore, the vowel identification scores had a strong correlation with L2 proficiency.

Previous studies mainly focused on the comparison of native and non-native speakers. They found that non-native listeners had more difficulty in recognizing speech than native listeners (Shimizu *et al.*, 2002; Bradlow & Bent, 2002; Lecumberri & Cooke, 2006; Pinet *et al.*, 2011), and the disadvantage was much larger in noise than in quiet. Few research has studied the non-native speakers with different proficiency levels, especially Chinese English learners. The present results indicated that L2 proficiency might be one influence factor for non-native speakers' perception. However, whether it makes a difference in each of the three conditions in the present study is still to be analyzed.

Besides, as shown in the table, the SD of both groups were large, then the individual differences within groups were considerable. According to former research, individual differences among non-native speakers were mainly the age of onset and second language experience, which were already controlled in the experiment. All the participants started English learning as a second language at 7-9 years old, and they all did not have ever been exposed to English environment for a long time, for example, living abroad in an English-speaking country. The only criterion for grouping was their grades in OQPT. Their grades were a continuum and to divide the participants into two proficiency groups at one point along the continuum was not as precise as possible. Besides, more various grouping criteria should be adopted to ensure the proficiency discrepancy between the two groups.

Table 4. Descriptive Statistics of Mean Identification Rates in Quiet

	N	Minimum	Maximum	Mean	SD
HG	30	35.00	100.00	73.00	19.23
LG	30	20.00	75.00	54.33	16.49

In quiet, the mean accuracy of the HG and LG group was 73% and 54.33%. Under both conditions, listeners with higher proficiency levels outperformed listeners with low proficiency levels significantly ($p<0.01$). This is consistent with the overall situation above. However, the overall identification accuracy was not as high as expected. Although the percentages were slightly higher in quiet, the situation was not optimistic. In previous research, the mean identification accuracy in a quiet environment by L2 learners was at least over 75% (Van Engen, 2010; Zhou *et al.*, 2010; Xu *et al.*, 2018), while in the present study, the accuracy was much lower with 73% by higher proficiency level learners as the highest. There may be two reasons to explain that. First, the participants are all freshmen in the university and their English competence does not achieve a relatively high level. Second, although all the participants are guaranteed to have known all the English vowels and be able to recognize them before the experiment, some students were still not very familiar with them at the time of data collection. Therefore, under the pressure of taking tests, it would probably be challenging for them to make the best choice during the experiment process.

4.2 Vowel Perception Under Different Types of Noise

Table 5. Descriptive Statistics of Mean Identification Rates in Quiet and Noise

	N	Minimum	Maximum	Mean	SD
Quiet	60	20.00	100.00	63.67	20.10
EN	60	10.00	100.00	56.25	18.99
CN	60	15.00	100.00	60.08	20.30

The mean identification rates by participants in quiet, EN, and CN were respectively 63.67%, 56.25%, and 60.08%. Data showed that due to the masking effects of noise, participants performed worse in two types of noisy environments than in quiet. What was interesting about the data in the table was that the mean identification rate in

Chinese noise was slightly higher than that in English noise, though the difference was not significant (*p*=0.288). It meant that there may be effects of noise types, i.e., the languages of maskers, in influencing participants' different performances, but whether the situation is the same for the higher or lower proficiency group is still to be analyzed. The results were partly in agreement with Van Engen's (2010) study, in which 20 native English speakers and 20 non-native English speakers whose native language was Chinese were recruited to do second-language recognition test in both English and Chinese 2-talker babble noise. Van Engen found that both groups of participants experienced greater difficulty in English noise than in Chinese noise, and this was consistent with the present study. The greater interference from English noise for both groups suggested that acoustic and/or linguistic similarity between the speech signal and the noise might be the most essential factor in driving noise language effects.

However, Garcia Lecumberri and Cooke (2006) did not observe the different effects of L1 and L2 noise types. They investigated the performance of Spanish speakers on English consonant perception respectively in Spanish and English competing speech and found that there was no significant differences in the masking effects of the two languages. The disagreement in findings may be accounted by that their study differed from the present study and Van Engen's study in that the former used single talker's speech as the noise while the latter respectively used 4-talker babble and 2-talker babble. Since the signal density increased, greater energetic masking could be induced, and then renders informational masking effects more observable (Van Engen, 2010).

4.3 Interplay Between L2 Proficiency and Masker Types

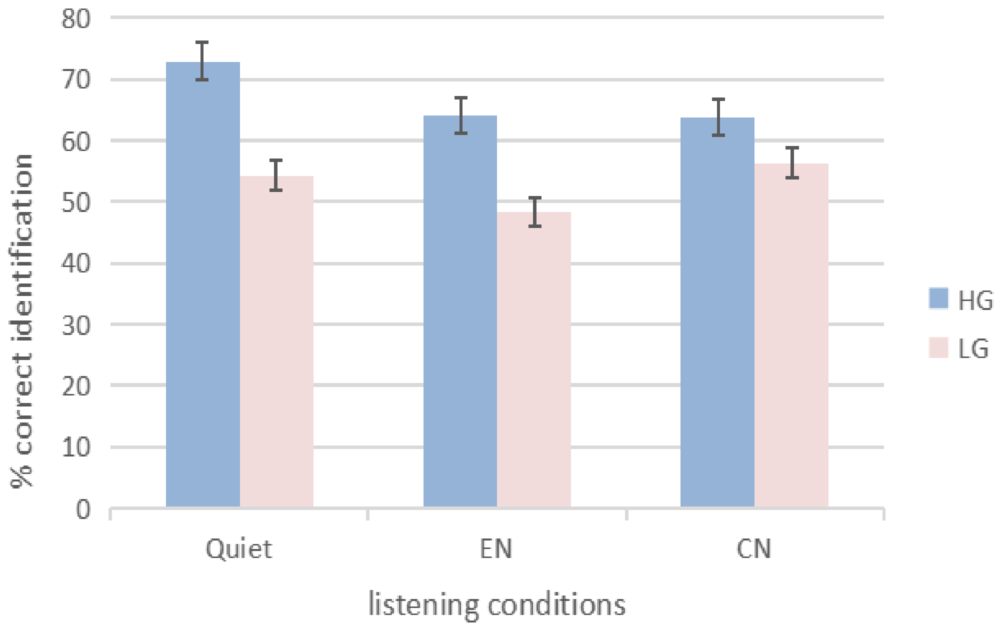

Figure 1. Mean Identification Rates by HG and LG in Three Conditions

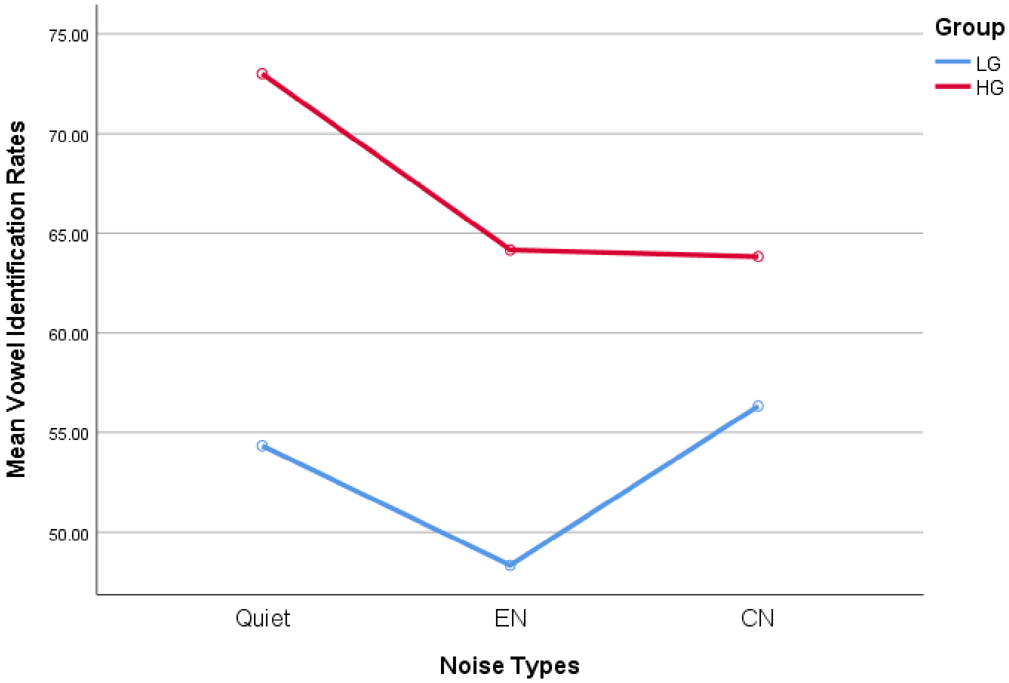

Figure 2. Changes of Mean Identification Rates Across Three Conditions

A repeated-measures ANOVA with L2 proficiency as a between-subject factor and noise type as a within-subject factor revealed that the effect of noise type was significant (F(2,57)=5.411, p<0.01, η²=0.16). There was also an interaction effect between noise type and L2 proficiency (F(2,57)=3.168, p≤0.05, η²=0.1). The main effect of L2 proficiency was significant (F(1,58)=12.362, p<0.01, η²=0.176).

Since there was an interaction effect between L2 proficiency levels and noise types, a simple effect analysis was performed after correction by the Bonferroni method. Results showed that in quiet and in EN, the simple effects of L2 proficiency were significant (F(1,58)=16.284, p<0.01, η²=0.219; F(1,58)=12.448, p<0.01, η²=0.177), while in CN, it was not significant (F(1,58)=2.084, p=0.154, η²=0.035). Pairwise comparisons indicated that among HG learners, their mean identification score in quiet was the highest and showed strong advantages over CN and EN environments (p<0.05). Though they performed slightly better in EN than in CN, data showed no significant difference between the two noise types (p>0.05). However, among LG learners, their identification score in CN was significantly higher than that in EN (p<0.05), and there was no significant difference between that in quiet and in EN, or in CN.

4.3.1 The Effects of L2 Proficiency

Table 6. Descriptive Statistics of Mean Identification Rates in Noise

	N	Minimum	Maximum	Mean	SD
HG	60	15.00	100.00	64.00	20.58
LG	60	10.00	80.00	52.33	16.96

In the two types of noisy environments, the mean vowel identification rates by HG and LG were respectively 64% and 52.33%. Compared with those in quiet, there was an apparent decrease. Likewise, participants in HG outperformed the others significantly, which meant that in noise, second language proficiency still had an impact on participants' vowel perception. However, on account that two different types of maskers were employed in the study, it should be analyzed separately whether L2 proficiency played a role in each noisy environment.

Table 7. Descriptive Statistics of Mean Identification Rates in EN

	N	Minimum	Maximum	Mean	SD
HG	30	35.00	100.00	64.16	18.15
LG	30	10.00	65.00	48.33	16.57

In English noise, the mean identification rates of the HG and LG group were 64.17% and 48.33%.

Participants of higher proficiency levels generally got significantly higher scores than those of lower proficiency levels. Thus, it could be concluded that both in quiet and English noise, L2 proficiency would influence the English vowel perception of Chinese learners. The more proficient participants were at the second language, the better they performed in speech perception.

Table 8. Descriptive Statistics of Mean Identification Rates in CN

	N	Minimum	Maximum	Mean	SD
HG	30	15.00	100.00	63.83	23.06
LG	30	30.00	80.00	56.33	16.65

On the contrary, in Chinese noise, the mean vowel identification rates of the HG and LG groups were respectively 63.83% and 56.33%. There was no significant difference between the two groups (p=0.154). This indicated that though L2 proficiency played an important role in influencing listeners' performance in quiet and in English noise, it seemed to have no impact on English vowel perception by Chinese EFL learners when there was Chinese babble noise.

4.3.2 The Effects of Noise Types

Table 9. Descriptive Statistics of Mean Identification Rates of LG

	N	Minimum	Maximum	Mean	SD
Quiet	30	20.00	75.00	54.33	16.49
EN	30	10.00	65.00	48.33	16.57
CN	30	30.00	80.00	56.33	16.65

The vowel identification rates by the LG participants in quiet, English noise, and Chinese noise were respectively 54.33%, 48.33%, and 56.33%. Performance in quiet, which was similar to performance in CN, did not have apparent advantages over the two other conditions. It was due to their poor performance in quiet environment, which was probably caused by participants' unfamiliarity with the English vowels. There was a significant correlation between vowel identification accuracy and noise types among listeners with lower proficiency levels (p<0.01). This indicated that LG learners were more likely to be affected by environment and noise types. As noted in the previous chapter, Van Engen (2010) investigated native English speakers and Chinese native speakers and found that both groups suffered more from the interference by English noise than Chinese noise. It was worth noting that Chinese speakers experienced a smaller masking release in Chinese noise relative to English noise. In other words, compared with English native speakers, non-native speakers did not show relatively significant advantages in L2 perception in Chinese noise. Lecumberri *et al.* (2010) have tried to explain this by proposing the multiple effects of informational masking. On the one hand, due to the intelligibility of masking sounds, listeners faced a great challenge when processing the target sounds and masking sounds simultaneously. On the other hand, if listeners were familiar with the masking sounds, they would suffer greater interference. As Chinese speakers were more familiar with their mother tongue than L2 (English), it might be more challenging for them when the masking sounds were Chinese.

In the present study, participants were all EFL learners, so they were assumed to be more familiar with Chinese than with English. Therefore, they were expected to suffer greater interference in Chinese noise rather than in English noise. However, the current results suggested that participants showed strong advantages in Chinese noise compared with English noise. As in English babble noise, the speech materials and maskers were the same languages. Identifying the target sounds required more cognitive resources when there was a similarity between target sounds and masking sounds, so English noise would cause more interference for listeners when perceiving English vowels. In conclusion, for participants of lower L2 proficiency levels in this study, the effects of typological similarity between the target sounds and masking sounds played a relatively dominant role over listeners' degree of familiarity with noise.

Table 10. Descriptive Statistics of Mean Identification Rates of HG

	N	Minimum	Maximum	Mean	SD
Quiet	30	35.00	100.00	73.00	19.23
EN	30	35.00	100.00	64.16	18.15

CN	30	15.00	100.00	63.83	23.06

The vowel identification rates by the HG participants in quiet, English noise, and Chinese noise were respectively 73%, 64.17%, and 63.83%. Participants performed best in quiet, and significantly better than in noise (p<0.01), while in contrast to the results of participants at lower proficiency levels, there was no significant difference between EN and CN environments (p=0.951). The results suggested that although the overall data showed a relative disadvantage in EN, learners of HG seemed to be less susceptible to the influence of different masker types, or languages of noise. The findings were in agreement with the study of Garcia Lecumberri and Cooke (2006), who found no apparent differences in Spanish speakers' perception of English consonant in Spanish and English noise. Besides, Van Engen and Bradlow (2007) used 6-talker babble and also found no effects of languages of noises. On the contrary, many other studies have reported that there were effects of noise types. Van Engen (2010), for example, claimed that listeners' language experience and signal similarity would modulate the interference they suffered when they perceived speech in different noises and further contribute to different performances. One possible reason that may account for the different outcomes was the different speech perception tasks adopted. Van Engen (2010) investigated keyword recognition in sentences, while the present study and Garcia Lecumberri and Cooke (2006) measured identification of individual phones, vowels and consonants respectively. Then it seemed that, when participants listened to sentence-length materials, they were more likely to be interfered by noise in a specific language. As participants tried to identify keywords in sentences rather than consonants, more linguistic structures needed to be processed, so processing inefficiencies across levels of linguistic processing (Cutler *et al.*, 2004) would accumulate for non-native listeners. Moreover, in the perception of individual phones in isolation, it did not involve much cerebral activity to process and comprehend the target sounds and masking sounds.

In conclusion, L2 proficiency and noise types respectively played a dominant role in certain conditions. Results showed that for the non-native listeners at lower proficiency levels, interference from a 4-talker masker in the target language (English) was greater than interference from the listeners' native language (Chinese). This finding suggests that signal similarity (a match between target and noise languages) is at least as important as second language proficiency in driving noise language effects in general. For Chinese EFL learners of higher proficiency levels, their familiarity with their native language would offset the effects of signal similarity so that no significant difference in English versus Chinese noise was observed. These results supplemented and expanded the findings of Van Engen (2010) and Lecumberri *et al.* (2006) to a certain degree, in that similarity and familiarity would differ in their degree of influences on non-native speakers of different proficiency levels.

5. Conclusion

5.1 Major Findings of the Study

The primary goal of this study is to examine the effects of noise types and second language proficiency in influencing Chinese EFL learners' perception of English vowels. Subjects' English abilities were measured at first. Then they participated in three vowel identification tests, respectively in quiet, English babble noise and Chinese babble noise. After the experiment, participants' answers were scored and analyzed. The experimental results prove that noise does have masking effects on English vowel perception by Chinese EFL learners. The effect varies in degree across different kinds of noise types and different L2 proficiency levels. Based on the three research questions, the research findings of the present study are as follows:

Listeners at higher L2 proficiency levels perceive English vowels better than listeners at lower L2 proficiency levels. In quiet, English babble noise and Chinese babble noise, the higher proficiency learners all possess the advantages.

The effects of noise are different for listeners at higher proficiency and lower proficiency levels. For the former, the presence of noise does have an impact on vowel perception, but changing the noise types will not cause much difference to the interference. However, for the latter, English noise is proven to be more challenging for them to resist than Chinese noise.

There are interactive effects between the two factors, L2 proficiency, and noise types. If the listeners are at a high proficiency level, noise types will have little effect because as previous research suggests, advanced English learners are less likely to be influenced by noisy environments. However, if the listeners are at a lower proficiency level, the change in noise types may cause greater influence. As it has been difficult for them to identify the vowels compared with advanced learners, the interference caused by English noise, which is similar to the target sounds, requires much more cognitive resources to process the information.

5.2 Implications of the Study

According to the results and discussion above, some practical implications of the present study are addressed briefly. Firstly, this study has shed light on some notable problems in non-natives' English vowel perception, especially in unfavorable listening conditions such as noise. Speech perception is the premise of speech acquisition and speech production. It determines learners' L2 pronunciation acquisition. In the current study, the accuracy of speech perception is proven to be influenced by L2 proficiency. This further improves that individual differences, including L2 experience, age of arrival and L2 proficiency, are internal factors that will influence learners' speech perception.

Then, the ability to deal with different noisy environments when perceiving English vowels also differs across L2 proficiency groups. Thus, in a real teaching situation, learners with different proficiency levels should adopt relevant measures, such as controlling the listening conditions and using group instruction, to guarantee teaching quality and results.

Furthermore, from a more clinical perspective, the findings in this study can be extended to some research on subjects with hearing disorders, including patients with hearing degradation or impairment and so on.

5.3 Limitations and Suggestions for Future Studies

Due to the restriction of actual conditions, the present study has limitations and needs to be improved.

First, the total number of participants is sixty. They are divided into two groups and thus, each group consists of thirty subjects. This is a relatively small number for such an experiment which requires participants to listen and judge, so subjectivity is a potential factor to influence the results. Future studies should increase the number of subjects in order to minimize the effect of individual differences.

Second, as mentioned above, the overall accuracy is not ideal. This is owing to the generally low second language proficiency level of participants recruited in the present study. Besides, the only criterion for grouping is the grades of OQPT and their latest English exam. This may lead to irrationality and one-sidedness to some extent. Future studies should consider strictly selecting subjects and taking various measures to determine the differences among their proficiency levels.

Third, the present study focuses on second language learners' vowel perception in multi-talker babble noise. In real studying and working conditions, there may be various kinds of noise. Thus, more noise types should be considered in future studies. Whether there is a difference in identification accuracy across the 20 vowels also should be further investigated.

References

Bent, T., & Atagi, E. (2015). Children' s perception of nonnative-accented sentences in noise and quiet. *The Journal of the Acoustical Society of America, 138*(6), 3985–3993.

Black, J. W., & Hast, M. H. (1962). Speech Reception with Altering Signal. *Journal of Speech Language and Hearing Research, 5*(1), 70–75.

Bradlow, A. R., & Alexander, J. A. (2007). Semantic and phonetic enhancements for speech-in-noise recognition by native and non-native listeners. *The Journal of the Acoustical Society of America, 121*(4), 2339–2349.

Bradlow, A. R., & Bent, T. (2002). The clear speech effect for non-native listeners. *The Journal of the Acoustical Society of America, 112*(1), 272–284.

Broersma, M., & Scharenborg, O. (2010). Native and non-native listeners' perception of English consonants in different types of noise. *Speech Communication, 52*(11-12), 980–995.

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America, 109*(3), 1101–1109.

Calandruccio, L., Bradlow, A. R., & Dhar, S. (2014). Speech-on-speech Masking with Variable Access to the Linguistic Content of the Masker Speech for Native and Nonnative English Speakers. *Journal of the American*

Academy of Audiology, 25(4), 355–366.

Cooke, M., Garcia Lecumberri, M. L., & Barker, J. (2008). The foreign language cocktail party problem: Energetic and informational masking effects in non-native speech perception. *The Journal of the Acoustical Society of America, 123*(1), 414–427.

Cutler, A., Garcia Lecumberri, M. L., & Cooke, M. (2008). Consonant identification in noise by native and non-native listeners: Effects of local context. *The Journal of the Acoustical Society of America, 124*(2), 1264–1268.

Cutler, A., Weber, A., Smits, R., & Cooper, N. (2004). Patterns of English phoneme confusions by native and non-native listeners. *The Journal of the Acoustical Society of America, 116*(6), 3668–3678.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *The Journal of the Acoustical Society of America, 106*(5), 2973–2987.

Hillenbrand, J. M., Clark, M. J., & Houde, R. A. (2000). Some effects of duration on vowel recognition. *The Journal of the Acoustical Society of America, 108*(6), 3013–3022.

Ji, X., Zhang, H., Li, A. & Gong, J. (2018). An empirical study of English intonation perception among learners at different proficiency levels. *Foreign Language Teaching and Research, 50*(03), 393–406+480–481.

Jin, S.-H., & Liu, C. (2012). English sentence recognition in speech-shaped noise and multi-talker babble for English-, Chinese-, and Korean-native listeners. *The Journal of the Acoustical Society of America, 132*(5), EL391–EL397.

Ladefoged, P. (2006). *A Course in Phonetics (5th Ed.)*. Boston: Thomson Wadsworth.

Lecumberri, M. L. G., & Cooke, M. (2006). Effect of masker type on native and non-native consonant perception in noise. *The Journal of the Acoustical Society of America, 119*(4), 2445–2454.

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication, 52*(11-12), 864–886.

Lewis, D., Hoover, B., Choi, S., & Stelmachowicz, P. (2010). Relationship Between Speech Perception in Noise and Phonological Awareness Skills for Children with Normal Hearing. *Ear and Hearing, 31*(6), 761–768.

Lin, S. (2011). The Effect of Age on the English Speech Perception and Production. *Journal of Tibet University, 26*(03), 175–178.

Mackay, I. R. A., Meador, D., & Flege, J. E. (2001). The Identification of English Consonants by Native Speakers of Italian. *Phonetica, 58*(1-2), 103–125.

Mao, Y., & Xu, L. (2016). Lexical tone recognition in noise in normal-hearing children and prelingually deafened children with cochlear implants. *International Journal of Audiology, 56*(2), 23–30.

Masuda, H. (2016). Misperception patterns of American English consonants by Japanese listeners in reverberant and noisy environments. *Speech Communication, 79*, 74–87.

Mayo, L. H., Florentine, M., & Buus, S. (1997). Age of Second-Language Acquisition and Perception of Speech in Noise. *Journal of Speech Language and Hearing Research, 40*(3), 686–693.

McGowan, K. B. (2015). Social Expectation Improves Speech Perception in Noise. *Language and Speech, 58*(4), 502–521.

Mi, L., Tao, S., Wang, W., Dong, Q., Jin, S.-H., & Liu, C. (2013). English vowel identification in long-term speech-shaped noise and multi-talker babble for English and Chinese listeners. *The Journal of the Acoustical Society of America, 133*(5), EL391–EL397.

Pinet, M., Iverson, P., & Huckvale, M. (2011). Second-language experience and speech-in-noise recognition: Effects of talker-listener accent similarity. *The Journal of the Acoustical Society of America, 130*(3), 1653–1662.

Ren, H. (2022). L2 Speech perception and its Correlation with Production：Evidence from Speech Acquisition of Japanese Gemination by Chinese Learners. *Foreign Language Learning Theory and Practice*, (03), 115–127+162.

Rogers, C. L., Lister, J. J., Febo, D. M., Besing, J. M., & Abrams, H. B. (2006). Effects of bilingualism, noise, and reverberation on speech perception by listeners with normal hearing. *Applied Psycholinguistics, 27*(03), 465-485.

Shimizu, T., Makishima, K., Yoshida, M., &

Yamagishi, H. (2002). Effect of background noise on perception of English speech for Japanese listeners. *Auris Nasus Larynx, 29*(2), 121–125.

Strori, D., Bradlow, A. R., & Souza, P. E. (2020). Recognition of foreign-accented speech in noise: The interplay between talker intelligibility and linguistic structure. *The Journal of the Acoustical Society of America, 147*(6), 3765–3782.

Van Dommelen, W. A., & Hazan, V. (2010). Perception of English consonants in noise by native and Norwegian listeners. *Speech Communication, 52*, 968–979.

Van Engen, K. J. (2010). Similarity and familiarity: Second language sentence recognition in first- and second-language multi-talker babble. *Speech Communication, 52*(11-12), 943–953.

Van Engen, K. J., & Bradlow, A. R. (2007). Sentence recognition in native- and foreign-language multi-talker background noise. *The Journal of the Acoustical Society of America, 121*(1), 519–526.

Van Wijngaarden, S. J., Steeneken, H. J. M., & Houtgast, T. (2002). Quantifying the intelligibility of speech in noise for non-native listeners. *The Journal of the Acoustical Society of America, 111*(4), 1906–1916.

Watson, C. S. (2005). Some comments on informational masking. *Acta Acustica United with Acustica, 91*(3), 502–512.

Weiss, W., & Dempsey, J. J. (2008). Performance of bilingual speakers on the English and Spanish versions of the hearing in noise test (HINT). *Journal of the American Academy of Audiology, 19*, 5–17.

Winters, S., & O'Brien, M. G. (2013). Perceived accentedness and intelligibility: The relative contributions of F0 and duration. *Speech Communication, 55*(3), 486–507.

Xu, C., Yang, X., Wang, Y., Zhang, H., Ding, H. & Liu, C. (2018). Informational Masking of Six-talker Babble on Mandarin Chinese Vowel and Tone Identification: Comparison between Native Chinese and Korean Listeners. *Studies of Psychology and Behavior, 16*(01), 22–30.

Yang, X. & Zhao, Y. (2014). The effects of background noise on second-language speech perception. *Advances in Psychological Science, 22*(06), 934–942.

Yang, X., Zhao, Y. & Wang, Y. (2018). Differences in duration processing in the production and perception of English /e/–/æ/ by Chinese college students. *Foreign Languages and Their Teaching*, (02), 79–89+149.

Yang, X., Zhao, Y., Xu, L., Tang, W., Zhu, X., Bao, A. & Zhao, Z. (2017). Effects of Noise on English Listening Comprehension among Chinese College Students. *Contemporary Foreign Languages Studies*, (01), 12–17+109.

Yang, Z., Zhang, T., Song, Y. & Li, L. (2014). Subcomponents of auditory informational masking: Evidence from behavioral and neuroimaging studies. *Advances in Psychological Science, 22*(03), 400–408.

Zhou, W., Shao, P. & Chen, H. (2010). An empirical study of RP English vowel perception among English-major university students. *Foreign Languages Bimonthly, 33*(06), 45–49+128.