

ATTENTION, PERCEPTION, AND PRODUCTION OF THE ENGLISH VOICELESS
INTERDENTAL FRICATIVE BY CHINESE LEARNERS OF ENGLISH

By

Lianye Zhu

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Teaching English to Speakers of Other Languages—Master of Arts

2014

ABSTRACT

ATTENTION, PERCEPTION, AND PRODUCTION OF THE ENGLISH VOICELESS INTERDENTAL FRICATIVE BY CHINESE LEARNERS OF ENGLISH

By

Lianye Zhu

This study investigated the relations among speaking styles, Chinese ESL learners' attention to the voiceless interdental fricative ([θ]) sound, the learners' perception of English native speaker's speech, learners' self-perception, and the learners' production. Thirty-four Chinese ESL learners participated in three production tasks and two perception tasks. A stimulated recall, questionnaire, and interview were also conducted. The results of both quantitative and qualitative studies showed that a) the production accuracy of [θ], and the amount of attention paid to it, were positively related to the difficulties of the speech styles; b) the more demanding the speech style was and the less salient the word position was, the less attention was paid to the production of [θ] and the less accurately it was produced; c) the participants produced [θ] better than they perceived it; d) the participants' NS-perception was better than their self-perception.

ACKNOWLEDGMENTS

First, I would like to express my sincere gratitude to my academic adviser, Dr. Debra M. Hardison, who provided me with very helpful and thoughtful ideas and advice during this study and the composition of this thesis.

Second, I would also want to thank all the participants and raters participated in this study, without whose cooperation many interesting findings in this study would never be noticed.

Third, my great appreciation goes to my loving parents and my boyfriend, Zheng Fan, who have been supporting my research and bringing sunshine into my life.

TABLE OF CONTENTS

LIST OF TABLES	vi
LIST OF FIGURES	viii
KEY TO ABBREVIATIONS	ix
INTRODUCTION	1
CHAPTER 1	2
LITERATURE REVIEW	2
Inaccurate [θ] Sound Produced by Chinese Learners of English	2
Interlanguage Variation	3
Attention and Monitoring	6
Distraction	8
Production and Perception	9
CHAPTER 2	11
RESEARCH METHOD	11
Participants	11
Materials and Procedures	11
Data Analysis	14
CHAPTER 3	17
QUANTITATIVE ANALYSIS	17
The Effect of Speaking Style on Speaking Production of [θ]	17
The Effect of Word Position on the Production of [θ] in Word List Reading	18
The Effect of Speaking Styles and Word Position on Production of [θ]	19
The Effect of Auditory Distraction on [θ] in Word List Reading	21
The Effect of Auditory Distraction and Word Position on Production Accuracy of [θ]	22
The Difference in Accuracy between Self-Perception and Perception of NS	24
Relationship between Production and Perception	25
CHAPTER 4	28
QUALITATIVE ANALYSIS	28
The Relationship between the Participants' Attention to [θ] in Various Word Positions and Speaking Styles	28
Factors Influencing the Participants' Attention to Production	31
Relationship between Attention and Monitoring of Speech Production	33

Relationship between Monitoring Strategies and Distraction.....	35
Factor Influencing the Participants' Attention to Perception.....	35
Relationship between Production and Perception	38
CHAPTER 5.....	40
RESEARCH DISCUSSION	40
What is the effect of speaking style, word position and auditory distraction on production accuracy of [θ]? Do these factors interact? If yes, how?	40
Speaking Style	40
Word Position.....	41
Interaction between Speaking Style and Word Position	42
Auditory Distraction.....	43
Interaction between Auditory Distraction and Word Position	45
Is there a relationship between L2 learners' perception and production?	45
What is the effect of distraction on participants' monitoring strategies?	46
CHAPTER 6.....	47
CONCLUSION AND IMPLICATION	47
APPENDICES.....	52
APPENDIX A: WORD LIST READ BY THE ENGLISH NATIVE SPEAKER	53
APPENDIX B: PICTURE PROMPTS FOR TELLING THE STORY	
"THREE LITTLE PIGS"	55
APPENDIX C: THE SECOND WORD LIST READ BY THE PARTICIPANTS	58
APPENDIX D: THE FIRST WORD LIST READ BY THE PARTICIPANTS	60
APPENDIX E: QUESTIONNAIRE ABOUT THE PARTICIPANTS' ATTITUDE AND ATTENTION PAID TO VOICELESS INTERDENTAL FRICATIVES	62
APPENDIX F: THE MAIN INTERVIEW QUESTIONS	65
REFERENCES	67

LIST OF TABLES

Table 1: The Procedures of the Study	14
Table 2: The Rating Arrangement.....	15
Table 3: Descriptive Statistics for Inaccuracy Percentage of [θ] and Speaking Style	17
Table 4: Paired-Samples <i>t</i> Test Between the Inaccuracy Production of [θ] in Two Speaking Styles.....	18
Table 5: Descriptive Statistics for the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Word List Reading (Without Distraction)	19
Table 6: Friedman Test of the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Word List Reading (Without Distraction).....	19
Table 7: Descriptive Statistics for the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Story Telling	20
Table 8: Wilcoxon Test of the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Word List Reading and Story Telling	21
Table 9: Descriptive Statistics of Inaccuracy Percentage of [θ] in Two Word List Readings	21
Table 10: Paired-Samples <i>t</i> Test Comparing the Inaccuracy Production of [θ] in Two Word List Reading Tasks	22
Table 11: Descriptive Statistics for the Inaccuracy Percentage of [θ] in Initial, Medial, and Final Word Positions in the Two Word List Readings	23
Table 12: Wilcoxon Test for Inaccuracy Percentage of [θ] in Word Initial, Medial, and Final Positions in the Two Word List Reading Tasks	24
Table 13: Descriptive Statistics for the Inaccuracy Percentage of the Participants' Perception of NS' Speech and Self-perception of Word List Reading	24
Table 14: Paired Sample <i>t</i> Test for Inaccuracy Percentage of [θ] Perception in the Word List Reading (without distraction) and the NS's Word List Reading	25
Table 15: Descriptive Statistics for the Inaccuracy Percentage of [θ] Production and	

NS-Perception.....	25
Table 16: The Correlation between the Inaccuracy Percentage of [θ] in Participants' Speaking Production and NS Perception (NS)	26
Table 17: Descriptive Statistics for the Inaccuracy Percentage of Participants' Self-perception of [θ]	26
Table 18: The Correlation between the Inaccuracy Percentage of [θ] in Participants' Speaking Production and Self-Perception (NNS)	27
Table 19: Descriptive Statistics of the Questionnaire Results (Item 7-Item 12).....	28
Table 20: Reliability Analysis of the Questionnaire	29
Table 21: Difference between Attention Paid to Initial, Medial, and Final [θ] in Word List Reading without Distraction and Story Telling.....	30
Table 22: Questionnaire about the Participants' Attitude and Attention Paid to [θ].....	63

LIST OF FIGURES

Figure 1: Picture Prompts for Telling the Story “Three Little Pigs”	56
---	----

KEY TO ABBREVIATIONS

1. ELC = English language center
2. ESL = English as the second language
3. IL = interlanguage
4. L2 = the second language
5. L1 = the first language
6. LOR = length of residence
7. NS(s) = English native speaker(s)
8. NNS(s) = English non-native speaker(s)
9. OT = Optimality Theory
10. SLA = second language acquisition

INTRODUCTION

For most of the learners of English, the foreign accent has always been a difficulty that could not be easily overcome. Numerous research studies have been carried out to study the reasons for this issue as well as the methods to help learners to produce spoken English close to native-like, if not to completely eliminate the foreign accents. Among the difficulties learners encounter in acquiring the second language (L2) native-like speaking production, the voiceless interdental fricative (i.e., [θ]) has been considered one of the most difficult sounds to acquire by most of the learners of English. Many studies have identified various substitutions for [θ] in the speaking production of learners from different first language (L1) backgrounds. For example, Rau, Chang, and Tarone (2009) reported that Thai, Russian, and Hungarian learners of English tended to substitute [t] for [θ] while [θ] was usually replaced with [s] by speakers from Asian countries such as Japan, Korea, and China (Lee & Cho, 2002; Rau et al., 2009). In order to probe the causes of these problematic performances, one plausible way is to seek the relation between learners' perception and production in that it is generally believed that there is a positive correlation between a speaker's perception and production so that the improvement of one part will facilitate the development of another. However, the perception alone by no means determines the production. Yang (1997) indicated that speakers' attention, an important factor involved in one's cognitive process of speaking activity, also influences a speaker's perception and production. The following literature review introduces related theories and previous studies that contributed to this research topic.

CHAPTER 1

LITERATURE REVIEW

Inaccurate [θ] Sound Produced by Chinese Learners of English

Numerous problematic phonetic substitutions have been observed in the study of English voiceless interdental fricatives produced by Chinese learners, evidencing that this phoneme is mostly replaced with [f], [s], or [t]. For example, Deterding (2006) analyzed the pronunciation of thirteen young Mandarin Chinese speakers through recording their passage reading and short interviewing. He determined that [θ] was mostly replaced by [s], which confirmed the results of Hung's (2005) study. Similar substitution errors were also found by Cheng and He (2008), Chen and Bi (2008). The value of these two studies is their research methods. Cheng and He (2008) developed a four-year longitudinal study of English pronunciation of 14 English major university students in mainland China to observe how participants' English pronunciation improved. Instead of relying on English native-speakers' (NS) judgment, the correctness of participants' pronunciation was acoustically analyzed via PRAAT, an articulation analysis software so that inaccurate pronunciations were further analyzed through the comparison of phonetic parameters between sounds produced by participants and NSs. Chen and Bi (2008) did their research not through recruiting participants to record speaking productions but by analyzing an English speaking production corpus consisting of speech samples by 200 Chinese university students (50 English major students and 150 non-English major students). The huge amount of data collected from a quite variable controlled group increased the generalizability of the study.

However, regardless of the different research methods applied in these studies, they shared one limitation -- the researchers attributed the difficulty of acquiring the target [θ] sound to the difference between the L1 (i.e., Chinese) and L2 (i.e., English). They claimed that the lack of the equivalence of [θ] in Chinese led to Chinese speakers' inaccurate production of that sound (Gao, 2002; Wu, 2008). Basing the analysis and discussion on Contrastive Analysis only might have overlooked other factors and may fail to explain other findings also observed in the studies. For example, Rau and Chang (2009) discovered that Chinese speakers performed differently under different circumstances. Specifically, the accuracy rate of the interdental fricative produced by them was higher in formal speaking than that in casual speaking. Such findings as interlanguage (IL) variation could not be completely explained by Contrastive Analysis. In order to compensate for its limitation, interlanguage variation, which could be traced back to speaking style (Labov, 1966) and Optimality Theory (OT) (Prince & Smolensky, 1993) has received increasing attention in the study of second language acquisition (SLA).

Interlanguage Variation

Interlanguage variation has been studied from two perspective points of view which are OT and speaking style shifting. OT, proposed by Prince and Smolensky (1993), is constraint-based and output-oriented (Hsu, 2013). Specifically, it states that the phonological output is the speakers' optimal choice out of all the potential candidates with markedness and faithfulness being concerned. Advocates of an OT model proposed that learners with different L1 backgrounds produced English interdental fricatives differently

because their optimal choice of target sounds (i.e., interlanguage) was influenced by their constraint rankings of the L1. Besides, such output is variable because learners would gradually re-rank their linguistic constraints, which would eventually be the same as the constraints of the L2 (Lee, 2006; Lombardi, 2003; Wester, 2007; Yildiz, 2002). The detailed research of the OT model is beyond the scope of this study because the emphasis will be placed on speaking style shifting.

"Speaking style" was first analyzed by Labov (1966) and extensively studied both in L1 and L2 acquisition. Later with the increasing interest in IL, speaking style was believed to significantly influence learners' IL (Dickerson & Dickerson, 1977) as well. In general, style shifting is caused by the change of speakers' attention paid to speaking (Labov, 1966, 1970). Based on this principle, Tarone (1982, 1983, 1988) drew a continuum of style shifting with vernacular speaking and careful speaking at each pole respectively and proposed that the more vernacular speech was, the less accurate it would be. However, Labov and Tarone's theory about "style shifting" was challenged by Dowd et al. (1990) and Major (2001) because it was hard to determine the boundary between two adjacent speaking styles. In other words, how many different variables should two speech samples have so that they could be identified as two different speaking styles? In addition, both Tarone (1979) and Labov (1984, p. 30) admitted that it was challenging to observe "real vernacular speaking" to which speakers paid very limited attention because of the presence of researchers, the speakers' awareness of participating in the research, and the application of the sound recorders. In other words, the "observer's paradox" was inevitable.

To specify the definition of style shifting, some researchers proposed that it was the amount of attention paid to speaking that determined the accuracy of speaking (Ellis, 1994; J. Hulstijn & W. Hulstijn, 1984; Major, 2001). Variationists proposed that the more formal the speaking style was, the more attention would be allocated to pronunciation, and therefore the more accurate the pronunciation would be (Dowd, Zuengler, & Berkowitz, 1990; Major, 2001). Commonly, the formality² of four speaking styles was studied. From the most formal to the least formal, they were: word-list reading, paragraph reading, picture describing, and free talking or interviewing (Thompson & Brown, 2012). The English interdental fricative, because of its variation, has always been a popular target sound in the study of the relationship between the formality of speaking and its variation. For example, as early as 1977, Schmidt (1977) investigated the English interdental fricatives produced by Egyptian Arabic speakers and discovered that participants were more likely to produce [θ] in more formal speaking. Similarly, Coyne (2008) studied the English interdental fricatives produced by Cajun people whose L1 was French in word list reading and paragraph reading and discovered that the participants substituted [t] for [θ] less often in the word list reading task. Therefore, Coyne concluded that higher formality level might lead to higher accuracy of speaking production. However, such correlation between formality and production accuracy in Coyne's study might not be generalizable because of two limitations. First, only four participants were involved in the study. Second, it might be questioned if word list reading had a significantly higher level of formality than the paragraph reading task. A more generalizable study was done

² Because the concept of "formality" is controversial, in the current study, the variable is referred to as speaking style; style and task are used interchangeably.

by Rau, Chang, and Tarone (2009) in which they investigated 27 Chinese university students' production of the [θ] sound under four circumstances with different levels of formality. Their results also indicated a positive correlation between the formality of the speaking and the accuracy of the target sounds' pronunciation.

However, the negative evidence about correlation between accuracy and formality was also found in Thompson and Brown's research (2012) in which they observed the speaking production of a Spanish learner of English and found that she pronounced most accurately not in reading minimal pairs but in reading passages. Coyne (2008) also discovered that one of her participants had the highest accuracy in the passage reading task. These unexpected findings indicated that the attention allocated to pronunciation might not necessarily be related to the form of speaking but other factors may also influence speakers' pronunciation as well.

Attention and Monitoring

Monitoring during speaking was viewed as production-based (Kormos, 1999) because it functioned as a "mental eye" through which speakers held control of their utterance (Berg, 1986). Because the Perceptual Loop Theory (Levelt, 1989, 1993) posited that speaking production was parsed into three steps, namely, pre-articulation, articulation, and post-articulation, some researchers claimed that speaking monitoring should also be studied in these three steps respectively (Kormos, 2000). The pre-articulatory monitoring was studied by Baars et al. (1975) and Motley et al. (1982). They believed that speakers would monitor the accuracy of their utterances before the speaking was articulated. This

hypothesis led to a further assumption that inaccurate pronunciation was caused either by the lack of attention or by the failure of monitoring. The former was related to the allocation of attention consciously or subconsciously controlled by speakers (de Bot, 1992) while the latter was caused by self-perception errors.

Kormos (2000) reviewed the previous studies (Tarone, 1983; Tarone & Parrish, 1988) and summarized that the accuracy of speaker's speech production was influenced by the amount of attention the speaker paid to it. In order to identify the amount of attention allocated to different aspects of speaking such as the lexicon, semantics, and/or phonetics, she recruited 40 Hungarian learners of English to accomplish an information-gap role play and a retrospective interview afterwards. She regarded the instances of self-repair as the signal of the existence of attention. Her analysis showed that attention was paid first and foremost to lexicon and then to grammar. She believed that such a priority hierarchy was shared by learners of all the levels. Although speakers' attention paid to pronunciation was not considered in Kormos' study, it had been taken into consideration in Wheeldon and Levelt's study (1995) five years earlier.

Wheeldon and Levelt (1995) focused on how speakers monitored phonological encoding. They asked the Dutch participants to silently translate the L2 English stimulated words they heard into their L1 Dutch and during the translation, to press the button whenever they noticed that they encountered the target Dutch phoneme which they were required to monitor. A comparison between the participants' response time when the target Dutch phoneme was in the word initial position with the response time when it was in other word positions showed that the participants monitored word initial

phonemes significantly faster than other phonemes. Wheeldon and Levelt's findings illustrated that monitoring was influenced by word position of the target segment. Such implication was evidenced by Rau et al. (2009)'s study in which Chinese learners of English were observed to monitor their words' initial interdental fricative the most.

Distraction

Two predominant methods of studying participants' internal attention are thinking aloud and stimulated recall. However, since it is impossible for participants to think aloud when they are speaking and the details participants provide during the stimulated recall might potentially be incomplete or incorrect due to limited short-term memory, an alternative research method is needed. Though it is probably hard to guarantee that participants pay as much attention to L2 data as researchers presume, it may be safe to assume that participants will pay less attention to L2 data if they are distracted by other tasks. Based on previous research, Zeamer and Fox Tree (2013) posited that auditory distraction would cost people extra cognitive effort and shift their attention from the focal task to the distractions. Al-Hejin (2005) further noted the concepts of attention in SLA that the more demanding the task was, the more attention people needed to pay to it. These studies implied that if distracted by another task, participants would pay less attention to the primary task.

Because of the salient influence that distraction has on attention, it will be included in the study as an independent variable which might differentiate participants' performance.

Production and Perception

Sometimes speakers may not be able to identify production errors even though they are carefully monitoring their pronunciation. In other words, it is not the lack of attention but the failure of self-perception that leads to inaccurate production. In fact, the relation between perception and production has been studied by many researchers. One popular belief is that perception influences production. For example, Brannen (2011) analyzed the perception and production of English interdental fricatives by participants from Japan, Russia, France, and Québec Canada (French speaker). He observed a relationship between participants' perception and production of target sounds and posited that the improvement of perception could facilitate the development of production. Similarly, Fu (2011) discovered a positive relation between Taiwanese ESL learners' production and perception of interdental fricatives.

However, opposite results were also found by other researchers such as Lee (2011) who found no relation between a group of advanced Korean EFL learners' perception and production of interdental fricatives. Both Syed (2013) and Owolabi (2012) found that learners had difficulty only in producing but not perceiving English interdental fricatives.

Another issue in studying the relation between perception and production is that not much attention has been paid to learners' self-perception. In early research, usually it was assumed that learners' ability to perceive their own production was the same as that to perceive the production of native speakers of the target language. Therefore, it might be insightful to see if there is any difference between these two types of perception.

To conclude this literature review, the research questions of this study are listed below.

1. What is the effect of speaking style, word position and auditory distraction on production accuracy of [θ]? Do these factors interact? If yes, how?
2. What is the relationship between L2 learners' perception and production of [θ]?
3. What are the roles of monitoring, attention, and distraction in L2 learners' perception and production of [θ]?

The researcher hypothesized that the speaking style, the word position, and the auditory distraction could affect the production accuracy of [θ]. Specifically, first, the more formal the speaking style is, the more accurate the production will be. Second, the participants might produce the [θ] in the word initial position the best. Third, the auditory distraction will influence the production accuracy of the [θ] probably because it affects the participants' monitoring of the production or decreases the amount of attention paid to the production. In addition, the relationship between L2 learners' perception and production of [θ] could be expected.

CHAPTER 2

RESEARCH METHOD

Participants

In this study, thirty-four Chinese graduate students (25 female and 9 male) enrolled at a large university in the United States were recruited through the university mail-list that most of the Chinese students at that university joined. Their first language was Mandarin. When they participated in this study, their length of residence (LOR) in the United States was approximately from six months to three years with the mean as 13.24 months (SD = 9.43 months). Coincidentally, about two-thirds of the participants were engineering students who had comparatively limited opportunities to speak English with English native speakers (NSs).

Five NSs participated as raters. Four of them were ESL instructors at the university which the participants attended. Each had at least two years of ESL teaching experience. The fifth rater was a program coordinator who frequently communicated with international graduate students.

Materials and Procedures

Before asking the participants to complete the experiment tasks, the researcher asked a rater to read 15 groups of the minimal pairs with two to three words in each group (41 words in total) for the perception test in this study. The minimal pairs were adopted from Rau et al. (2009)'s study and [θ]s were included in each pair (see Appendix A). Her voice was recorded into a SONY Recording Pen and was later edited by the software Gold Wave so that the volume was amplified while the noise was eliminated.

The experiment tasks had three phases: the production test, the perception test, and the interview. Each phase was further divided into several steps.

First, the participants were required to tell the story *The Three Little Pigs*, which was adopted from Rau et al. (2009)'s study, based on eight picture prompts (see Appendix B). They were given one minute to plan the story and they were expected to include as much information as indicated from the pictures in their stories.

Second, the participants were asked to listen to the recordings of their story telling and to recall whether they paid attention to [θ] when they produced it. If they paid attention to the target [θ], they were also required to explain their strategies for monitoring the production accuracy of this sound. The stimulated recall was recorded through a recording pen.

Third, half of the participants were provided with a word list containing 37 words divided into fifteen groups of minimal pairs, all of which included the [θ] sound. This word list was adopted from Rau et al. (2009)'s study (see Appendix C). A piece of Chinese news was also selected from a China Central Television (CCTV) News Report and edited through Gold Wave to reduce the background noise. The CCTV News Report was chosen because the reporter's voice was clear and speaking rate was moderate. In addition, the language spoken by the reporter was standard Mandarin. The participants were required to listen to this news report played in the headphones while reading aloud the minimal pairs on the word list at a normal speed. They were asked to remember as many details of the news report as they could because immediately after they listened to the news, they were expected to repeat it to the digital recording pen.

Fourth, that group of the participants was told to read another 15 groups of minimal pairs, which included two to three words in each group (44 words in total, see Appendix D). Like the first word list, [θ] was also involved in each minimal pair. This time, no additional task was required. The other half of the participants were required to go through the fourth step first and then the third step. The counterbalanced experiment could increase the reliability of the test and avoid the test effect.

Fifth, the participants were asked to complete a 12-item questionnaire (see Appendix E). Half of the items addressed their overall attitude towards accented English spoken by Chinese learners of English, especially concerning [θ] and the other half concerned their attention paid to [θ] during story telling and word list reading.

Sixth, the participants were required to listen to 15 groups of minimal pairs, which had two to three words in each group, read by the fifth rater and to write down whatever words they heard. They were allowed to use IPA symbols, which was shown to them for reference, if they did not know how to spell the words. After that, they were asked to listen to all the words read by themselves during step three and step four. They were also required to write down the words they heard. In order to reduce the influence of participants' memory for these minimal pairs, they were not told that they were listening to their own recordings.

Last, a one-to-one interview was conducted between the participant and the researcher in Chinese because it was the first language of the interviewer and the interviewee. The major questions discussed during the interview are shown in Appendix F.

All the participants' speaking productions were recorded using a SONY Recording Pen

and were stored on a private password-protected laptop for data analysis. The

procedures of the study are summarized in Table 1.

Table 1.

The Procedures of the Study

Steps	Methods of Collecting Data	Materials
Step One: Story telling	Recording	Appendix B
Step Two: Stimulated recall	Recording	Recordings of Story telling
Step Three: -Word list reading with auditory distraction - News retelling	Recording	Appendix C, headphones
Step Four: Word list reading without auditory distraction*	Recording	Appendix D
Step Five: Questionnaire	Writing down answers	Appendix E
Step Six: other-perception, self-perception	Writing down the words perceived	Recordings of the word list reading
Step Seven: Interview	Recording	Appendix F

*Step three and four were performed in reverse order by half of the participants

Data Analysis

In order to answer the research questions, both quantitative and qualitative data were analyzed. The first step in the quantitative data analysis was the evaluation of the participants' production and perception. Four ELC teaching assistants, who were NSs of English, were assigned to evaluate the participants' speaking production. Each rater was responsible for evaluating 17 participants' productions of [θ] in their story telling and word list reading. The rating arrangement was carefully designed so that each participant's production could be rated by two raters. Table 2 illustrates which rater was assigned to evaluate which participants.

Table 2

The Rating Arrangement

Raters	Participants		Raters	Participants	
Rater 1	1~10	18~24	Rater 3	11~17	25~34
Rater 2	1~10	11~17	Rater 4	18~24	25~34

The raters were provided with the transcriptions of the participants' story telling and the two word lists the participants read. The raters were required to accomplish two tasks: first, to identify the [θ] sounds that were mispronounced as other sounds by the participants; second, to spot the sounds that were mispronounced as [θ] by the participants. The raters worked separately to record the rating results on the rating sheets. After comparing the rating results of these four raters, the fifth rater was called in to judge any discrepancies between each pair of raters. The final results were based on a consensus reached by at least two raters.

The researcher investigated the results of the participants' perception tests through two steps: first, the researcher compared the words produced by the NS and the corresponding words written down by the participants; second, the researcher compared the words the participants were supposed to produce, the words the raters determined that they had actually produced, and the corresponding words the participants wrote down during the perception tests. Each discrepancy was counted as one perception error and the percentages of inaccuracy were calculated through dividing the number of erroneous words by the total number of the words.

Based on the ratings provided by the raters, the researcher further categorized the erroneous words, which were either produced or perceived incorrectly, into three groups according to where the [θ] was positioned in the word: initial, medial, or final. For each

group, the percentage of erroneous words was also calculated.

In addition, the researcher transcribed the interviews and stimulated recalls for a qualitative study. The quantitative study was conducted with the help of SPSS while the qualitative study was accomplished through interpreting the interviews, questionnaires, and stimulated recalls.

CHAPTER 3

QUANTITATIVE ANALYSIS

The quantitative analysis was conducted by using SPSS 19.0. The data were analyzed to answer each research question and the analysis results are presented in this chapter.

The Effect of Speaking Style on Speaking Production of [θ]

In order to investigate the possible influence of speaking style on the participants' production of [θ], the percentages of mispronounced sounds related to [θ] out of the total number of words were calculated. Participants' inaccuracy percentage was used in this study because the participants produced different numbers of words in story telling and therefore calculating the percentage was more reliable than counting the number of words containing erroneous [θ]. Table 3 shows the descriptive statistics for the inaccuracy percentage of [θ] produced in three speaking styles. It was anticipated that the inaccuracy percentage of [θ] produced in the word list reading could be lower than that produced in the story telling.

Table 3

Descriptive Statistics for Inaccuracy Percentage of [θ] and Speaking Style

	Mean	Standard Deviation	Lower Bound	Upper Bound
Word list reading without distraction	0.09	0.01	0.06	0.12
Story telling	0.18	0.02	0.15	0.21

A paired-samples *t* test was conducted to investigate if there was a significant difference between the inaccuracy percentages of [θ] produced in different speaking styles (i.e., word list reading and story telling). The results of paired-samples *t* test in

Table 4 reveals a significant difference between the inaccuracy percentage of [θ] produced in the participants' word list reading (without news) and in the story telling, $t(33) = -5.34, p < .001, r = .68$.

Table 4

Paired-Samples t Test Between the Inaccuracy Production of [θ] in Two Speaking Styles

		Paired Differences					t	df	Sig.(2-tailed)
		Mean	Std. Devia- tion	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	Word List Reading (without distraction) – Story Telling	-.09193 82	.10048 46	.0172330	-.12699 90	-.05687 75	-5.335	33	.000

The Effect of Word Position on the Production of [θ] in Word List Reading

Because many previous studies concluded that the second language learners produced the target sound in different phonetic positions in a different way, this study examined whether a significant difference would occur among the inaccurate productions of [θ] found in three word positions: initial (e.g., thank), medial (e.g., enthuse), and final (south). Table 5 illustrates the descriptive statistics for the inaccuracy percentage of [θ] in the initial, medial, and final word positions. It was predicted that the participants would produce the [θ] in the word initial positions better than they produced the [θ] in the word final positions.

Table 5

Descriptive Statistics for the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Word List Reading (Without Distraction)

	Median	IQR	Minimum	Maximum
[θ] in the initial position	0.35	0.50	0.00	1.00
[θ] in the medial position	0.00	0.20	0.00	0.50
[θ] in the final position	0.50	0.34	0.00	1.00

Because the data did not show a normal distribution, the nonparametric Friedman test was conducted (see Table 6), which reveals a significant difference in the participants' production inaccuracy of [θ] according to the position of [θ] within words, $\chi^2(2) = 20.77$, $p < .001$. The post hoc Wilcoxon tests show significant differences between inaccuracy perception of [θ] in the word positions of initial and medial ($z = 2.31$, $p = .021$, $r = .28$), and between the inaccuracy perception of [θ] in the word positions of medial and final ($z = 4.08$, $p < .001$, $r = .49$), confirming the hypothesis that the word position affected the production accuracy of the [θ]s.

Table 6

Friedman Test of the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Word List Reading (Without Distraction)

Test Statistics ^a	
N	34
Chi-Square	20.774
df	2
Asymp. Sig.	.000
a. Friedman Test	

The Effect of Speaking Styles and Word Position on Production of [θ]

Since the data of the participants' inaccuracy percentage of [θ] in different word positions did not approximate a normal distribution, the nonparametric Wilcoxon test was

used to investigate if there was an effect of speaking styles on the production accuracy of [θ]. Table 7 shows the descriptive statistics for the inaccuracy percentage of [θ] in the initial, medial, and final word positions in the story telling task while Table 8 illustrates the Friedman test statistics. Note that inaccuracy percentage was the greatest in the word final positions.

Table 7

Descriptive Statistics for the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Story Telling

	Median	IQR	Minimum	Maximum
Story Telling Initial	.07	.45	.00	1.00
Story Telling Medial	.00	.04	.00	.50
Story Telling Final	.73	.50	.00	1.00

Table 8 reveals a significant difference between the inaccuracy percentage of [θ] in the word final position produced in story telling and in the word list reading (without distraction), $z = -2.86$, $p = .004$, $r = -.35$. This indicates that the participants paid significantly more attention to [θ] in the word-final position in the word list reading than in the story telling, indicating that the speaking styles influenced their attention, which further affected the accuracy production of [θ].

Table 8

Wilcoxon Test of the Inaccuracy Percentage of [θ] in the Initial, Medial, and Final Word Positions in the Word List Reading and Story Telling

Test Statistics ^c			
	Story Telling Initial - Word List Reading Initial	Story Telling Medial - Word List Reading Medial	Story Telling Final - Word List Reading Final
Z	-1.377 ^a	-1.266 ^a	-2.862 ^b
Asymp. Sig. (2-tailed)	.168	.205	.004
a. Based on positive ranks. b. Based on negative ranks. c. Wilcoxon Signed Ranks Test			

The Effect of Auditory Distraction on [θ] in Word List Reading

In order to examine if the auditory distraction would impact the accuracy production of [θ], the inaccuracy percentage of [θ] produced in the word list reading (without distraction) and the word list reading (with distraction) was calculated (see Table 9) and the results were analyzed through paired-samples *t* test (see Table 10).

Table 9

Descriptive Statistics of Inaccuracy Percentage of [θ] in Two Word List Readings (With/Without Distraction)

	Mean	Std. Deviation	Lower Bund	Upper Bund
Word list reading (without distraction)	0.09	0.01	0.06	0.12
Word List Reading (with distraction)	0.129	0.102	0.093	0.164

The Paired-samples *t* test shows a significant difference between the inaccuracy

percentage of [θ] produced in the word list reading (without distraction) and the word list reading (with distraction), $t(33) = -3.83$, $p < .001$, $r = .55$. This indicates the influence of the auditory distraction on the speech production. Specifically, the auditory distraction decreases the production accuracy of the target [θ] sounds.

Table 10

Paired-Samples t Test Comparing the Inaccuracy Production of [θ] in Two Word List Reading Tasks

		Paired Difference							
				Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Mean	SD						t	df	Sig. (2- tailed)
Pair 1	Word List Reading (without distraction) –Word List Reading (with distraction)	-.04038 53	.061530 2	.0105523	-.06185 42	-.01891 64	-3.827	33	.001

The Effect of Auditory Distraction and Word Position on Production Accuracy of [θ]

It was hypothesized that the auditory distraction would interfere with the participants' monitoring of their speaking production by decreasing their attention to the target sound. Because previous research showed that attention allocated to the sounds in different word positions tended to be different, this study specifically analyzed in what word positions, [θ] would be significantly affected by the auditory distraction. The data of the word list readings both with and without distraction are presented in Table 11. Because the data did not satisfy the normal distribution, Wilcoxon T test was conducted to investigate the difference between the two speaking tasks for each word position (see Table 12).

Table 11

Descriptive Statistics for the Inaccuracy Percentage of [θ] in Initial, Medial, and Final Word Positions in the Two Word List Readings

	N	Median	IQR	Minimum	Maximum
Word List Reading (without distraction) Initial	34	0.35	0.50	0.00	1.00
Word List Reading (without distraction) Medial	34	0.00	0.20	0.00	0.50
Word List Reading (without distraction) Final	34	0.50	0.34	0.00	1.00
Word List Reading (with distraction) Initial	34	0.00	0.00	0.00	1.00
Word List Reading (with distraction) Medial	34	0.13	0.25	0.00	1.00
Word List Reading (with distraction) Final	34	0.82	0.41	0.00	1.00

Wilcoxon T test (Table 12) shows that the auditory distraction significantly affected the production accuracy of [θ] in the initial and final word positions. Specifically, the participants produced [θ] in the initial and final word positions significantly better in the word list reading (without distraction) than in the word list reading (with distraction). The results were $z = 2.77$, $p = .006$, $r = .34$; and $z = 2.80$, $p = .003$, $r = .36$, respectively.

Table 12

Wilcoxon Test for Inaccuracy Percentage of [θ] in Word Initial, Medial, and Final Positions in Two Word List Reading Tasks

Test Statistics ^c			
	Word List Reading (without distraction) Initial - Word List Reading (with distraction) Initial	Word List Reading (without distraction) Medial - Word List Reading (with distraction) Medial	Word List Reading (without distraction) Final - Word List Reading (with distraction) Final
Z	2.774 ^a	.455 ^b	2.979 ^b
Asymp. Sig. (2-tailed)	.006	.649	.003
a. Based on positive ranks. b. Based on negative ranks. c. Wilcoxon Signed Ranks Test.			

The Difference in Accuracy between Self-Perception and Perception of NS

Each participant's perception skills were investigated two ways: perception of the dictation of the word list produced by the NS and the perception of the word list reading (without distraction) produced by the participant (i.e., self-perception). The inaccuracy percentage of the perception of [θ] was used. Table 13 shows the descriptive statistics of the participants' perception of the production sources.

Table 13

Descriptive Statistics for the Inaccuracy Percentage of the Participants' Perception of NS' Speech and Self-perception of Word List Reading

	Means	Standard Deviation
Inaccuracy percentage of the perception of NS' word list reading	0.12	0.08
Inaccuracy percentage of the perception of NNS (participant)'s word list reading	0.18	0.09

Paired-samples *t* test shows a significant difference between the participants' self-perception and NS-perception, $t(33) = -4.50$, $p < .001$, $r = .62$ (see Table 14). The negative mean = $-.07$ ($SD = .08$) indicates that the participants could perceive [θ] sounds

produced by the NS better.

Table 14

Paired-Samples t Test for Inaccuracy Percentage of [θ] Perception in the Word List Reading (without distraction) and the NS's Word List Reading

		Paired Difference					t	df	Sig.(2-tailed)
		Mean	SD	Std. Error Means	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	NS- Perception Word List Reading - Self-perception Word List Reading (without distraction)	-.07	.08	.01	-.09	-.03	-4.50	33	.000

Relationship between Production and Perception

In this study, the participants' perception skills were divided into two parts: self-perception and NS-perception. Therefore, the relationships between production and both self-perception and NS-perception will be discussed as follows.

The relationship between production and NS-perception of [θ] was analyzed by Spearman rank-order correlation test, in that the data were not normally distributed. The descriptive statistics (see Table 15) show that the participants' perception of NS speech was better than their production of [θ] in the word list reading.

Table 15

Descriptive Statistics for the Inaccuracy Percentage of [θ] Production and NS-Perception

	Median	IQR	N
Production Word List Reading (without distraction)	.07	.07	34
Perception NS Word List Reading	.11	.08	34

Spearman rank-order correlation (see Table 16) reveals a significant correlation between the participants' production and NS-perception, $r_s = .57$, $p < .001$, $R^2 = .32$.

Table 16

The Correlation between the Inaccuracy Percentage of [θ] in Participants' Speaking Production and NS-Perception (NS)

			Production First Word List Reading	Perception NS Word List Reading
Spearman's rho	Production First Word List Reading	Correlation Coefficient	1.000	.570**
		Sig. (2-tailed)	.	.000
		N	34	34
	NS-Perception Word List Reading	Correlation Coefficient	.570**	1.000
		Sig. (2-tailed)	.000	.
		N	34	34

** . Correlation is significant at the 0.01 level (2-tailed).

Self-perception was another type of L2 perception skills that was analyzed in this study. Spearman rank-order correlation was conducted to see if there was also a relationship between the participants' production and self-perception. In this study, the inaccurate self-perception was defined as the words that the participants perceived that were different from what they produced. The descriptive statistics (see Table 17) illustrate that overall the participants produced the [θ] better (i.e., lower inaccuracy percentage) than they perceived it in their own speech.

Table 17

Descriptive Statistics for the Inaccuracy Percentage of Participants' Self-perception of [θ]

	Median	IQR	N
Production Word List Reading (without distraction)	.07	.07	34
Self-Perception Word List Reading (without distraction)	.17	.11	34

Table 18

The Correlation between the Inaccuracy Percentage of [θ] in Participants' Speaking Production and Self-Perception (NNS)

			Production First Word List Reading	Self- Perception First Word List Reading
Spearman's rho	Production Word List Reading (without distraction)	Correlation Coefficient	1.000	.715**
		Sig. (2-tailed)	.	.000
		N	34	34
	Self- Perception Word List Reading (without distraction)	Correlation Coefficient	.715**	1.000
		Sig. (2-tailed)	.000	.
		N	34	34
**. Correlation is significant at the 0.01 level (2-tailed).				

Spearman rank-order correlation revealed a significant correlation between the participants' production and self-perception, $r_s = .72$, $p < .001$, $R^2 = .51$.

CHAPTER 4

QUALITATIVE ANALYSIS

A qualitative study was conducted, including stimulated recall, a questionnaire, and an interview to analyze the relationship between attention and speech production or perception.

The Relationship between the Participants' Attention to [θ] in Various Word Positions and Speaking Styles

Item 7 to item 12 on the questionnaire concerned the participants' self-assessed amount of attention paid to the production of [θ] in three word positions (initial, medial, and final) in two speaking styles (the first word list reading and the story telling). A 5-point scale was used in which 1 represented the least attention paid to the target sound while 5 indicated the most attention paid to it. As shown in Table 19, there was little variation in median responses across items.

Table 19

Descriptive Statistics of the Questionnaire Results (Item 7 - Item 12)

	Median	IQR	Minimum	Maximum
The amount of attention I paid to word initial "th" in word list reading without distraction	3.00	2.00	1.00	5.00
The amount of attention I paid to medial position "th" in word list reading without distraction	3.00	1.25	2.00	5.00
The amount of attention I paid to final position "th" in word list reading without distraction	3.00	1.00	1.00	5.00
The amount of attention I paid to word initial "th" in story telling	3.00	1.00	1.00	5.00
The amount of attention I paid to medial position "th" in story telling	2.00	1.00	1.00	5.00

Table 19 (cont'd)

The amount of attention I paid to final position "th" in story telling	3.00	1.00	1.00	5.00
--	------	------	------	------

Table 20 reveals that Cronbach's alpha (α) for the subscale *Amount of Attention Paid to [θ] in Different Speaking Styles* was .87, which was considered reliable.

Table 20

Reliability Analysis of the Questionnaire

Cronbach's Alpha	N of Items				
.867	6				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted	
The amount of attention I paid to word initial "th" in first word list reading	14.18	15.665	.613	.856	
The amount of attention I paid to medial position "th" in first word list reading	14.41	18.553	.515	.868	
The amount of attention I paid to final position "th" in first word list reading	14.15	15.584	.712	.836	
The amount of attention I paid to word initial "th" in story telling	14.65	15.144	.729	.832	
The amount of attention I paid to medial position "th" in story telling	14.91	16.992	.687	.843	
The amount of attention I paid to final position "th" in story telling	14.76	14.791	.758	.827	

To distinguish if there was a significant difference between the attention allocated to [θ] in the participants' word list reading without distraction and story telling, the Wilcoxon test (see Table 21) revealed a significant difference between attention paid to [θ] in the word initial, medial, and final positions in the word list reading and in the story telling

respectively, $Z = 2.37$, $p = .018$, $r = .29$; $Z = 2.83$, $p = .005$, $r = .34$; $Z = 3.38$, $p = .001$, $r = .41$.

Table 21

Difference between Attention Paid to Initial, Medial, and Final [θ] in Word List Reading without Distraction and Story Telling

Test Statistics ^b			
	word initial position	Word medial position	Word final position
Z	2.368 ^a	2.825 ^a	3.384 ^a
Asymp. Sig. (2-tailed)	.018	.005	.001
a. Based on positive ranks.			
b. Wilcoxon Signed Ranks Test.			

Because the data yielded from the 5-point scale could not completely reflect the amount of attention the participants paid to [θ] in each word position, the researcher also relied on the interview and stimulated recall data to support the statistical data from the questionnaire. Extract 1 and Extract 2 below demonstrate that with limited attentional resources, the participants tended to give priority to the [θ] in the word-initial positions with least consideration to [θ] in the word-final positions.

Extract 1. (S)³

Participant (P): 我这里没有考虑在单词末尾的 th 因为我觉得就算说得不怎么准确别人也知道我在说什么。而且我也没功夫去想这个音怎么发。我要想着接下来应该说什么。这个不像后面的读单词, 我有充分的时间去考虑每一个音该怎么发音准确。

Participant (P): Here, I did not consider the pronunciation of “th” in the word final position because I thought that even if I did not pronounce it correctly, other people can understand me. Besides, I did not have enough time to consider its pronunciation. I was thinking about the following contents. This is not like the word list reading in which I had enough time to consider the pronunciation of each sound.

³ (s) = stimulated recall. This means that the extract comes from the stimulated recall

Extract 2. (I)⁴

P: 一般来说我首先关注的肯定是单词开始的 **th** 因为这个是最明显的, 而且如果第一个音发准确了别人就容易理解我在说什么了, 就算我后面的发音不是很标准也不影响理解。

P: In general, I definitely paid attention to the “th” sounds in the word initial positions first because the sound on this position is the most salient one. In addition, if I pronounced the first sound correctly, people would understand me more easily, even if I failed to pronounce the rest of the sounds correctly.

Factors Influencing the Participants’ Attention to Production

In addition to the impact of the speaking styles on the attention the participants paid to different word positions, comments revealed three other factors attracting participants’ attention: salience of words, familiarity of words, and word order as shown in the following Extracts. In Extract 3 and Extract 4, the participants explained why they paid attention to the words that were shown to them.

⁴ (I) = interview. This means that the extract comes from the interview

Extract 3. (S)

P: 我关注了这两个单词, Ruth 和 Keith. 因为这两个单词很明显的放在图片上了, 而且看上去还是两个人名, 所以我想我肯定需要把这两个人名包括进去(故事里)并且我都已经看到这两个单词了所以就知道单词的拼写上有 th. 我知道 th 这个音中国人特别难发所以我就注意了。

P: I noticed the (pronunciation) of these two words, Ruth and Keith because they were shown in the pictures and I knew that they were the names of two characters. I thought that I should include these two characters' names into my story. Also because I could see the spelling of these two words so I noticed that there was "th" in these two words. I knew that the Chinese people had difficulty in producing "th" so I paid special attention to it.

Extract 4. (I)

P: 我觉得如果能看得到单词, 就能刺激我去关注这个单词的发音。比如图片上这些个有 th 的单词就刺激我去关注 th 的发音。如果看不到这些单词我可能就不会去关注单词的发音了而只是想着怎么把内容说下来。

P: I think if I could see the words, I would be stimulated to notice the words' pronunciation. For example, these words containing "th" in the pictures inspired me to notice my production of "th." If I could not see the spelling of these words, I might not notice their pronunciation but only focus on the contents.

The sentence structure also determined the participants' allocation of attention. Four participants claimed that they only noticed the words that were "comparatively separated from the rest of the sentences." In other words, their attention was paid to the adverbials first, in that the adverbials "do not contain too many contents and the words are easy."

Extract 5 shows an example of this situation.

Extract 5. (S)

P: 这里我关注了 south 和 north 这两个单词的发音因为感觉上这两个单词在句子中相对比较独立。我的意思是他们看上去不属于句子的核心部分, 只是一个状语部分, 我说这个部分的时候就已经可以不怎么考虑句子的核心思想了, 所以就有额外的精力去关注单词的发音了。

P: Here I noticed "south" and "north" because it seems to me that these two words are separated from the rest of the sentences. I mean they are not the core of the sentences but are only the adverbials. When I spoke this part, I did not need to consider the core idea of the sentence so that I could spare

additional attention on the pronunciation of these two words.

Twelve participants mentioned that the familiarity of the words determined their attention paid to the words' pronunciation. Interestingly, half of these participants indicated that the familiar words attracted their attention while the other half of the participants thought the unfamiliar words attracted their attention. Their reasons are illustrated in Extract 6 and Extract 7.

Extract 6. (I)

P: 如果是我比较熟悉的单词我就会关注读音，因为我有足够的精力去想这个单词该怎么发音。如果是不熟悉的单词我就得先想着这个单词怎么用就来不及去考虑发音了。而且有时候单词我不熟悉并且发音很复杂不太容易发音的话我就干脆不去想怎么发音准确了，因为我认为不值得为了这个单词去仔细想发音，结果把后面的内容都忘记了连不上了。

P: If I was familiar with the words, I would focus on the words' pronunciation because I already knew how to use these words. However, if I was not familiar with the words, I had to first consider the words' meanings and then the usage of these words so I could barely spare any effort to think about their pronunciation. Besides, I would not bother to think about the pronunciation of the words that I was very unfamiliar with or that were too hard to pronounce because I thought it was not worth dwelling too much on them since I needed to first make sure that I expressed what I wanted to say.

Extract 7. (I)

P: 我会关注不熟悉的单词。因为熟悉的单词我就不会再去想发音了因为已经说习惯了成本能了。只有那些不是很熟悉的还没有形成我的发音习惯的我才会去关注发音保证发音准确。

P: I focused on the pronunciations of the unfamiliar words because I had been already accustomed to the pronunciations of the familiar words as if they had been my instinct. However, because I was not used to the pronunciations of the unfamiliar words, I had to spare extra effort on them to make sure I pronounced them correctly.

Relationship between Attention and Monitoring of Speech Production

In addition to what the participants paid attention to, the researcher also wanted to

know if the participants noticed the target sound ([θ]). Extract 8 and Extract 9 represented two ways of self-monitoring during speaking: Proprioceptive feedback and auditory feedback (self-perception).

Extract 8. (I)

P: 我知道发[θ]这个音的时候舌头要伸出来咬舌头，所以当我有时候说话说得快了可能舌头没有放到位我就知道自己肯定这个音没有发到位，因为舌头的位置就不对。

P: I know that I need to stretch my tongue out and put it between the upper and lower teeth to pronounce [θ]. So sometimes when I speak too fast that I don't have enough time to put my tongue at its right place to pronounce [θ], I know that I fail to pronounce it correctly because I know that the place of my tongue is not correct.

Extract 9. (I)

P: 因为我能听到自己的声音，所以通过比较我听到的自己发出的声音和我知道的应该是标准的发音发出来的声音，我就能知道自己发音是不是准确了。

P: Because I can clearly hear my own voice, through comparing my voice and the "standard" pronunciation, I can determine if my pronunciation is correct.

It can be concluded that when monitoring their own speech production, the participant in Extract 8 mainly emphasized the proprioceptive feedback while the participant in Extract 9 mainly noticed the auditory feedback (i.e., self-perception).

Interestingly, it was also revealed that the participants' monitoring strategies were related to how they were taught the target [θ] sounds. Specifically, those participants who mainly relied on the proprioceptive feedback admitted that they were taught the place and manner of the [θ] sound when they learnt it while those who depended on the auditory feedback indicated that they learnt it through repeatedly listening to the target [θ] sounds.

Relationship between Monitoring Strategies and Distraction

Among 21 participants who explicitly indicated their monitoring strategies, 13 participants relied on the proprioceptive feedback while 8 participants depended on the auditory feedback. It was hypothesized that the latter would be influenced by the auditory distraction more than the former in monitoring the production. Extract 10 and Extract 11 show how the auditory distraction influenced the participants' self-perception.

Extract 10. (I)

P: 新闻对我读单词的影响很大因为我根本就听不到自己读单词的声音，所以也根本没办法知道自己读的对不对，这样我就没办法纠正自己发音如果读错的话。

P: I think (listening to) the news influenced me greatly because I could not hear my own voice (when the news was played) so that I did not know if I pronounced the words correctly (without hearing my own voice) and therefore I did not know whether I needed to modify my pronunciation if I mispronounced the words.

Extract 11. (S)

P: 这里，我停顿了，因为我发现自己听新闻的时候就根本听不到自己在说什么了，所以我不能确定自己读的是不是对的。为了确保我念的是对的，我只好先听一下新闻，这里停顿一下，然后后来再接着读单词，可是我读单词的时候我就尽量不去听新闻了。

P: I stopped here because I found that I could not hear my voice when I was listening to the news so I did not know if I pronounced the words correctly. In order to make sure that I pronounced the words correctly, I had to stop when I was listening to the news and then continue to read the words. When I read the words, I tried not to listen to the news.

Factors Influencing the Participants' Attention to Perception

The stimulated recall in this study revealed that the acoustic contrast between [θ] and [s] attracted the participants' attention when they perceived the sounds. In Extract 12, the

participant looked through the words he wrote down and explained any change he made during the dictation. The participant also reported his confidence in the words he perceived.

Extract 12. (S)

P: 当我听到 **sink** 的时候我很确定我听到的是 **sink**。但是这个单词听上去和我之前听到的我认为是 **sink** 的那个单词听上去又不一样, 所以我觉得我前面听到的可能就不是 **sink**。一般来说能够混淆的也就只有 **think**, 所以我就把刚才写下来的 **sink** 修改成了 **think**。如果我后面没听到 **sink** 那我就不太可能会想到我前面可能听错了是 **think** 不是 **sink**。老实说如果你不把 **sink** 和 **think** 放在一起的话没有了对比我估计就不会意识到自己可能听错了。也就都听成 **sink** 了。

P: When I heard “sink” I was very sure that it was “sink.” However, since this word did not sound the same as the word I just heard before this, which I thought was “sink,” I thought maybe what I heard before this word was not “sink.” Also because in general I would only mistake “sink” with “think,” I decided to change “sink” that I just wrote down to “think.” Yes if I did not hear “sink” later, I would very likely not to realize that I might mistakenly perceive “think” as “sink.” Honestly, if you did not put “sink” and “think” together, without the acoustic contrast, I would not realize that I perceived it wrong. Probably I would perceive both of them as “sink.”

Extract 12 uncovers that other than confusing the participants, the similar sounds that were played to the participants successively actually facilitated the participants' perception of both NS and NNS' speaking. Interestingly, half of the participants (N = 17) tended to perceive both [θ] and [s] as [θ] while the half of the participants (N = 17) believed they heard [s].

Despite the acoustic contrast, the interview revealed that the clarity or the familiarity of the speech attracted the participants' attention as well. Extract 13 and Extract 14 exemplify these two cases respectively.

Extract 13. (I)

P: 很容易注意到美国人念的[θ]和[s]的区别因为这个区别非常大而且这种区别很

稳定，但是非英语母语者的发音在这两者的区别上就很模糊，有时候分辨的出来有时候就分辨不出来，相对来说不是很稳定吧。

P: It is easier to notice the difference between [θ] and [s] produced by the NS because he pronounced [θ] and [s] very differently. In addition, such difference is very stable. However, the NNS' production of [θ] and [s] do not show much salient difference. Sometimes I can tell the difference but sometimes they just sound the same to me. So I think probably the NNS' production is not stable enough.

Extract 14. (I)

P: 我觉得注意到非母语者说的[θ]和[s]的区别更加容易，我不是说非英语母语者的发音更加准确，只是因为我更加熟悉这个发音，我知道这个发音的特色对应的是哪个音标。但是辨析英语母语者的发音的时候因为不熟悉虽然我知道他的发音肯定准确得多，我还是辨析不出来哪个是[θ]哪个是[s]。

P: I think it is easier to notice the difference between [θ] and [s] produced by the NNS. I am not saying that the NNS' production is more correct than the NS' production but what I mean is that because I am more familiar with the voice of this NNS⁵ and I know the features of each sound (on IPA). However, because I am not familiar with the voice of the NS, I could not completely identify what he said, even though I knew his pronunciation should be correct and standard.

Extract 14's statement that the familiarity with the features of the sounds produced by the NNS attracted the participants' attention to the difference between [θ] and similar sounds implied that perhaps some participants depended on incorrect criteria (i.e., the features of the sounds that were generalized by the participants but were not accurate) to distinguish between [θ] and [s], two of the most frequently misperceived English fricatives by Chinese learners of English. This misunderstanding of the sounds might influence their perception performance. The interview revealed that commonly, the participants misunderstood that the difference between [θ] and [s] was either the stress or the length of the sounds. Extract 15 and Extract 16 exemplified these misunderstandings.

⁵ This participant already knew that the NNS he heard was himself.

Extract 15. (I)

P: [θ]这个音要比[s]这个音轻，所以我听自己的录音的时候就能很明显听出来区别。但是那个美国人念的时候我感觉听上去的音调强弱都是一样的，所以我就听不出来到底是[θ]还是[s]了。

P: (I think) that the [θ] sounds should be less accented than the [s] sounds. Because I will pronounce [s] more accented than I pronounce [θ] so when I listened to my own recordings, I could tell a salient difference between two. However, I felt like that the American speaker (NS) produced the two sounds with the same stress so I could not tell if he produced [θ] or [s].

Extract 16. (I)

P: [θ]要比[s]发音听上去要短很多。我能听出来我自己发音的时候一长一短的区别。但是美国人发音的时候这两个音的长度区别不是很大，所以我觉得区分起来很困难。

P: [θ] sounds shorter than [s]. When I listened to my own recordings, I could identify these two sounds according to their length. However, the lengths of these two sounds produced by the NS seem to me approximately the same so I have difficulty in distinguishing which one is which.

If the participants relied on the wrong criteria to perceive the target sounds, chances were that they would also depend on these same incorrect criteria to produce the sounds, thus leading to the incorrect pronunciation. To investigate the interaction between the participants' production and perception, the interview question 6 was discussed and the findings are presented in the next section.

Relationship between Production and Perception

The hypothesis posited in the previous section that the participants' incorrect production of sounds might lead to their inaccurate perception was confirmed by the interview. Extract 17 and Extract 18 reflected two common mistakes in the participants' speech production: relying on the stress or length of the sounds.

Extract 17. (I)

P: 我发[θ]这个音的时候就比较弱化，而发[s]的时候就比较重读。我就是靠重音强弱来确保我读的这两个音是有区别的。

P: When I pronounced [θ], I pronounced it softer than I pronounced [s]. I relied on the stress of these two sounds to make sure that I pronounced them differently.

Extract 18. (I)

P: 如果我念[θ]这个音我会念得比较收，也就是比较短，但是念[s]的时候就会比较长，尤其是当这两个音都在结尾的时候，我发[θ]就明显短于发[s]。

P: If I pronounced [θ], I would stop the sound very soon. In other words, I would pronounce it comparatively short. But I would produce [s] comparatively longer. Especially when these two sounds are at the end of the word, I would produce [θ] saliently shorter than I produced [s].

The participants in Extract 17 and Extract 18 were the same ones who reported Extract 15 and Extract 16 correspondingly. The association between Extract 15 (perception strategy) and Extract 17 (production strategy), as well as between Extract 16 (perception strategy) and Extract 18 (production strategy) suggested a relationship between the participants' speaking perception and production.

CHAPTER 5

RESEARCH DISCUSSION

The data analyzed in this study yielded many noteworthy findings. The purpose of this section is to interpret these findings to answer the research questions of this study.

What is the effect of speaking style, word position and auditory distraction on production accuracy of [θ]? Do these factors interact? If yes, how?

Speaking Style

The quantitative study revealed a significant effect of speaking style on the production accuracy of [θ]. As anticipated before the study, the participants produced [θ] better in the word list reading without distraction than in the story telling. However, in this study, there was no support for the use of the term formality as a distinguishing characteristic of story telling and word list reading. The interview showed that most of the participants regarded both speaking tasks as English speaking tests.

Yet the difficulties of these speaking tasks were by no means the same. Specifically, the story telling was much more challenging than the word list reading. Since the questionnaire results evidenced that the participants paid less attention to the production accuracy of [θ] in the story telling than they did in the word list reading without distraction, it could be deduced that the amount of attention paid to [θ] correlated with the challenging nature of the speaking style.

This inference was supported by the participants' comments such as "I didn't have enough time to carefully think about how to pronounce the words when I was thinking about the following content;" "I was so struggling at describing the pictures with

appropriate words that I barely concerned any pronunciation issue.”

Having said that, some participants in this study admitted that they tended to care less about the accuracy of the pronunciation when they chatted casually with others in English or they chose not to self-correct their inaccurate pronunciation even if they were aware of the errors as long as the erroneous pronunciation did not interfere with communication. On the contrary, they tried to pronounce as correctly as possible when they made presentations, had interviews, or attended meetings. These statements indirectly indicated that at least there was a relationship between the type of speaking task and the extent of the speakers’ effort to pronounce accurately.

Word Position

The quantitative study showed that the participants were poorest at producing the [θ] in the word-final positions in the word list reading. The reason for this phenomenon was found in the qualitative study, in which the participants explained that the inaccurate pronunciation of the sound in the word-final positions had the least influence on the intelligibility of the speech. In addition, some participants explained that they had more time to allow their muscles and tongues to approach the correct position before they produced the initial [θ]s, if it was the first sound in the word, but they might fail to quickly switch their muscles and tongue from the previous positions to the correct one to produce the final [θ]. As one participant claimed, “If my tongue could be more agile, my pronunciation would be better.”

What deserves further study was that unexpectedly the [θ] sounds in the medial position of words (e.g., enthuse) actually were mispronounced the least. Following the

comments of some of the participants in the current study, it should be more difficult to pronounce the [θ] sounds positioned between two other sounds than to pronounce them either in the initial or final position, at least in word list reading.

The possible reasons might be that the words chosen in this study that contained the [θ] in the medial position were either so salient to the participants because they were infrequently used (e.g., *enthuse*) that the sufficient amount of attention attracted by such salience offset the complexity of the pronunciation, or because they were so familiar to the participants (e.g., “birthday” and “something”) that the participants had already been skillful at pronouncing these words correctly. This assumption was supported by the discussion results that a), the more salient the target sounds appeared to the participants, the more they would attract attention; b), half of the participants tended to pay more attention to the familiar words; and c), the production accuracy was positively associated with the amount of attention allocated to it.

Interaction between Speaking Style and Word Position

The statistical analysis yielded a significant difference between the participants' inaccuracy percentage of [θ] in the word-final position in the story telling and in the word list reading without distraction, suggesting the influence of the speaking style on production accuracy. The quantitative study revealed the participants' hierarchy of allocating attention, specifically the participants ranked [θ] in word-final position as the least important sound to be pronounced correctly.

Two inferences could be drawn. First, the more demanding the speaking style was, the fewer attentional resources were available to be paid to the pronunciation of [θ]; second,

the less impact the inaccurate production of [θ] in certain word positions had on the intelligibility of the speech, the less attention it would gain from the learners.

Auditory Distraction

The statistical analysis revealed a significant difference between the production accuracy of [θ] in the word list reading with and without distraction, suggesting that the auditory distraction significantly impacted the production accuracy of [θ] in the word list reading with distraction, in that all the variables such as the content, speaker, and length of the speech were strictly controlled; only the auditory distraction was the difference.

This finding was also supported by the qualitative study, in which most of the participants commented that due to the dual tasks (i.e., speaking and listening simultaneously), they could not “focus on” pronouncing the words correctly or they were “distracted” when they read the words. In addition to the participants’ concentration on the reading task, the auditory distraction also interfered with the participants’ self-perception, which was closely related to the monitoring of production and therefore played a role in the inaccurate production of [θ]. During the stimulated recall, some participants complained that they could not “hear” their “voice” because of the auditory distraction so they found it hard to “judge” if their pronunciation was correct.

In addition to the effect of reducing the participants’ attention to pronunciation, the results uncovered three factors that significantly attracted their attention to the production and perception of [θ]: the salience of the words, the salience of the sounds, and the words’ familiarity.

First, the visual information such as the words printed on the paper greatly attracted the

participants' attention paid to the pronunciation when they spoke. This implied that the visual input, especially the spelling of "th" within the words, activated the participants' cognitive response to [θ], or alerted them to be careful about a coming sound that they would have difficulty in pronouncing.

Second, the acoustic contrast of words attracted the participants' attention to perceiving the subtle acoustic difference between two similar sounds like [θ] and [s], which was both evidenced by the participants' perception performance (i.e., changing the previous word after hearing the latter word) and confirmed by the participants' comments during the interview. This finding implies that the acoustic contrast between two similar sounds attracts the listeners' attention by increasing the salience of both sounds.

These two factors indicate that attention would be primarily allocated to the most salient information. In other words, the element, which endows the information receivers with the most input, wins the battle in the competition for limited attentional resources. This implication was also supported by the participants' feedback on the dual task (i.e., listening to the news while reading the second word list). The participants who performed poorly in recalling the news explained that: "When I read the words, I could not help concentrating all my attention on them even though I knew that I needed to remember what I heard. I felt like I could not control my brain during that test. I assumed that to me, the visualized input is stronger than the acoustic input." This feedback clearly showed that the allocation of attention was not completely a matter of control but instead, was greatly determined by the comparison of salience among various inputs.

The third factor is the word's familiarity. Half of the participants noticed the familiar

words while the other half noticed the unfamiliar words. It was possible that the familiar words were more salient to the former while the unfamiliar words were more salient to the latter. If this assumption is verified, this factor, together with the previous two factors, could be categorized as one general factor: salience.

Interaction between Auditory Distraction and Word Position

The Wilcoxon test demonstrated a significant difference between the inaccuracy percentage of [θ] in both the word-initial and word-final positions in the word list reading with and without distraction. The production of [θ] in the word-final position appeared to be slightly more affected by the auditory distraction than the production of [θ] in the word-initial position.

Perhaps because of the stress on the word initial phonemes, the [θ] in this position sounded more salient to the participants than the [θ] in the word final position and it attracted more of the participants' attention.

Is there a relationship between L2 learners' perception and production?

In general, the participants' L2 production was significantly and positively correlated with perception. However, the results of this study generated more fruitful implications than this simple conclusion. The relationship between the participants' self-perception and NS-perception was investigated, as well as the relationship between their self-perception or NS-perception and their speech production. This is discussed further below.

First, generally speaking, the participants' perception of the NS was better than their self-perception. This might be caused by two reasons. The NS' production was clearer

and more standard than that of the NNS and therefore the former was more identifiable to the participants than the latter. In addition, some participants seldom monitored their recorded speech production, therefore, they were actually less familiar with their own voices than with the voice of a NS.

Second, nearly half of the participants' incorrect pronunciations might be caused by the incorrect perception of the NS speech, which was suggested by comments from the interview. Extract 15, 17 and Extract 16, 18 were two examples typically showing that because the participants perceived [θ] and [s] produced by the NS based on incorrect acoustic criteria, they also produced the target sounds in an incorrect way.

Third, a few participants ($N < 5$) specifically denied the relationship between their speaking production and perception. They claimed that even though they knew what the correct [θ] should sound like, they could not pronounce it correctly in that they had no idea how to pronounce it. One participant explicitly stated that to him, "the speaking is entirely separated from the listening," which implied that at least to some participants, the correlation between the speaking production and perception was weak, if not completely non-existent.

What is the effect of distraction on participants' monitoring strategies?

Two monitoring strategies were discovered in this study: proprioceptive feedback and auditory feedback. The results showed that the participants who relied on auditory feedback were influenced by the auditory distraction more than the participants who depended on proprioceptive feedback. This is not unexpected, for the auditory distraction occupied the participants' attentional resources.

CHAPTER 6

CONCLUSION AND IMPLICATION

In this study, several research questions were discussed related to the L2 English learners' speech production, the perception of speech produced by a NS, the perception of their own speech, and the attention paid to pronunciation in different speaking styles. The summary of the results is presented in this section, followed by the implications, and limitations of the research, and options for future study.

First, the relationship between the formality *per se*, of the speaking style and the accuracy of the production of [θ] was neither rejected nor supported directly by the quantitative study but was indirectly implied by some participants' comments given during the interview. However, what the statistical analysis and the qualitative study (i.e., interview and stimulated recall) supported was a relationship between the difficulties of the speaking tasks (i.e., speaking styles) and the amount of attention paid to [θ] sounds during the speech, as well as the production accuracy of [θ]. For example, the participants noticed the pronunciation of [θ] more and performed better in reading the word list, a less challenging speaking task, than in telling the story, a more challenging speaking task.

Second, a relationship was found between the speaking styles and the production accuracy of [θ] in different word positions. The more demanding the speaking style was and the less salient the word position was, the less attention was paid to the production of [θ] and the less accurately the [θ] was produced.

The unexpected result was that the [θ] sound in the medial position of the words was

produced the best. The discussion concluded that since the word position was not the only factor that influenced the production accuracy of the target sound, other factors such as the salience of the phones and the familiarity of the words might offset the negative impact the comparatively complicated phonetic environment had on the production of [θ] sounds.

Third, the L2 learners produced [θ] better than they perceived it. A further investigation showed that the L2 learners' perception of the NS was better than their self-perception, probably because the NS's production was clearer and more standard than the NNS's production.

Fourth, relationships were discovered between the salience of the audio or visual input of the target [θ] sounds and the amount of attention paid to them. Relationships were also found between the auditory distraction and the participants' monitoring strategies.

The conclusions of the study implied several points for future study. First, as Tarone (1979) said, it is hard to observe the participants' real casual speaking as long as they are aware that the research experiments are being conducted. In this study, although the speaking tasks (or styles) were supposed to have different degrees of so-called formality, which was proposed in the previous studies, it was not convincing that the participants regarded the story telling as a less formal speech task than the word list reading. The researcher had hypothesized that other than the formality, it was the challenges of the speaking task that determined the attention paid to the pronunciation as well as the pronunciation accuracy. It might be interesting to see, if possible, how the second language learners perform in a real "vernacular" speech task.

Second, through investigating the influence of the auditory distraction on the accuracy of the L2 learners' speaking production, the researcher discovered that when the participants were listening to the news and reading the words simultaneously, the allocation of their attention was partially determined not by themselves consciously but by the comparison of salience between the auditory and visual inputs. In other words, the more noticeable the input was, the more attention it won. In order to further clarify the competition for attentional resources observed in this study, a quantitative study with well-designed laboratory experiments is needed. For example, by controlling the inputs' features including the volume and the font, and by using fMRI, the researchers might have a clearer idea about the correlation between the types and features of the input and the amount of attention they draw.

Third, this study revealed that in terms of the speaking production and perception, the L2 learners could be categorized into two types: the one mainly relying on the auditory feedback and the one mostly focusing on the proprioceptive feedback. The former was clear about what a correct [θ] should sound like while the latter knew how it should be produced and how it would feel. The interview revealed that almost all the participants who relied on the first monitoring strategy (i.e., auditory feedback) learnt the target [θ] sound through listening to the model sound produced either by the English NS or by the teacher whose pronunciation could be incorrect, while the participants using the second monitoring strategy (i.e., proprioceptive feedback) were taught explicitly the place and manner of articulation of the target [θ] sounds. Pedagogically, this finding implies a significant relationship between the pedagogical methods and the learners' second

language performance. In other words, to some extent, the learners' problematic second language speech production and perception could be caused by the incorrect instruction. Another pedagogical implication is that it might be helpful if the diagnostic assessment of the causes of the learners' pronunciation issue could be placed before any pronunciation training is carried out. Further study might focus on figuring out the pedagogical methods that are specifically effective to each type of L2 learner.

Fourth, the effect of the word positions on the participants' speech production of the target [θ] sounds shows a possible influence of phonetic context in the production of the fricative by L1 Mandarin speakers. Relevant finding was obtained by Trofimovich, Gatbonton, and Segalowitz (2007). They studied the French speakers' production of the voiced interdental fricative in different phonetic contexts and discovered the relation between the phonetic contexts and the production of the target sounds. However, the limitation of their study is that the contexts were limited to words in the determiner class, and they did not take into account the fact that the production of [θ] in "to the" and "at the" differs not only in terms of the preceding sound (vowel vs. alveolar stop) but "at the" also involves co-articulation. In a future study of the influence of phonetic context on the speech production of the English fricatives by L1 Mandarin speakers, researchers could take adjacent sounds, co-articulatory phenomena, and word position into account.

Admittedly, this study had several limitations that should also be improved in the future study. First, though 34 participants already yielded a large amount of data, they are still not enough if more robust statistical analysis results are needed. Second, though nearly half of the participants did not realize that the recordings of the word list reading were

produced by themselves, it could not be guaranteed that the participants' memory of the words they just read did not influence the research results. It would be better if there were an interval between the tests of production and perception so that it might be less possible for the participants to rely on memory to perceive the words.

APPENDICES

APPENDIX A

WORD LIST READ BY THE ENGLISH NATIVE SPEAKER

Thigh	Tie	Die	Sigh
Both	Boat		
Thread	Dread	Tread	
Dead	Debt	Death	
Thin	Tin	Sin	
Maths	Mats		
Deem	Seem	Theme	Team
Worse	Worth	Word	
Fate	Faith	Face	Fade
Ether	Eater		
Thorn	Saw	Torn	
Martha	Martyr	Masa	
Thug	Dug		
Bird	Birth		
Youth	Use		

APPENDIX B

PICTURE PROMPTS FOR TELLING THE STORY “THREE LITTLE PIGS”



Language Learning 59:3, September 2009, pp. 581–621

612

Figure 1. Picture Prompts for Telling the Story “Three Little Pigs”

These pictures were borrowed from the article Think or Sink: Chinese Learners’ Acquisition of the English Voiceless Interdental Fricative, by D. V. Rau, H-H A. Chang, and E. E. Tarone in *Language Learning* 59:3, September 2009, pp. 581-621, © 2009 Language Learning Research Club, University of Michigan, published by Wiley-Blackwell, and were used with permission.

Figure 1 (cont,d)



These pictures were borrowed from the article Think or Sink: Chinese Learners' Acquisition of the English Voiceless Interdental Fricative, by D. V. Rau, H-H A. Chang, and E. E. Tarone in *Language Learning* 59:3, September 2009, pp. 581-621, © 2009 Language Learning Research Club, University of Michigan, published by Wiley-Blackwell, and were used with permission.

APPENDIX C

THE SECOND WORD LIST READ BY THE PARTICIPANTS

Welsh	Wealth	Welt	Wells
Worth	Worse	Wordsworth	
Breath	Bread		
Keith	Keys		
Moss	Moth		
Mouth	Mouse		
Pat	Path	Pass	
South	Southern		
Teeth	Tease		
Truth	Truce		
Youth	Use		
Nothing	Nodding		
Something	Everything	Anything	
Plaything	Placing		
Healthy	Healthier	Wealthy	Wealthier

APPENDIX D

THE FIRST WORD LIST READ BY THE PARTICIPANTS

Thick	Tick	Sick		
Sink	Think	Link	Sinking	Thinking
Shank	Sank	Thank	Tank	
Tin	Sin	Shin	Fin	Thin
Thing	Sing			
Sunder	Thunder			
Enthuse	Ensues			
Three	Tree	Free		
True	Through			
Thought	Sought	Fought	Taught	
Earth	Errs			
Forth	Fort	Force		
Norse	North			
Strength	Strengths	Strengthen		
Tenth	Tense	Tent		

APPENDIX E

QUESTIONNAIRE ABOUT THE PARTICIPANTS' ATTITUDE AND ATTENTION PAID TO THE VOICELESS INTERDENTAL FRICATIVE

Table 22

Questionnaire about the Participants' Attitude and Attention Paid to [θ]

Statement	Com- pletely Agree	Agree	Neutral	Dis- agree	Com- pletely Dis- agree
1. I want to get rid of my Chinese accent in my English spoken language	1	2	3	4	5
2. I pay attention to "th" sounds when I speak English, in general	1	2	3	4	5
3. I noticed that English native speakers produce "th" differently than I do	1	2	3	4	5
4. I believe that the most important reason why I have Chinese accent when I speak English is that I cannot pronounce "th" in a native-like manner.	1	2	3	4	5
5. I didn't have time to correct my pronunciation of "th" when I described pictures in this experiment , even if I knew I didn't pronounce it correctly	1	2	3	4	5
6. I paid special attention to pronounce "th" correctly when I read the word-list in this experiment because I had enough time to do so	1	2	3	4	5

Table 22 (cont'd)

Amount of attention	most	a lot	medial	not much	least
7. The amount of attention I paid to word initial “th” in word list reading without distraction (eg. <u>th</u> ink)	5	4	3	2	1
8. The amount of attention I paid to medial position “th” in word list reading without distraction (eg. en <u>th</u> use)	5	4	3	2	1
9. The amount of attention I paid to final position “th” in word list reading without distraction (eg. fa <u>th</u>)	5	4	3	2	1
10. The amount of attention I paid to word initial “th” in <u>story telling</u> (eg. <u>th</u> ink)	5	4	3	2	1
11. The amount of attention I paid to medial position “th” in <u>story telling</u> (eg. en <u>th</u> use)	5	4	3	2	1
12. The amount of attention I paid to final position “th” in <u>story telling</u> (eg. fa <u>th</u>)	5	4	3	2	1

APPENDIX F

THE MAIN INTERVIEW QUESTIONS

1. Did listening to the news report influence your reading of the second word list? Why?
2. Did you know that the second dictation was produced by yourself?
3. Whose speaking production of [θ] do you think was easier to identify, the native English speaker or the non-native English speaker? Why?
4. What is your strategy for monitoring your speaking production?
5. What strategies did you use during the word list reading to make sure that the minimal pairs you produced were intelligible and distinguishable?
6. Please evaluate your production of “th” sounds compared with the NS’s production of them.
7. When you communicate with others in English, do you notice their pronunciation of “th” sounds?
8. In your daily life, do you imitate the “th” sounds produced by the English native speakers? Why or why not? How do you imitate?

REFERENCES

REFERENCES

- Al-Hejin, B. (2005). Attention and awareness: Evidence from cognitive and second language acquisition research. *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics*, 4(1).
- Baars, B. J., Motley, M. T., & Mackay, D. (1975). Output editing for lexical status from artificially elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behaviour*, 14, 382-391.
- Berg, T. (1986). The problem of language control: Editing, monitoring and feedback. *Psychological Research*, 48, 133–144.
- Brannen, K. (2011). The perception and production of interdental fricatives in second language acquisition. *Dissertation, McGill University, Linguistics*.
- Broadbent, D. E. (1958). *Perception and communication*. NewYork: Pergamon
- Chen, C. & He, A. (2008). An analysis of oral segmental errors of advanced English learners——A corpus-based study. *Journal of PLA University of Foreign Languages*, 1, 38-42.
- Chen, H. & Bi, R. (2008). English Majors' Pronunciation and Intonation in Read Speaking: A Longitudinal Study. *Journal of PLA University of Foreign Languages*, 4, 43-49.
- de Bot, Kees (1992). A bilingual production model: Levelt' s “Speaking” model adapted. *Applied Linguistics* 13 (1), 1–24.
- Deterding, David (2006) The pronunciation of English by speakers from China. *English World-Wide* 27, 175–98.
- Dickerson, L., & Dickerson, W. B. (1977). Language variation in applied linguistics. *ITL Review of Applied Linguistics*, 35, 43–66.
- Dowd, J., Zuengler, J., & Berkowitz, D. (1990). L2 Social Marking: Research Issues. *Applied Linguistics*, 11(1), 16-29.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press
- Fu, J. S. (2011). A comparison between Taiwanese ESL learners' perception and production of English pronunciation. *ProQuest, UMI Dissertation Publishing*.

- Gao, H. (2002). How to pronounce the English fricatives. *Journal of LuLiang Education Institute*, 19(4), 34-35.
- Gathercole, S. E., & Baddeley, A. (1994). *Working memory and language*. Hillsdale, NJ: Lawrence Erlbaum
- Hsu, H. (2013). Interaction between markedness and transfer in English final nasals: Taiwanese learning English as a foreign language. *Studies in Media and Communication*, 1, 64-90.
- Hulstijn, J. H. & Hulstijn, W. (1984), grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*, 34, 23–43.
- Hung, T. T. N. (2005) Phonological features of 'New Englishes'. *Paper presented at the the 1st International Conference on the Linguistics of Contemporary English*, Edinburgh
- Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning* 49 (2), 303-342.
- Kormos, J. (2000). The role of attention in monitoring second language speaking production. *Language Learning*, 50 (2), 343-384
- Kraguievac, S. (2011). Factors influencing Serbian learners' production accuracy of English interdental fricatives. *English Language and Anglophone Literatures Today (ELALT)*.
- Labov, W. (1966). *The Social Stratification of English in New York City*. Washington, DC. Center for Applied Linguistics.
- Labov, W. (1970). The study of language in its social context. *Studium Genemie*, 2 (3), 30-W.
- Labov, W. (1984). Field methods of the project on linguistic change and variation. In J. Baugh and J. Sherzer (Eds.), *Language in use: Readings in sociolinguistics* (pp. 28-53). Englewood Cliffs, NJ: Prentice-Hall.
- Lee, S. (2011). Perception and production of English fricative sounds by advanced Korean EFL learners. *Studies in Phonetics, Phonology and Morphology* 17 (2), 259-281.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT.

- Levelt, W. J. M. (1993). Language use in normal speakers and its disorders. In G. Blanken, J. Dittmann, H. Grimm, J. C. Marshall, & C. W. Wallesch (Eds.), *Linguistic disorders and pathologies* (pp. 1–15). Berlin, Germany: de Gruyter
- Lee, S., & Cho, M.-H. (2002). Sound replacement in the acquisition of English consonant clusters: A constraint-based approach. *Studies in Phonetics, Phonology and Morphology*, 8(2), 255–271.
- Lee, J. (2006). Typology of interdental fricatives with reference to loanword adaptation. *Studies in Phonetics, Phonology and Morphology*, 12J, 127–148.
- Lombardi, Li. (2003). Second language data and constraints on manner: explaining substitutions for the English interdentals. *Second Language Research*, 19 (3), 225–250.
- Major, R. C. (2001). *Foreign accent: The ontogeny and phylogeny of second language phonology*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speaking production: evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior*, 21, 578–594.
- O’ Brien, I., Segalowitz, N., Freed, B., & Collentine, J. (2007). Phonological Memory Predicts Second Language Oral Fluency Gains in Adults. *Studies in Second Language Acquisition*, 29 (4), 557–581.
- Owolabi, D. (2012). Production and perception problems of English dental fricatives by Yoruba speakers of English as a second language. *Theory and Practice in Language Studies*, 2 (6), 1108–1113.
- Prince, A., & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar*. Cambridge, MA: MIT Press
- Rau, D. V., Chang, Hui-Huan. A., & Tarone, E. E., (2009). Think or Sink: Chinese Learners’ Acquisition of the English Voiceless Interdental Fricative. *Language Learning*, 59 (3), 581–621.
- Reis, M., Kluge, D., & Bettoni-techio, M. (2007). Different task type and the perception of the English interdental fricatives. *New Sounds 2007: Proceedings of the Fifth International Symposium on the Acquisition of Second Language Speaking*, 408–419.
- Robinson, P. (1995). Attention, memory and the “noticing” hypothesis. *Language Learning*, 45, 283–331.

- Schmidt, R. 1977. Sociolinguistic variation and language transfer in phonology. *Working Papers in Bilingualism*, 12, 79-95.
- Syed, N. A. (2013). Acquisition of English dental fricatives by Pakistan learners. *Newcastle Working Papers in Linguistics*, 19 (2), 59-73.
- Tarone, E. (1982). Systematicity and attention in interlanguage. *Language Learning*, 32, 69-82.
- Tarone, E. (1983). On the variability of interlanguage systems. *Applied Linguistics*, 4, 142–163.
- Tarone, E., & Parrish, B. (1988). Task-related variation in interlanguage. The case of articles. *Language Learning*, 38, 21–44
- Thompson, G. & Brown, A. (2012). Interlanguage variation: the influence of monitoring and contextualization on L2 phonological production. *VIAL*.
- Trofimovich, P., Gatbonton, E., & Segalowitz, N. (2007). A dynamic look at L2 phonological learning: seeking processing explanations for implicational phenomena. *Studies in Second Language Acquisition*, 29, 407- 448.
- Wester, F., Gilbrs, D., & Lowie, W. (2007). Substitution of dental fricatives in English by Dutch speakers. *Language Sciences*. 29. 477-491.
- Wheeldon, L. R. & Levelt, W. J. M. (1995). Monitoring the time course of phonological encoding. *Journal of Memory and Language*, 34, 311-334
- Wu, J. (2008). Comparison between Chinese and English fricatives. *Journal of Mudanjiang Teachers College (Philosophy Social Sciences Edition)*, 5, 69-70.
- Yang, H. (1997). A sociolinguistic study of native Taiwanese speakers' nonstandard pronunciation of the retroflex sounds in mandarin. The University of Texas at Arlington. *ProQuest Dissertations and Theses*, 106-106.
- Yildiz, Y. (2002). *The acquisition of English interdental by Turkish learners*. (Unpublished PhD dissertation). University of Essex. Essex, England.
- Zeamer, C. & Fox Tree, J.E. (2013). The process of auditory distraction: disrupted attention and impaired recall in a stimulated lecture environment. *The Journal of Experimental Psychology: Learning, Memory and Cognition*.