THE PRODUCTION AND PERCEPTION OF ENGLISH SIBILANT FRICATIVES BY MANDARIN SPEAKERS

SEYED SEPEHR SADRIEH

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2014

THE PRODUCTION AND PERCEPTION OF ENGLISH SIBILANT FRICATIVES BY MANDARIN SPEAKERS

SEYED SEPEHR SADRIEH

DISSERTATION SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF LINGUISTICS

FACULTY OF LANGUAGES AND LINGUISTICS
UNIVERSITY OF MALAYA
KUALA LUMPUR

2014

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **Seyed Sepehr Sadrieh** Pasport No: X**95384966**

Registration/Matric No: **TGC080022**

Name of Degree: **Master's of Linguistics**

Title of ~~Project Paper~~/~~Research Report/~~Dissertation~~/Thesis~~ ("this Work"):

The Production and Perception of English Sibilant Fricatives by Mandarin Speakers

Field of Study: **Phonetics**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;

(2) This Work is original;

(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;

(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;

(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;

(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date: 18.7.2014

Subscribed and solemnly declared before,

Witness's Signature Date: 18.7.2014
Name: **Assoc. Prof. Dr. Stefanie Shamila Pillai**
Designation: **Supervisor**

ABSTRACT

The use of English language is mostly restricted to classrooms in China and many Chinese learners have difficulties in the actual use of the language with one of the prominent areas of difficulty being pronunciation. Although there have been previous studies in the area of pronunciation difficulty faced by Chinese learners of English whose first language is Mandarin, there is a lack of research on a common area of difficulty for such speakers when it comes to the pronunciation of English: fricatives in general and English sibilant fricatives in particular. This study attempts to fill this gap through two main tasks in order to investigate and compare the production and perception of English sibilant fricatives by eight male and eight female native Mandarin speaking college students from China. In the production task, the participants were recorded reading a list of 31 English words with the four sibilants in different word positions. Five native speakers of standard American English then rated the sibilants produced by the participants in relation to the degree of 'nativeness' of their pronunciation. The results indicate that the participants had the most difficulty with /ʒ/ followed by /z/, /ʃ/ and /s/. In the perception test, the same Chinese participants listened to and selected the odd item in each instance of the perception task. The results indicate that the participants had the most difficulty distinguishing English /s/ from Mandarin /s/. However, they correctly picked most of the odd items out in words containing the rest of the sibilants. This suggests that the participants' perception of English sibilant fricatives was generally better than their production of the same sounds except for /s/. This can be attributed to the fact that the difference between English and Mandarin /s/ is only phonetic and hard to detect for both English and Mandarin L1 speakers. Based on Flege's Speech Learning Model (1995), this could be the result of filtering out of L2 sound features through Ll phonology due to the fact that these features are phonetically important but not phonologically so. The findings of this study help us understand the

nature of Chinese language leaners' production and perception of English sibilant fricatives and also contribute to the growing body of research on the production and perception of these sibilants in different varieties of English.

ABSTRAK

Penggunaan Bahasa Inggeris di China masih terhad kepada penggunaan di dalam bilik darjah. Ramai pelajar China menghadapi kesukaran dalam penggunaan Bahasa Inggeris khususnya dari segi sebutan. Walaupun terdapat kajian ke atas pelbagai kesukaran sebutan Bahasa Inggeris bagi pelajar China yang berbahasa ibunda Mandarin, namun tidak banyak kajian ke atas kesukaran lazim yang dihadapi oleh penutur tersebut: sebutan frikatif, khususnya frikatif bunyi berdesis. Kajian ini bertujuan memenuhi jurang tersebut dengan dua tugasan utama, mengkaji dan membanding hasil dan persepsi frikatif bunyi berdesis Bahasa Inggeris oleh lapan lelaki dan lapan perempuan pelajar kolej yang berbahasa ibunda Mandarin. Peserta direkod membaca 31 perkataan Bahasa Inggeris yang mengandungi empat frikatif bunyi berdesis dalam kedudukan perkataan yang berlainan. Lima penutur jati Bahasa Inggeris variasi America menilai bunyi berdesis yang dihasil oleh peserta berdasarkan tahap keaslian sebutan. Keputusan yang diperolehi menunjukkan peserta menghadapi kesukaran dalam /ʒ/, /z/, /ʃ/ and /s/ menurut tahap kesukaran. Bagi ujian persepsi, peserta dengar dan pilih perkara berlainan dari setiap tugasan persepsi. Keputusan ini menunjukkan peserta menghadapi paling banyak kesukaran dalam pembezaan /s/ Bahasa Inggeris dengan /s/ Mandarin. Walau bagaimanapun, mereka berjaya memilih dengan betul kebanyakan perkara berlainan di kalangan perkataan yang mengandungi bunyi berdesis lain. Ini bermaksud persepsi peserta terhadap frikatif bunyi berdesis Bahasa Inggeris secara umumnya, adalah lebih baik daripada hasil sebutan bagi semua bunyi kecuali /s/. Ini adalah kerana perbezaan /s/ di antara Bahasa Inggeris dan Mandarin hanyalah dari segi fonetik dan adalah ia sukar dikenalpasti oleh kedua-dua penutur jati L1 Bahasa Inggeris dan Mandarin. Berdasarkan model *Speech Learning Model* oleh Flege (1995), ini mungkin disebabkan oleh sebab saringan ciri-ciri bunyi L2 dari fonologi Ll kerana kenyataan ciri-ciri ini penting dari segi fonetik dan kurang penting di dalam fonologi. Hasil kajian

penyelidikan ini membantu kami memahami latar belakang hasil sebutan Mandarin dan persepsi frikatif bunyi berdesis Bahasa Inggeris dan menyumbang kepada perkembangan kajian produksi dan persepsi bunyi berdesis pelbagai variasi Bahasa Inggeris.

ACKNOWLEDGEMENTS

First and foremost, I would like to show my greatest appreciation to my advisor Prof. Dr. Stefanie Pillai for her support of my mater's dissertation and also for her patience, encouragement, and insight without which I would not be able to finish this research. I would also like to express my gratitude to my family and friends who believed in me and supported me all throughout this journey and whose prayers of support I felt even though they were not all here with me.

I am deeply indebted to Steve, Nicole, Joshua, Emily, and William as well for their invaluable help in rating the data. It was a long and difficult process but they did it with great precision and patience. And last but not least, I would like to thank all the participants who took part in this study and everyone else who aided me along the way.

TABLE OF CONTENTS

`

`

LIST OF FIGURES

LIST OF TABLES

`

CHAPTER 1

INTRODUCTION

1.1 English in China

Pronunciation is a source of concern for many EFL learners who often associate mispronunciation with misunderstanding and embarrassment. Those with accurate pronunciation may be considered as more professional while not being able to speak a second/foreign intelligibly can bring about social and financial consequences; this is especially true in the case of students and immigrants to English-speaking countries. As an example, Da´vila, Bohara & Saenz (1993) discovered that foreign accentedness negatively affects overall earnings. Fluency and intelligibility in a second language (L2) can help people integrate faster into the culture of that L2. However, that is not the case in some EFL contexts. For instance, not a lot of attention has traditionally been paid to teaching pronunciation in many English classrooms in China. In relation to this, one problematic area for L2 learners of English from China is the pronunciation of English consonants including many fricative sounds.

The English teaching context in China is not very easy to describe. China is a huge country with an enormous population. This has resulted in some very big differences between language teaching in major cities versus small cities and rural areas where there is often a lack of enough qualified teachers. Nevertheless certain themes and important periods in the history of ELF in China can be recognized. This is because some of the most prevalent methods and beliefs about foreign language teaching (including the teaching of pronunciation) in China since the first time English was taught there in the middle of the 19th century are still upheld in one form or another by some teachers in certain areas there, as will be seen. This, in turn, can inform us about the needs of English language learners from China as relevant to the area of pronunciation. Most of the English teaching approaches and methods that have been used in China were tested on small groups of English L2 learners for relatively short periods of time without producing any tangible results. A few of these methods, however, have been used for a much longer time.

One such method is Grammar Translation which was most widespread during the first few decades of the 20th century in China when the Chinese education system was modeled on the Japanese one. It has also been the most extensively used method across much of the history of EFL there and was the result of a combination of local Chinese and overseas thinking and practices on education (Hu, 2002). In this method, as the name implies, translation of written language plays an important role in teaching and learning. Grammatical rules and structures are explicitly explained, analyzed and practiced, and vocabulary is learned by rote with the unfortunate result that the aural/oral aspects of the language are almost always neglected. This neglect is mostly due to the fact that Grammar-translation does not require a vast knowledge of L2 language on the teacher’s part.

Another popular EFL approach employed in China is the audio-lingual method. It is important to briefly look at this method here because it was the first time aural-oral practice took center-stage in many EFL classrooms in China and language learners practiced listening to and repeating sounds and their combinations in language laboratories for long periods of time. Originating in the United States in the 50s, ALM was an entirely foreign approach to language teaching in China and was first employed in the higher education system there in early 60s where it was generally accepted. It then rapidly moved down from there to other levels of education. While not all ALM classroom practices were compatible with the traditional approaches in China, others were highly consistent with them. This resulted in the widespread acceptance of ALM in classrooms there. Some teachers even integrated ALM practices into the Grammar-Translation Method and successfully used this combination in Chinese classrooms (Hu, 2002).

Despite the favorable opinion toward ALM in China, this method has not been without its problems (e.g. inability to use classroom skills in real communicative situations) which take us to the last hugely popular ELT approach in China, namely 'communicative language teaching'. CLT has as its primary objective 'communicative competence'; this indicates the importance of teaching pronunciation as well as other skills in this method. CLT first appeared in China in late 70s but it did not generate the same kind of widespread acceptance that ALM did. In fact, it was met with skepticism and apathy at first. There have been on-going discussions among both Chinese and Western ELT experts about the effectiveness of CLT in the context of China. One major outcome of these discussions has been the many attempts at the integration of traditional practices with CLT in different ways (Adamson, 2001). Ultimately higher education teachers in China started adopting CLT in their lessons and this was followed by secondary and primary teachers as well. English textbooks too gradually became more communicatively oriented (Hu, 2002). The introduction of CLT to China meant that there was a lot more pronunciation teaching in context as opposed to the mechanical drills of the audio lingual method.

More foreign language experts started to appreciate CLT in China around 1977 which also saw the end of Cultural Revolution. It was during this time that Policies on education and foreign language learning started to change fundamentally in China as it opened its doors to the international community. The government and people started to realize the importance of learning English and it became one of the important subjects at schools. In 1978, a major foreign language teaching conference was held in China. A few years later, English became a fixed part of the university entrance examination. Since that time, an ever increasing number of people in China have been trying to learn English through the media or various programs. The demand for English has risen exponentially after the turn of the century. Tourists to China are often approached by enthusiastic people, especially younger Chinese generation, who are hoping to practice their English with them. On the other hand, the government has facilitated the process of learning English through a variety of different programs and has set specific goals and linguistic criteria to be achieved at each school grade.

As well as being a compulsory subject in primary and secondary levels, English is a compulsory subject in colleges and universities throughout China too. Not only is English important at schools and colleges in China, but it is also increasingly important when one is trying to find a job and later on for promotion in the work place. According to the Chinese Ministry of Education "College English is not only a language course that provides basic knowledge of English, but is also a capacity enhancement course that helps students to broaden their horizons and learn about different cultures in the world" (cited in Xu & Connelly, 2009, p. 220). This is reflected in a new curriculum developed in China in 2003 which emphasizes the incorporation of five important areas of "linguistic knowledge, linguistic skills, affect, learning strategies and intercultural awareness" (Wang & Lam, 2009, pp.75-76).

In relation to that, the Chinese government has taken some practical steps to achieve the goals mentioned in the 2003 curriculum. For instance, it announced in 2010 that it is compulsory for all state employees under the age of 40 to learn a minimum of 1,000

English phrases by the year 2015 and English teaching must start at kindergarten by that same year. The need for newer language learning models and textbooks is being felt and in order to facilitate that the government is currently organizing extensive teacher training programs in China. In addition to that, the number of private language schools has skyrocketed in recent years and those who have the financial means are opting to send their children (sometimes as young as 2 years old) to such schools (Ward & Francis, 2010). What all of these mean for the teaching of pronunciation in China is that there is an ever increasing emphasis on intelligibility and learning the correct pronunciation in context. However, as was mentioned earlier, this is still not exactly the case in many parts of China where more trained teachers are required. This issue will be elaborated on in the next section.

1.2 Rationale and Statement of the Problem

At present, the teaching of English pronunciation starts at Primary school in China but as was hinted earlier, not every language learner in China has the same opportunities nor do they always receive the same quality of English teaching. These differences mostly appear to be between coastal and inland regions where the former are generally wealthier; similarly, these differences can be found between urban and rural areas too (Nunan, 2003). Shanghai is a very good example of a region in which the amount of exposure to English is much higher than that of many other parts of China with the socioeconomic and educational background of people in Shanghai playing an important role in creating these differences. Additionally, many parents in Shanghai realizing the significance of learning English are willing to take any necessary steps for their children to learn it given the abundance of options and opportunities there (Zou & Zhang, 2011). While the English textbooks at schools in cities such as Shanghai and rural areas might not necessarily be very different due to China’s centralized approach to education, a visit to rural schools gives us a clearer picture of the differences in classrooms in such places and the ones in wealthier urban areas. Through a series of case studies of rural classrooms, Wang (2006, p. 41) concluded that teaching and learning methods have not changed much since a long time ago:

> In class, the principal way in which students learn is by sitting still and listening quietly, while teacher-controlled indoctrination is the main model of instruction. Learning is characterized by repeated reading and rote memorizing. Exams are administered every week, and students are ranked by their test scores. Teaching style is monotonous, and learning is boring.

Considering all of this, it is evident that there is not an equal emphasis on speaking and pronunciation in classrooms across China; while this is the case, a large number of people in China go to other countries (especially to English speaking ones) to continue their education every year. There are also many people who are trying to further their careers and businesses by learning English and reaching out to other nationalities both inside and outside of China. In relation to this, Wei and Ding (2011), in a study of 81 Chinese EFL learners' attitude toward their accent, reported that more than half the learners felt native speakers would respect them more if they had a good English pronunciation which shows they realized the social implications associated with accentedness even though they were not willing to openly admit it. This was further confirmed when all the learners indicated their excitement at the prospect of being able to sound like a native speaker when speaking English.

In relation to that, it is obvious that speaking intelligibly and being understood clearly when conversing with non-Chinese, especially native speakers of English, takes on extra importance for language learners from China. An important part of achieving this objective is to learn the sound system of English language properly. It is important at this point to establish the variety of English used as the benchmark in this study because different Englishes can have very different sound systems and phonemes. The variety of English chosen for this study is Standard American English. There are several reasons for this; first of all, the participants in this study were part of a twinning program which required them to do their last two years of their studies in the United States. In fact, a recent report (Jia, 2013) has shown that the United States is the number one destination for Chinese students from China with about 28% of Chinese students contemplating pursuing higher education abroad going to the United States as opposed to only 18% who opt for United Kingdom which is the second major destination for China students. On top of that, Chinese students from China by far comprise the largest group of international students in the United States too (Song, 2013):

> Overall, the number of international students in U.S. institutions increased by about 7% last year to nearly 820,000. The largest group came from China, which sent about 236,000 students, nearly double the number of students from India, the second-largest group.

Further, despite much talk about English becoming a lingua franca, standard American and British English are still the two popular varieties of English in China. He and Li (2009, p. 79), for example, found that “teachers and learners of college English alike are generally in favor of adopting ‘Standard English’ (most probably British or American English) as the pedagogic model for college English in China”. In Hong Kong, a former British colony, American English – American English accent in particular – is slowly gaining more popularity over British English (Hutchison, 2013). The article quotes Dr. Rodney Jones, the Acting Head of English at Hong Kong's City University, as saying: "There's no doubt that the American accent is becoming more prevalent here. The main reason is because people are more exposed to it”. This desire to sound like a native speaker of English, and in particular American English, is not restricted to students in Hong Kong (Hutchison, 2013):

> Hong Kong recruitment consultant Adam Bell agrees that sounding American can help boost a candidate's employability -- particularly if the job is with a US firm. [Bell states] “There's a degree of prestige associated with both the UK and the US accents compared to a Hong Kong accent as it suggests they are from a good background and can afford to study at school or university abroad”.

The situation in Hong Kong is important for this research when we consider the fact that more and more mainland Chinese go to Hong Kong to study English (Hutchison, 2013). On the other hand, it is believed that “the rise of American idiomatic expressions and the American accent in other areas is often ascribed to the world's increased exposure to American culture, especially through movies, videos, computer games and the Internet” (Hoke, 2013).

Based on what has been mentioned so far in this chapter, an explanation of the way Native Mandarin speakers produce and perceive English sounds is in order as doing so will highlight some of the problematic areas faced by these speakers in pronouncing and perceiving English sounds. One such area of difficulty for the EFL learners in China is the production and perception of English fricatives, and in particular English sibilant fricatives. As will be discussed later in this chapter, not enough attention has been paid to the production and perception of these fricatives. It is, therefore, the aim of the current study to explore these two aspects of English sibilants by native speakers of Mandarin from China, as will be seen in the following section.

1.3 The Objectives of the Study

This study is an attempt to investigate and compare the production and perception of English sibilant fricatives by native Mandarin speaking college students from China with the following objectives:

i. To examine the production of English sibilant fricatives by Mandarin speakers from China.

ii. To examine how these speakers perceive English sibilant fricatives.

iii. To examine if there is a correlation between the production and perception of the sibilant fricatives.

In connection with the objectives, the current research strives to provide answer to these questions:

i. How close to General American English sounds do the participants produce English sibilant fricatives?

ii. How do the participants perceive English sibilant fricatives?

iii. To what extent is there a difference between the students' perception and production of English sibilant fricatives?

Apart from the fact that a large number of Chinese students from China attend American Universities in the United States and the relative popularity of American English among English EFL learners in China, General American sounds were chosen as target sounds in this study because the participants had to spend at least two years studying in the United States. In addition, they had to take and pass Test of English as a Foreign Language (TOEFL) which employs "General American" as the model standard. It should, however, be noted that there is not much difference between sibilant fricatives in General American and Received Pronunciation (RP) – also known as Standard British English – anyway. This is all the more reason why these areas of difficulty should be addressed in English classes in China regardless of which variety (American or British) is taught. Consonantal differences between General American and Standard British English will be discussed in Chapter 2, section 2.2.

1.4 The Significance of the Study

There are not a lot of published studies on English pronunciation of fricatives by native speakers of Mandarin from China and among these, very few have discussed the relationship between the production and perception of the fricatives, in particular the sibilant fricatives. This study will contribute to the body of knowledge on the production and perception of English sibilant fricatives as produced by Mandarin speakers from China. In addition, this study attempts to fill the research gap by comparing the results of the production and perception of the sibilants. The findings of this study provide insights into the way native Mandarin speakers from China produce and perceive English sibilant fricatives and how close their production is to the American English variety. Such knowledge can potentially lead to an improvement in the teaching and learning of pronunciation to Chinese language learners from China.

1.5 The Limitation of the Study

Every research study is bound by limitations that can potentially affect the results of the study. The first limitation here is the sample size of 16 participants. The next limitation is related to the fact that a list of 31 words was used to elicit the productions. Having included any additional reading material to elicit more productions would most probably take this research beyond the scope of a master’s dissertation. To sum up, even though the findings of this study reveal certain patterns in the production and perception of English sibilants among students in China, they cannot be generalized from the small sample here to the entire student population in China.

1.6 Organization of the Study

This study contains five chapters. The first chapter mainly discusses the objectives of the study. Related literature is reviewed in the second chapter and in the third one the methodology for obtaining and analyzing the data is detailed. The findings of the study will be presented and discussed in Chapter Four and finally Chapter five summarizes the findings and their significance as relevant to teaching and learning of English by native Speakers of Mandarin from China.

CHAPTER 2

LITERATURE REVIEW

The first main section of this chapter tackles the issue of foreign accents related to the production and perception of non-native speech sounds followed by a detailed discussion of related theoretical framework. A discussion of fricatives, in particular English and Mandarin fricatives, will follow next. Finally, related studies on the production and perception of fricatives (in particular English and Mandarin fricatives) will be reviewed.

2.1 Foreign Accent and L2 Production and Perception

Reaching an "acceptable" and "intelligible" level of pronunciation has always been one of the major goals of language learners and despite the fact that many such learners have managed to successfully pick up aspects of English language such as grammar, vocabulary, writing, and reading as close to native speakers' as possible, not as many have been able to do the same with pronunciation. The issue with pronunciation is that those who speak English with an exotic accent, especially in English speaking countries, apart from potential misunderstandings and breakdowns in communication (as was mentioned in Chapter 1), might get stigmatized (Rajadurai, 2007) which further explains why General American was chosen in this study as the target model; however, a more basic question which has fairly recently been poised by some researchers in the area is: 'Why is there a need for native-like model in the teaching of English?' Such studies often point to the fact that English has become a Lingua Franca and is no longer only widely spoken in English speaking countries; as a result, they claim, many people actually use English in communication with non-native speakers (see for example Jenkins, 2000, 2002). However, the picture in this study is not exactly the same. For one thing, the participants in this study needed to communicate with native speakers of American English both in and outside of an academic setting. More generally, however, as was pointed out in chapter one, a much larger number of Chinese students find their way to colleges and universities in the United States in comparison to other countries and they also form the largest group of overseas students in there, hence the use of a native-speaker model in the current study.

The more important aspect of this research, however, deals with individual L1 and L2 sounds and includes areas such as the potential constraints of L2 speech learning as opposed to L1 speech learning, the plausibility of such constraints, differences in production and perception constraints, and the possibility of L2 learners being destined for a difference in comparison to native speakers. These are the areas for which a plethora of proposals have been offered. In particular, there has been much research investigating the production and perception of L2 phonetic segments. We will look at these issues in detail in the next few sections.

2.1.1 Speech Production

What is prevalent in the production of language learners is a divergence from phonetic norms of the second/foreign language. Indeed many studies have shown this in the production of L2 consonant and consonant clusters (see Leather & James, 1996, for review). One of the early theories that addressed this issue, though, was contrastive analysis hypothesis (CAH) which stated that cross-language differences result in learning difficulty. Lado (1957, preface) maintained that:

> …we can predict and describe the patterns that will cause difficulty in learning, and those that will not cause difficulty, by comparing systematically the

> language and the culture to be learned with the native language and culture of the student.

What this meant for learning L2 speech sounds was that, L2 categories that have no equivalent in the L1 are more difficult to learn than those that sound similar to an L1 sound. And although there have been some studies supporting this view of CAH (e.g. Broselow, 1984; Erdmann, 1973; Lehn & Slager, 1959), it has failed to properly account for many observations in L2 phonetic segment learning. Klein (1986, p. 25) suggests:

> A major reason for this relative failure lies in the fact that structural similarities and dissimilarities between two linguistic systems and the processing of linguistic means in actual production and comprehension are two quite different things. Contrastive linguistics was concerned with the former; acquisition, however, has to do with the latter. It is not the existence of a structure as described by the linguist that is important, but the way the learner deals with it in comprehension and production. Therefore, comparison of structures may totally miss the point.

One of the debates that has often come up in the study of L2 production, especially after CAH started to lose its appeal, is the notion of critical period which centers around the question 'is there a noticeable difference in the L2 production of the learners who learn the second/foreign language in childhood versus the ones who learn it in adolescence and adulthood in terms of how close they are to the phonetic norms of the L2?' Studies investigating this difference have mostly found that early learners, as they are often called, have an advantage over late learners in this regard (see Long, 1990, for a review). It has, therefore, been suggested that the critical period resulting from reduced ‘neural plasticity’ has a negative effect on the ability of late learners to master L2 speech sounds and pronunciation (McLaughlin 1977; Patkowski 1989). The proponents of this hypothesis believe that speech learning and indeed language learning in general becomes increasingly difficult after the critical period. DeKeyser (2000, pp. 518-519), for instance, proposed that:

> Somewhere between the ages of 6-7 and 16-17, everybody loses the mental equipment required for the abstract patterns underlying a human language, and the critical period really deserves its name … It may be that the severe decline of the ability to induce abstract patterns implicitly is an inevitable consequence of fairly general aspects of neurological maturation and that it simply shows up most clearly in language acquisition.

It should be mentioned that when we talk about critical period hypothesis, there is not always a clear distinction between production and perception in the literature but as Scovel (1988, p. 62) points out:

> Pronunciation is the only part of language which is directly “physical” and which demands neuromuscular programming. Only pronunciation requires an incredible talent for sensory feedback of where the articulators are and what they are doing. And only pronunciation forces us to time and sequence motor movements. All other aspects of language are entirely “cognitive” or “perceptual” in that they have no physical reality.

What this implies is a distinction between segmental production and perception when it comes to the limitations that the critical period imposes on the two. It was proposed by Bever (1981), for example, that organizing and maintaining a mutual relationship between segmental production and perception is done through a “psychogrammar” which is employed in the process of L1 acquisition for the purpose of creating "conjoint" representations of perception and production. According to him, what indicates the closure of the 'window of opportunity' at the end of the critical period is the fading of the psychogrammar which is in turn due to completion of L1 phonology acquisition. From then on, language learners "often learn to discriminate sounds ... they cannot distinctively produce" because the psychogrammar is not there to align speech production and perception anymore resulting in a divide in the development of the two Bever (1981, p. 196).

The connection between segmental production and perception has been further discussed by others. According to Pisoni (1995), while being related to each other in a "complex" way, production and perception represent features of a "unitary articulatory event". Pisoni hypothesized those acoustic differences critical in the process of perceptual analysis are also employed by speakers in production. He also claims that category systems of speech production and perception are in a "unique" relation not seen among other systems. However, this last point (alignment of perception and action) might actually indicate a general feature of the way brain works. Edelman (1989, pp. 54-56), in his theory of neuronal group selection states that "dynamic loop ... continually matches gestures and posture to several kinds of sensory signals," which means perception "depends upon and leads to action" with the result that motor activity can be considered as "[an] essential part of perceptual categorization". Further, according to Churchland (1986, p. 473) "evolution [has] solved the problem of sensory processing and motor control simultaneously" so that "theories [must] mimic evolution and aim for simultaneous solutions as well".

In another observation Kuhl and Meltzoff (1996, p. 2425) pointed to a need for "exquisitely detailed" information which exactly defines auditory-articulatory connections; they further claimed that adults might even possess an "internalized auditory-articulatory 'map' that specifies the relations between mouth movements and sound". Having mentioned this, they concede that "formation of memory representations ... derives initially from perception of the ambient input and then acts as guides for motor output". What this means is that there is an asymmetry between auditory-articulatory aspects early on in the acquisition of L1 sounds and phonology. It was proposed by Flege (1995) that the accuracy of L2 learners' productions is constrained by how accurately they perceive L2 sounds. He went on to say that perceptual representation of L2 phonetic segments correspond with their productions in the sense that the latter is usually not more native-like than the former. In fact, productions of beginner L2 learners can even be less native-like compared to their perception. As a case in point, Rochet (1995), using a synthetic French continuum, /i/-/y-/u/, investigated the perception of these sounds by Brazilian Portuguese and Canadian English speakers with the result that the Portuguese speaking participants often identified and produced /y/ quality as /i/ while the English participants identified and produced the same sound with /u/. Based on the results, Rochet (p. 404) attributed these errors to "the target phones having been assigned to an L1 category."

2.1.2 Speech Perception

The constraining effect imposed by the native language experience on the perception of L2 or foreign language phonetic segments, especially consonants, has been documented in many studies almost all of which point to an influence by L1 phonological system. For example, a number of studies have investigated the effects of Japanese and Korean L1 on the perception of American English approximant contrast /ɹ/-/l/ (e.g. Gillette, 1980; Ingram & Park, 1998; MacKain, Best & Strange, 1981; Miyawaki et al., 1975; Sheldon & Strange, 1982; Takagi & Mann, 1995; Yamada & Tohkura, 1992). They found that there is no ‘contrastive function’ between /ɹ/ and /l/ in Japanese and Korean. However, they found that the contrastive function exists for the two other English approximant contrasts /w/-/ɹ/ and /w/-/j/. Trubetzkoy (1939/1969), argued that those acoustic differences that are phonemic in L2 (but not so in L1) are "filtered out" as a result of the phonology of L1. The L1 phonological system acts as a kind of “sieve” which L2 sounds must pass through. And this is why, according to him, Japanese speakers cannot discriminate English /ɹ/ from /1/. In a similar study, Michaels (1974) looked at the pronunciation of English /θ/ by Russian and Japanese learners and found that while Russian speakers replaced this sound with /t/ (a non-strident sound), Japanese speakers replaced it with /s/ (a continuant sound) despite the fact that both /t/ and /s/ exist in Russian and Japanese. Based on this, he concluded that the Russians substituted the closest non-strident Russian sound, /t/, because they perceived “non-stridency” in English /θ/ while the Japanese perceived “continuancy” in English /θ/ which prompted the use of /s/ in its place. This shows that what influences L2 segmental perception by different learners is the varying distinctive features in their first languages.

In another study, Flege and Port (1981) investigated the production of /p/ (a sound that does not exist in Arabic) in the speech of a group of adult Saudi Arabians. Flege and Port noted that even though they stayed in the United Sates for a few years, they still produced /p/ with the closure voicing similar to /b/. We have to bear in mind here that all the distinctive features of English /p/ separately exists in Saudi Arabian Arabic sound system (it has a voiced bilabial stop /b/ and voiceless stops, /t/ /k/); what this study shows, then, is the non-commutability of distinguishing properties because if that was the case, Saudi Arabian speakers would be able to produce English /p/ with little difficulty. As to the nature of this difficulty, Flege and Port pointed to the participants' difficulty in adjusting the glottal and supraglottal gestures necessary for the production of /p/. It is also possible that they did not accurately perceive the features of English /p/. Employing young listeners, Sebastián-Gallés and Soto-Faraco (1999) found that acoustic properties of speech are modified by children in alignment with their L1. And these modifications will later in life act as a "sieve" through which L2 speech passes. It is possible for the modifications to get realigned so that non-native phonetic segments can be processed but that is dependent upon the age at which individuals are exposed to the L2.

Specifically focusing on adults, though, the question is, does age hinder adults' perception of contrast in non-native speech sounds? Basically, many studies have documented strong L1 effect on the ability of adults to perceive non-native speech sound contrasts (e.g. Abramson & Lisker, 1970). However, there are different accounts of this phenomenon purposed by researchers. The traditional view as was mentioned earlier in this section points to a sieving effect from native-language when an adult individual tries to produce non-native speech sounds (Polivanov, 1931; Trubetzkoy, 1939/1969). Aside from the widely cited example of adult Japanese speakers having difficulty discriminating English /r/-/l/ (also mentioned earlier in this chapter), researchers have also pointed to non-native speech perception difficulty among speakers of other languages. As an example, Werker, Gilbert, Humphrey & Tees (1981) demonstrated the difficulty of English speakers in perceiving the contrast between Hindi retroflex versus dental stops.

This apparent difficulty faced by adults has led some to point to a lack of exposure to an L2 early in life as the cause. According to such researchers, this early exposure is the key because the sensori-neural mechanisms responsible for speech sound discrimination are tuned mainly during early life. Eimas (1991), for example, claimed that the sensori-neural mechanisms for perception start to get tuned to universal settings during infancy before they become specialized as a result of exposure to group of speech sounds. Another group of researchers (see for example Aslin & Pisoni, 1980), on the other hand, believe that infants come prewired with certain “psychophysical” mechanisms (no longer active in adults) which are tuned and reinforced when infants are exposed to certain acoustic features in the environment. This in turn causes a particular respond to those acoustic features. In that sense, such researchers have taken a rather acoustic approach than a linguistic one in an attempt to explain young children’s seemingly effortless process of mastering and discriminating the speech sounds that they are exposed to.

Such accounts, however, have failed to give a complete picture of non-native segmental perception by adults. Many studies (see for example, Lively, Logan & Pisoni, 1993; MacKain, Best & Strange, 1981; Strange & Dittmann, 1984) have demonstrated the ability of adults to perceive non-native segmental contrasts which are not only attributed to rigorous training but also to a large amount of exposure in everyday life. This implies that early life exposure is not a must in segmental perception and discrimination. This led some researchers (e.g. Werker & Tees, 1984) to propose that exposure to an L2 mainly engages higher-level processes (e.g. phonological encoding or memory retention) which are adaptable even in adults. This is contrasted with lower-level sensorineural processes which will not remain as adaptable for long and are largely affected by language experience early in life.

Additionally, some studies have shown that whether an individual is exposed to certain speech sounds or acoustic features early in life or not, does not necessarily affect the accuracy of perception and discrimination of those sounds one way or another. Abramson and Lisker (1970), for example, demonstrated that even though voicing contrast exists in American English stops, American English listeners had a hard time discriminating voicing contrast in non-native stops. In another study, Polka (1991), investigated the Native English speakers' perception of Hindi dental-retroflex stops with different voicing features and found that the English speakers' ability to discriminate the stops varied considerably from not being able to discriminate at all to good discrimination. Contrary to the results of the last two studies just mentioned, Best, McRoberts & Sithole (1988) found that Zulu click consonants were perceived and discriminated easily by American English listeners. What these three studies (among others) show is the unreliability of age in making generalizations about the ability of individuals (especially adults) to perceive and discriminate non-native segmental speech sounds.

So now the question is, as the traditional models fail to explain adults' discrimination of non-native speech sounds, what *does* explain this discrimination? In order to account for this phenomenon, a number of theoretical models have been proposed which basically point to prior L1 linguistic knowledge as affecting the ability of adult listeners (either positively or negatively) in discriminating non-native speech sounds through a perceptual framework. We will look at three of these frameworks next.

2.1.2.1 Perceptual Assimilation Model

In a bid to explain the accurate discrimination of Zulu clicks by English listeners (as mentioned in the previous section), Perceptual Assimilation Model (PAM) was introduced (Best, 1994a, 1994b, 1995; Best et al., 1988). It was hypothesized that this phenomenon was a result of the listeners perceiving the clicks as non-speech (Best et al., 1988). Two of the important features of PAM involved in its hypotheses are: 1) the fact that it predicts the accuracy with which different listeners with different L1s should be able to assimilate and discriminate non-native speech sounds and 2) making use of articulatory phonology in this process.

PAM emphasizes listeners' experience with L1 "phonological equivalence classes" as heavily influencing their perception of non-native speech sounds. It further maintains that non-native sounds are often assimilated into native sounds according to the shared features of the two sounds (Best, 1994a, 1994b, 1995). Best suggested six different perceptual assimilation patterns in her model in order to explain the different ways non-native sounds are assimilated into L1 sounds. These six patterns, therefore, reveal the level of difficulty in L2 contrasts relative to the L1 categories (cited in Altmann & Kabak, 2011):

1. Two-category assimilation (TC) – there are two L2 phonemes, each assimilated into a different L1 sound; dissemination is excellent; as an example, English alveolar /s/ and /t/ are assimilated into Persian dental /s/ and /t/ categories, respectively.

2. Category-goodness difference assimilation (CG) – the two L2 phonemes are assimilated into a single L1 category but one is perceived as closer to the L1 sound than the other. Discrimination is normally moderate to very good. For instance, L1 speakers of Spanish assimilate English /ɪ/ and /i/ to Spanish /i/ but perceive English /i/ as a better example of Spanish /i/.

3. Single-category assimilation (SC) – the two L2 phonemes are assimilated to a single L1 sound and are perceived as equal in terms of how well or how poorly they resemble the L1 sound. Discrimination is normally not that good. As an example, /æ:/ and /ɑ:/ in American English are assimilated into a single Japanese phoneme, /a:/; both American English vowels here are considered equally good instances of the Japanese phoneme.

4. Both uncategorizable assimilation (UU) – the two L2 phonemes are phonologically possible in L1 but there are no L1 categories they can assimilate to and so discrimination varies from poor to very good. For instance, English dental fricatives /θ/ and /ð/ are uncategorizable for L1 speakers of Arrente (an Australian language) which does not have fricatives but includes dental place of articulation for non-fricative contrasts.

5. Uncategorized versus categorized assimilation (UC) – only one of the two L2 phonemes is assimilated to an L1 sound. Discrimination is normally very good. For example, some Danish and French listeners perceived the two Norwegian /ʉ/-/u/ as a two-category or categorized-uncategorized contrast in their L1s.

6. Nonassimilation (NA) – Neither L2 phonemes is phonologically possible in L1. Discrimination can range from good to very good. An often cited example is the Zulu clicks for native speakers of English.

2.1.2.2 Native Language Magnet

Another speech perception model introduced by Kuhl and Iverson (Iverson & Kuhl, 1996; Kuhl, 1991, 1992; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992) is the Native Language Magnet (NLM). The NLM emphasizes early language experience as a very important factor in perceiving acoustic features of speech sounds which will ultimately result in changes to perception and also production of spoken language. Through perception of spoken language and the frequency of speech sound properties therein, infants organize phonetic segments into categories. The outcome of this process is the formation of a mapping between L1 speech sound categories and the spoken language they hear around them. This mapping is, therefore, language specific. As an example of how this works, Kuhl et al. (1992) looked at how two groups of infants at the age of 6 months, one being brought up in Sweden and exposed to Swedish and the other in an English speaking environment in the United States differed in their perception of a number of synthetic high front vowels. Swedish infants' responses indicated that they perceived Swedish /y/ and its variants to be the same sound in significantly more instances than they perceived English /i/ and its variants to be the same. And while this was the case for Swedish infants, the opposite held true for English infants. Among other findings, this study demonstrated the alteration of phonetic perception through linguistic experience. According to Kuhl (2000, p. 11854) the mapping infants hear around them alters "the acoustic dimensions underlying speech, producing a complex network, or filter, through which language is perceived". This adjusting to native language categories through perception has the potential to influence the perception of L2 speech sounds later in life.

Further, acquiring two languages early in life can lead to the formation of two different mappings without much interference effects; however, individual learning a second language later in life might have trouble in separating the two mappings because a neural "commitment" to the familiar mapping which is already in place for L1 may shape L2 speech sounds processing. As might have been inferred from the discussion on

NLM so far, this model attributes L2 perception difficulty to L1 experience and not as a function of brain plasticity level.

In a study supporting the claims of NLM on the nature of differences between L1 and L2 perception, Iverson et al. (2001) studied the perception of English /ɹ/ and /l/ by a group of native English and a group of native Japanese adults residing in Japan. The stimuli consisted of synthesized /ɹa/ and /la/ tokens varying only in F2 and F3 transition frequencies and resembling the occurrences of such tokens in an American English speaker's natural speech. The participants were first asked to listen and identify each stimulus based on the phonemes of their own language and then rate the same stimulus in terms how close they thought it was that to the category that they first identified it with. The results of multidimensional scaling analyses for the Japanese listeners pointed to a distortion in perceptual space along the acoustic-phonetic dimensions for these phonemes. The differences in F3 are critical in order to successfully perceive the contrast between English /ɹ/ and /l/ (as was the case for English listeners). However, the Japanese listeners, ignoring those differences due to a warped perceptual space, mostly relied on acoustic variation and F2 frequency which are largely irrelevant for discriminating between English /ɹ/ and /l/. Based on this, Iverson et al. noted that perceptual maps used by L1 speakers of Japanese may hinder discrimination of English /ɹ/-/l/ (L1 interference effects) which led the authors to conclude that overemphasis on F2 frequency variations to the neglect of F3 might actually cause the learners to form "erroneous" long-term memory if and when they develop new categories for these sounds. Iverson et al. also argued that as individuals acquire their first language, interference effects may become increasingly stronger.

For those Japanese adults who are not able to perceive the same "auditory distribution" of F3 differences in English /ɹ/ and /l/ the same way young learners acquire English as native language do, the interference effects may get automatically reinforced. This, however, does not mean that adults can never perceptually learn L2 speech sounds. Kuhl (2000) claimed that children acquiring two languages at the same time in early childhood may have little influence due to past experience, provided that "two different mappings" for the speech sounds of the two languages are acquired. Kuhl further proposed that the best way adult L2 learners can bypass interference effects might be to get "exaggerated acoustic cues, multiple instances by many talkers, and massed listening experience" which is something that infants regularly experience when acquiring their L1 speech.

In order to recap, two broad theories were elaborated as to why L2 learners often speak with a foreign accent. The first one proposed the presence of a critical period beyond which accurate speech production may be hampered. And the second one claimed that how accurately learners produce L2 speech sounds is a function of how close they are to the perceptual representations developed by the learners for L2 phonetic segments. Consequently several proposals have been put forth as to why perception of L2 speech sounds is sometimes not accurate. These hypotheses all revolve around the idea that L2 learners might not always have access to features or properties necessary for creating accurate perceptual representations. However, an important issue remains here, and that is whether the effects of the constraints described in the literature are permanent in regards to L2 perceptual learning putting aside all the other variables of age, L2 learning context, or the amount and type of L2 exposure. Best and Strange (1992, p. 327) proposed that a "reorganization of perceptual assimilation patterns" might be possible through L2 experience resulting in changes in discriminability. But they did not give details as to how much change is feasible or what the conditions for the alteration of perceptual assimilation patterns might be.

2.1.2.3 Speech Learning Model

Speech Learning Model (SLM) was introduced by Flege (Flege, 1986, 1990, 1995). What is special about this theoretical model is it specifically concentrates on non-native speech learning with the major goal of explaining the changes that happen in segmental learning (both in production and perception) over the life span of individuals. This model identifies two general assumptions: 1) Bilinguals' L1 and L2 phonetic subsystems cannot exist completely independently from each other because they share a "common phonological space", and 2) The capacities leading to successful acquisition of L1 speech stay unchanged throughout individuals' life span. These capacities are: accurately recognizing featural patterns in spoken language, organizing a broad array of phonetic segments with shared features into categories, and the ability to associate speech production to the features recognized in speech input. This second assumption, however, is controversial and is in contrast to the critical period hypothesis.

The notions of L2 speech filtering or warping are not refuted by the SLM. The filtering out of the phonetic features used in the discrimination of L2 (but not L1) speech sounds seems to be a plausible argument. Munro (1993), for example, looked at a group of male L1 speakers of Arabic who learned to produce English /i/ and /ɪ/ accurately with a "native-like spectral difference". The closest vowels to English /i/ and /ɪ/ is Arabic /i/ and /i:/ which are spectrally different from the English pair. English /i/ and /ɪ/ are tense and lax vowels but the participants did not seem to perceive that and instead relied mainly on temporal difference between the two as though "phonologically long and short Arabic vowels" were being produced.

Some studies (e.g. Flege, 1984) have shown that phonetic differences in other languages might sometimes be perceived by untrained listeners, and also that minor deviations from phonetic norms of an L1 can be detected by adult speakers of that L1. In view of such observations, the SLM proposed that as individuals learn a detailed network of L2 lexical items that needs phonetic discrimination, the filtering of spoken L2 starts to fade. For example, in a study involving L1 English and Spanish speakers who had lived in Stockholm for a long time, McAllister, Flege & Piske (2002) observed that although the main cue to distinguishing vowels in English and Spanish is not vowel duration, the participants learned to discriminate Swedish words that had distinct phonological quantity.

Some studies (e.g. Francis & Nusbaum, 2002; Gottfried & Beddor, 1988) also pointed to a change in feature weighting as a result of L2 learning. Gottfried and Beddor, for example, studied the perceptual effects of synthetic versions of French /o/-/ɔ/ differing orthogonally in temporal and frequency formants (F1, F2). The authors found that while native French speakers did not rely on vowel duration as a cue in the perception of the vowels, English speakers (with no prior knowledge of French) were sensitive to the vowel difference between the two; this is because vowel duration plays a more important role in the identification of vowels in English than in French. English speakers, who had some prior experience with French, also relied on duration of French vowels when discriminating them. On the other hand, the native speakers of English who had an advanced knowledge of French language were more similar to the French speakers in their perception of the two vowels by not relying on the vowel duration very much.

Contrary to the SLM, Best and Strange (1992) proposed that learners assign a new category to those L2 sounds that are recognized as poor instances (but nevertheless as instances) of an L1 category rather than to those recognized as remote to the closest L1 sound. The SLM, on the other hand, believes that the more distant an L2 sound is from the closest L1 sound, the higher is the possibility of forming a new category for an L2 sound. Flege (1987), for instance, noted that French /y/ was easier for adult L1 English learners of French to produce than French /u/. Flege concluded that this was due to French /y/ being perceived by the learners as having greater phonetic dissimilarity to the nearest English vowel compared with French /u/ to the nearest English vowel.

As previously mentioned, the NLM emphasizes L1 phonetic acquisition rather than a loss of neural plasticity as a source of difficulty for L2 learners in perceiving and discriminating non-native speech sounds. This is a view that is shared by SLM too. SLM also maintains that the capacities infants and children have in successful L1 speech acquisition is preserved by adults who are able to form new phonetic categories for L2 speech sounds they hear around them. SLM, however, acknowledges that this ability decreases with age. What is more, with the gradual development of L1 phonetic categories from childhood into early adolescence, there will be a higher possibility for the assimilation of L2 speech sounds into the existing categories. If any of the L2 consonants or vowels gets continually perceived as an L1 sound, no new category will be formed for it. A possible shortcoming of this model, however, is that it does not offer a way to measure cross-language phonetic differences in order to predict when those differences are big enough to cause the establishment of new categories, and what role, if any, age or L1 system development play in this regard.

To test the proposal that L2 speech sounds get assimilated into existing L1 categories as they develop, Baker, Trofimovich, Mack & Flege (2002) conducted a perceptual assimilation experiment looking at English vowels. The participants were L1 speakers of Korean (adults and children) who had resided in the United States for 9 months. The findings indicated that Korean adults assimilated English vowels into Korean vowels more often than did Korean children. Further experiments were carried out investigating perceptual assimilation by L1 speakers of Koreans residing in the United States with a length of stay of about nine years and an average arrival age of nine (early learners) and nineteen (late learners). One of the experiments involved English vowels /i/-/ɪ/, /ɛ/-/æ/ and /u/-/ʊ/. These three pairs had each been assimilated into a single Korean vowel. Employing a categorial discrimination test, the authors found that the early learners showed a more accurate discrimination of the English vowels as compared to the late learners. What is more, the early learners' discrimination was not significantly different from that of L1 English speakers. These findings were supported by a second experiment which investigated the production of the same six English vowels. The authors also observed some between-group differences in the production and perception of these vowels which they attributed to the degree of perceptual assimilation of the vowels into Korean vowels stemming from the differences in the age of the participants.

As was discussed earlier, infants start to adjust into L1 speech sounds. It is generally believed that the acquisition of L1 phonemes happens at around age eight but it seems that the two aspects of 'speech motor control' and 'perceptual representations of native phonetic segments' will not complete until adolescence (see Johnson, 2000; Walley & Flege, 2000). However, the exact age at which L1 speech reaches its full development is still not clear. This means that the critical period hypothesis with its emphasis on brain maturation as affecting L2 speech learning on the one hand and the SLM or the NLM with their developmental view of age effects on the other would be hard to separate from each other if indeed it is found that the age at which L1 speech reaches its full development happens around the same age when the critical period for the acquisition of a second language comes to an end (Scovel, 1988; Patkowski, 1989). There is, however, one way to separate the two views and that is through examining the effects that L2 learning has on the production and perception of L1 speech sounds.

According to the SLM, there is inevitable interaction between bilingual individuals' L1 and L2 phonetic subsystems because the phonetic segments that comprise the two subsystems occupy a shared phonological space. The SLM maintains that the way bilinguals try to distinguish between phonetic segments in the L1 and L2 subsystems closely resembles the way phonetic segments are distinguished within a single system. The SLM proposes that the interaction between phonetic categories is facilitated by means of a couple of mechanisms named "phonetic category assimilation" and "phonetic category dissimilation". When an L2 phonetic segment gets assigned a new category in a phonetic space belonging to an L1 speech sound, the dissimilation of the two categories might occur in which case the new category and the pre-existing one will not be exactly the same as the categories acquired by monolinguals. Phonetic category dissimilation is something that has been discussed in previous hypotheses advocating a critical period or L2 sound filtering/distorting.

SLM proposes that category assimilation happens where no new category for an L2 phonetic segment is formed (provided that there is an audible difference between the L2 sound and the closest L1 sound to it). When this occurs, experienced L2 learners may form a "composite" category which comprises of the features of the L1 and L2 phonetic segments proportional to the language that learners hear around them and probably with a concentration of the features coming from more recent language input. This means, the L2 sound productions will continue to stay similar to the L1 sound and the L1 sound will ultimately resemble the L2 sound. For example, Flege (1987) investigating the production of L2 stops by a group of adult L1 speakers of French who had learned English, and another group of adult L1 speakers of English who had learned French, found that voice onset time (VOT) values for non-native stops were different from the VOT values for the same stops in the productions of native speakers.

A rise was observed in VOT in the productions of English stops by the French speakers; however, it still did not correspond to the VOT levels in the productions of the same stops by English monolinguals. Similarly, despite the fact that the L1 speakers of English reduced VOT when producing French stops, it was not enough to correspond to that of French monolinguals. We can indeed explain such differences between native and non-native learners through the critical period hypothesis or less directly ascribe them to L2 filtering/distorting, but such hypotheses cannot account for the modifications in the production of L1 speech sounds by native speakers of that particular L1 (that is, longer VOT seen in French stops in the production of the French speakers and a corresponding shorter VOT in English stops in the production of the English speakers).

The SLM describes L2 speech learning as a gradual process which needs much native-speaker input in order for it to lead to improvement. This idea develops from the fact that L1 speech is acquired across an extended period of time. Consequently, limitations associated with L2 speech learning are convincing only if they come from L2 learners who have had more or less the same amount of input in an L2 as have the children who successfully acquired that language as their L1 including its phonetic segments. A couple of studies showed similar results supporting the significance of L2 input. Both these studies investigated English /p t k/ as produced by adult native speakers of Spanish who had learned English in childhood. English stops produced by the Spanish speakers who had learned English in the United States from American teachers resembled English stops in that they had long VOT values (Flege 1991). On the other hand, English stops produced by Spanish speakers who had learned English from native Spanish speaking teachers in Brazil mainly had VOT values lying half way between native Spanish and English VOT values for the three stops (Flege & Eefting, 1987). It seems that the participants in Brazil had their productions of English /p t k/ affected by the non-native input they had regularly encountered around them.

In another study portraying the significance of input, Flege and Liu (2001) investigated the extent to which groups of adult L1 speakers of Chinese having resided in the United States from age two to seven could recognize English consonants in word-final position. Full-time students comprised half the participants in each group while the other half were professionals who did not have a lot of opportunities to interact with L1 speakers of English at work. A comparison between students who had stayed in the United States for a long time and the students who had lived there for a shorter time revealed that the former were able to identify the consonants in significantly more instances than were the second group. This difference, however, was not seen between the professional participants who had lived there for a short and long time. An important observation in this research was that how often the participants used English did not seem to affect the results of the study as much as the people that they interacted in English with. The authors reached this conclusion because in a self-reported questionnaire the students and professionals indicated more or less the same amount of English use.

This brings us to the end of the discussion on major theories related to speech production and perception in the literature. It is now appropriate to discuss the English model used in this study, namely General American (GA) before looking at the nature of fricative sounds.

2.2 General American

There are a variety of English accents in the United States. The variety known as General American is "…an accent that does not associate an individual with a particular region of the United States, ethnic group, or social class" (Meyer, 2009, p. 33). It is a major accent in the United States and is one that is usually used in national news and American radio broadcasts. According to Trudgill and Hannah (2002) there are minor regional differences in General American; however, it is most often distinguished from other major accents such as Southern, Northeastern and some other regional and social group accents, e.g. African-American Vernacular and is considered "the prestige accent in America" (Lam, 2007, p. 7).

General American in the United States functions more or less the same as British Received Pronunciation in England. In fact, these two accents (GA and RP) are two accents of English language that have been thoroughly described more than any others. However, the two accents have some prominent distinguishing characteristics of their own. According to Finegan (2008) intervocalic alveolar flapping, for example, is a phenomenon that happens in GA. It happens when /t/, /d/ or /n/ occur between two vowel sounds when the first vowel is stressed and the vowel preceding the consonant is unstressed. This results in a flap /ɾ/ which is produced in the place of the consonant, for example /t/ in 'butter' /bʌɾəɹ/, /d/ in 'muddy' /mʌɾi/ and /n/ in 'nanny' /næɾi/. Intervocalic alveolar flapping, however, does not happen when the consonant is immediately followed by a stressed vowel; for example, /t/ in the word ‘retain’ /ɹɪteɪn/ remains the same as it precedes a stressed vowel. A similar situation happens when /nt/ is in the same vowel context (preceded by a stressed vowel and followed by an unstressed one). In such a situation GA speakers usually replace /nt/ with a nasal stop, /ñ/, or even a nasal flap, /ɾ̃/, e.g. ‘enter’ /eɾ̃əɾ/ (Lodge, 2009).

The other difference involving consonants is the presence of coda /ɹ/ in GA which makes GA a rhotic accent. This, however, does not happen in RP where orthographic r in word final position and also before consonants (e.g. far /fɑːɹ/, part /pɑːɹt/, and barn /bɑːɹn/) is not pronounced. This is one of the main distinguishing characteristics English accents. In addition, GA employs two vocoids, /ɚ/ (occurring in stressed syllables) and /ɝ/ (occurring in unstressed syllables), that are normally described as rhotic. In relation to sibilant fricatives, Tottie (2002), compared the individual American and British English sounds and found no differences between the sibilant fricatives of the two varieties (pp. 16-19). And with that, it is now appropriate to look at the nature of fricative sounds and in particular English and Chinese sibilant fricatives.

2.3 Fricatives

Fricatives are speech sounds produced when a turbulent airstream is forced through a narrow constriction formed by two articulators that are close together. The most usual type of fricatives are central fricatives in which air flows across the center of the mouth over the tongue. Some languages have lateral fricatives as well. There is great articulatory precision in forming fricative constriction in the mouth; that is, even a minor alteration in the shape of the vocal tract when producing a fricative sound can cause a huge difference. Fricatives can be formed through the turbulence created at the narrowing of two articulators, or, in the case of sibilant fricatives when a rapid jet of air flowing through this narrowing strikes the edge of some obstruction in the vocal tract; this will be further discussed in the next section.

Table 2.1 Fricatives of English and Mandarin in IPA

	Labiodental	Dental	Interdental	Retroflex	Alveolar	Post-alveolar	Palatal	Velar	Glottal
English	f v		θ ð		s z	ʃ ʒ			h
Mandarin	f	s		ʂ ʐ			ɕ	x	

English has nine fricative sounds which, with the exception of /h/, are all grouped into four pairs according to their place of articulation: labiodental /f/ and /v/, interdental /θ/ and /ð/, alveolar /s/ and /z/, and post-alveolar /ʃ/ and /ʒ/. The fricatives in each pair contrast with one another on the basis of whether they are voiced or voiceless. Mandarin fricatives, on the other hand, do not make use of this contrast very often (are not differentiated based on voicing). In fact, there is only one voiced fricative,/ʐ/, among the six Mandarin fricatives. And as can be seen in Table 2.1, English and Mandarin fricatives are quite different from each other in the place of articulation too with only one common fricative, /f/, between them.

2.3.1 Sibilant Fricatives

Sibilant fricatives (as was just mentioned) are the result of "high velocity jet of air formed at a narrow constriction going on to strike the edge of some obstruction such as the teeth" (Ladefoged & Maddieson, 1998, p. 138). More specifically, sibilant fricatives are the result of a narrow constriction between the tongue and the ceiling of the vocal tract. The high-pitch hissing sound of sibilant fricatives is due to rapid air stream passing through this constriction (Li, 2008). As was mentioned in the previous section, there is great articulatory precision in forming fricative constriction; this is even more important in sibilants where the precise position of the vocal tract has to be maintained for a longer duration of time.

Sibilants are normally described according to their place of articulation; however, acoustically different sibilant fricatives are produced not only through various tongue positions but may also be produced by different tongue shapes. There are a rather large number of post-alveolar fricatives and therefore they should be categorized further into palato-alveolar, alveolo-palatal, and retroflex fricatives (Li, 2008). As these three categories of post-alveolar fricatives have more or less the same place of articulation, there must be some other factor distinguishing them; this is tongue posture (Ladefoged & Wu, 1984). Palato-alveolar fricatives occur at or around the alveolar ridge; /ʃ/, as the initial sound in the word ‘ship’, is an example of a palato-alveolar fricative. An alveolo-palatal constriction is marked by a raised tongue blade toward the alveolar ridge and a bunched tongue body (Li, 2008). An example of an alveolo-palatal fricative is Mandarin Chinese /ɕ/ occurring in the initial sound of the Chinese word 谢谢 (‘xie xie’ in pinyin, meaning ‘thanks’).

Retroflex sibilant fricatives involve the tip of the tongue being curled to some extent and generally raising toward the palatal region. The Mandarin Chinese /ʂ/ is an example of a retroflex fricative as the initial sound in the word 什么 (‘shenme’ in pinyin, meaning ‘what’). /s/ is a sibilant fricative with a dental place of articulation in some languages (e.g. the initial sound in the Mandarin Chinese word 思考, ‘sikao’ in pinyin, meaning ‘to think’) and an alveolar one in others (e.g. the initial sound in the word ‘see’) (Li, F., 2008).

Table 2.2 Sibilant Fricatives of English and Mandarin in IPA

	Dental	Retroflex	Alveolar	Post-alveolar	Palatal
English			s z	ʃ ʒ	
Mandarin	S	ʂ ʐ			ɕ

English has four sibilant fricatives: /s/, /ʃ/, /z/, and /ʒ/ which are voiceless alveolar, voiceless palato-alveolar, voiced alveolar, and voiced palato-alveolar respectively. Mandarin Chinese also has four: /ʐ/, /ʂ/, /ɕ/, and /s/; these sibilants are voiced retroflex, voiceless retroflex, voiceless alveolo-palatal, and voiceless dental respectively.

2.4 Instrumental Studies of Fricatives

Li, Edwards & Beckman (2007) looked at the production of English, Mandarin, and Japanese voiceless sibilant fricatives by thirty participants between the ages of 18 and 30 from the United States, Japan, and China – 10 participants (five males and five females) from each country. They used acoustic analysis measurements applied to tongue posture - whether the tongue has a palatalized or apical (retroflex) shape - as well as tongue position - place of articulation - in order to distinguish the sibilant fricatives of each language from each `other. They looked at /s/ and /ʃ/ in English, /s/ and /ɕ/ in Japanese, and /s/, /ɕ/, and /ʂ/ in Mandarin Chinese.

A list of words beginning with target sounds was used followed by one of the vowels /a/, /i/, /o/, /e/ or /u/ if permitted by the phonological rules of the language. Each sibilant was used at the beginning of three words for each vowel, and so there were three-word targets for each vowel context. The tokens were elicited through a word-repetition task. They, then, used two measures, the first being 'amplitude ratio' to determine the degree of palatalization and the second for the frequency band above the ‘F2 region’ to determine the place distinction.

They found that the retroflex post-alveolar sibilant fricative /ʂ/ differs from the other two Mandarin sibilant fricatives /s/ and /ɕ/ in terms of tongue position and also in terms of tongue posture where the tongue has an apical posture; this last distinguishing factor does not occur among English sibilant fricatives. This could potentially be one of the reasons why native Mandarin learners of English may sometimes confuse certain English sibilant fricatives (for example English /s/ and /ʃ/). The authors also found that English and Mandarin /s/ differ in their phonetic but not phonemic details. This indicates that the two sounds are not phonetically distant from each other. According to SLM, as was mentioned earlier in this chapter, section 2.1.2.3, L2 learners are more likely to develop a new category for an L2 sound if that sound is phonetically distant from the closest L1 to it. So, based on the results of this study, we can predict that native Mandarin speakers (including late learners of English with limited L2 use) will probably have special difficulty perceiving and consequently producing English /s/.

In another study, Chang, Haynes, Yao & Rhodes (2009), using acoustic measures, compared the production of five English and Mandarin voiceless sibilant fricatives (/ʂ/, /ɕ/, /ʃ/, English /s/, and Mandarin /s/) by heritage speakers of Mandarin (those who have been exposed to Mandarin as a child but have shifted to English for the majority of their communication needs), native Mandarin speakers, and native English speakers who had learned Mandarin as a foreign language. They looked at the place contrast, among others, between the two Mandarin post-alveolar fricatives, retroflex /ʂ/ and alveolo-palatal /ɕ/. They also investigated whether the participants could distinguish between the Mandarin /ʂ/ and English /ʃ/, Mandarin /ɕ/ and /ʃ/, and also between Mandarin /s/

and English /s/ and concluded that “These are all pairs of consonants that, due to their high degree of phonetic similarity, stand to undergo equivalence classification and thereby become indistinguishable from each other” (Chang et al., 2009, p. 38).

There were 18 participants in the study between the ages of 18 and 40; five of them were native speakers of Mandarin born and raised in a Mandarin-speaking country; eight were heritage speakers of Mandarin who were either born in or had moved to the United States before they were 10; and finally five were native speakers of English born and raised in the United States but learned Mandarin as a foreign language in high school or college. The stimuli consisted of 62 Mandarin words and phrases written in Chinese characters and pinyin and 35 English words which were presented to the participant randomly using individual index cards. Critical stimuli were fifteen monosyllabic words – ten Mandarin and five English – containing one of the two Mandarin fricatives /ʂ/ and /ɕ/, or one of the other fricatives /ʃ/, Mandarin /s/, or English /s/. They measured peak amplitude frequency (PAF) and centroid frequency as well as F1, F2, and F3 formants. By looking at formant transition data they found that /ɕ/ was the most “palatalized” fricative. The data suggested that, on average, English /s/ has a slightly higher centroid frequency than Mandarin /s/. In reference to the three other post-alveolar fricatives, the centroid for Mandarin /ɕ/ is the highest of the three with /ʃ/ and /ʂ/ coming second and third. The /ɕ/ - /ʂ/ difference was statistically significant for 17 of the speakers. Only one participant – a late learner of Mandarin – could not differentiate the two fricatives.

When it came to /ɕ/ and /ʃ/, all the participants could differentiate between the two. In fact, for all the eight heritage speakers of Mandarin together with seven other participants the difference between /ɕ/ and /ʃ/ was significant based on both PAF and centroid; but the rest could differentiate between the two to a significant degree based on either PAF or centroid but not both. However, when it came to distinguishing between /ʃ/ and /ʂ/ only half the participants – most of them the heritage speakers – could do so. This is potentially an indication that the closest sound to /ʂ/ in English is /ʃ/. It should also be mentioned here that /ʂ/ is represented by letters ‘sh’ in pinyin which makes it more likely for L1 speakers of Mandarin to used /ʂ/ in place of /ʃ/ when speaking English. The last pair of fricatives that were studied was Mandarin /s/ and English /s/. Most of the heritage speakers and also late learners of Mandarin could distinguish the two while the majority of native Mandarin speakers could not. The fact that most of the heritage speakers could distinguish English /s/ and Mandarin /s/ is consistent with the SLM view that bilinguals who learn their L2 early in life have separate phonetic categories for an L2 and the closest L1 sound to it even if the two sounds are not phonetically distant from each other. However, for late learners the kind of L2 they are exposed to (native input versus non-native input) and then the amount of exposure seem to be the important factors in being able to distinguish a non-native phoneme from the nearest native sound to it. But there was no mention of the kind or amount of exposure to Mandarin by the late Mandarin learners. If the late learners in this study have received native Mandarin input for an extended period of time, we can accept the results as corresponding to the SLM proposals.

In conclusion, the majority of the heritage speakers in this study were able to distinguish all the pairs. The authors offered two possible explanations for this in line with the SLM claims. First, it is possible that due to their exposure to English and Mandarin at an early age the heritage speakers were able to produce the fricatives in both languages accurately though not completely identical to target sounds. The second explanation proposes that when two languages are learned at an early age, a shared phonological system is formed and dissimilation of similar categories of sounds happens; in other words, it is proposed that for heritage speakers there is phonetic distance between categories of sounds that may sound similar or even identical to others. However, the authors believe that the first hypothesis might explain better why more heritage speakers were able to distinguish similar categories of sounds than non-heritage speakers in this study because they did not find a lot of evidence that proves there is phonetic distance at work for the heritage speakers. Additionally, the authors pointed out that the participants' differences in PAF or centroid could show that the speakers merged the articulations of similar categories of sounds but these differences did not prove whether their pronunciations of similar categories were identical nor did they show what categories they merged closer to.

Johnson and Babel (2010) conducted a perceptual similarity rating task and a speeded discrimination task to see how differences in phonetic inventories as well as phonological alternations (different phonological rules) impacted on the ability of Dutch and American English speakers when discriminating the six English and Dutch voiceless fricatives, /f θ s ʃ x h/.

There were 16 (five males and eleven females) American English and 12 (six males and six females) Dutch speaking participants in the similarity rating task. The Dutch participants were all living in the United States at the time of the study and they all spoke English though their English proficiency was not rated. The majority of them started learning English at the age of twelve or earlier. Most of them had also been living in the United States for at least four years or more. In this task, the voiceless fricatives were placed between three pairs of vowels, [a_a], [i_i], [u_u]; hence there were 6 tokens for each vowel context. Tokens of the same vowel context were paired with each other and repeated several times. The main author, a native speaker of American English, was recorded saying them.

Overall, the listeners listened to 378 trials of the stimuli and compared and rated each pair on a 5-point scale from ‘‘very similar’’ to ‘‘very different’’. After analyzing the rating scores, pairs contrasting /s/ and /ʃ/, /s/ and /θ/, and /ʃ/ and /θ/ showed the most difference between Dutch and American participants with Dutch listeners rating them as more similar. There is no phonemic distinction between /s/ and /ʃ/ in Dutch and this can explain why the Dutch listeners rated the pairs as more similar even though they are familiar with /ʃ/. Looking at the vowel contexts, On the other hand, Dutch and American listeners differed in 3 of the pairs. This difference was significant in the case of i_i for s/θ, s/ʃ, and θ/ʃ pairs. In u_u and a_a the difference could be seen in the s/θ and s/ʃ pairs respectively. In all these cases the Dutch rated the pairs as more similar. According to the authors, "for highly confusable pairs such as [f]/[θ] the rating task is not sensitive to language differences because a floor effect on rating scores obscures any differences between the Dutch and English listeners" (Johnson & Babel, 2010, pp. 130 & 132).

For the second experiment a speeded AX discrimination task was used with the same stimuli as experiment one's. The participants were fifteen Dutch speakers (nine males and six females) nine of whom participated in experiment 1 as well, and also nineteen American English speakers (twelve females and eight males). The majority of Dutch speakers in the second experiment had been living in the United States for two years or less; however, most of them reported starting to learn English at the age of 12 or earlier. Experiment two was done about three months after the first experiment. The participants had to decide whether 'X' – the second stimulus of a pair – was the same as or different from A – the first stimulus of the pair. In this experiment, the listeners had to respond to each pair of stimuli within 500 ms, hence the *speeded* AX discrimination task. This was done to prevent the participants from introspecting on the sounds. The findings of experiment 2 show that, regardless of their native language, the participants could correctly identify the pairs of stimuli as 'different' or 'the same' 95% of the time.

To sum up, the results of two experiments here showed on one hand that the Dutch speakers could discriminate the fricatives most of the time (experiment two) but they (given the fact that they spoke Dutch quite a lot in the United States and retained a strong Dutch proficiency) nevertheless had a tendency to rate most the fricatives put in pairs as phonetically closer to each other than did the English speakers. This is consistent with Flege and Eefting's (1987) findings, mentioned earlier in this chapter, showing that Spanish speakers who had learned English from native Spanish speaking teachers in Brazil produced English stops with intermediate VOT values (half way between VOT values for native Spanish and native English speakers). So we can conclude that this is further support for the SLM's proposal that L2 speech learning is a slow process which needs a lot of native-speaker input in order for it to be successful.

In a study to test the accuracy of PAM predictions, Best et al. (2001) investigated the ability of 22 L1 speakers of American English (fifteen female, seven male) to discriminate non-native Zulu consonants through an AXB categorial test. They also studied the assimilation patterns of these consonants into American English consonants by conducting a questionnaire test. They specifically chose three pairs of Zulu consonants that they predicted (according to PAM) to be assimilable into American English consonants. The three pairs of consonants consisted of: "i) voiceless versus voiced lateral fricatives (/ɬ/-/ɮ/); ii) voiceless aspirated versus ejective (glottalized) velar stops (/kh/-/k′/); iii) plosive versus implosive voiced bilabial stops (/b/-/ɓ/)" (p. 778). The first pair (lateral fricatives) employs a non-native place of articulation for English listeners. However, in the two remaining pairs, it is the type of distinction between the consonants (laryngeal distinction) that is non-native for the American listeners not the place of articulation. Based on PAM assimilability patterns of prediction (as was mentioned in section 2.1.2.1), there are three patterns that predict assimilation of non-native phonemes into native sounds; those three patterns, in order of how well non-native sounds are assimilated into native sounds, are: two-category assimilation (TC), category-goodness difference assimilation (CG) and single-category assimilation (SC).

The authors predicted that the lateral fricatives would be assimilated into two native categories by the majority of the participants (according to TC pattern). The two lateral fricatives would probably be perceived "as some phonological contrast in English, such as a voiceless apical fricative (e.g. /θ s ʃ/, perhaps clustered with /l/) versus /l/ (voiced lateral approximant) or some voiced apical fricative (e.g. /ð z ʒ/, perhaps clustered with /l/), which involve the same articulators (tongue tip and dorsum, glottis)" (p. 779). The authors further predicted that the listeners would have little difficulty discriminating the two lateral fricatives. The next pair, the velar stops, was predicted to fall under CG, so they would be assimilated into one native category but one would be a better example of the native phoneme, English /k/, than the other. Discrimination of the velar stops was expected to be moderately good but not as good as the first pair. Finally, the third pair, the bilabial stops, was predicted to follow a SC pattern of assimilation in which listeners would perceive them as more or less equally poor examples of English /b/. The discrimination of the two stops would also be poor.

An L1 speaker of Zulu produced 20 tokens of each target consonant in CV nonsense syllables. Ultimately six tokens for each consonant, perceived as the clearest productions, were chosen from among the 20 original tokens. The AXB categorial discrimination test was first conducted for all the three pairs in which A represented one of the consonants in one of the contrasting pairs and B represented the other consonant in the same pair. X marked the target consonant which was either the same as A or B. So listeners' job was to identify whether A or B was the same syllable as X. However, X was acoustically different from A or B.

In order to access the assimilation of the non-native consonants, the authors carried out a questionnaire task which required the participants to listen to the same syllables one more time and write down in English orthography what they heard. The listeners were told to strictly write down only the consonants that they perceived as an English consonant or approximated one. After this, the listeners answered further questions on the questionnaire about how they thought syllables sounded (e.g. whether they heard the stimuli as speech sounds or not or whether they thought the speaker was doing something while producing the sounds).

The results of the questionnaire task corroborated the predictions that were originally made about the assimilation of the Zulu consonants by native American English speakers. In short, the lateral fricatives were assimilated into two native consonants as TC pattern indicated. All the listeners perceptually assimilated velar stops into one English consonant consistent with CG pattern. Similarly, the bilabial stops were assimilated into a single native category, but this was consistent with the SC pattern; that is, the listeners considered neither bilabial stops as good examples of English /b/.

The results of the AXB discrimination test too confirm the predictions that the listeners were able to discriminate the lateral fricatives better than the other two pairs. Velar stops were the next pair in terms of how well they were discriminated by the listeners. This was followed by bilabial stops which were the least discriminated of the three contracting pairs.

The results of this study for the first pair (lateral fricatives) support the view by PAM that listeners can detect small differences in native consonants. The fact that the listeners could assimilate the lateral fricatives into two native consonants without much difficulty shows that the fricatives were perceived as close to native phonemes and so the discrimination of the fricatives would also be very good. Best, who developed PAM, not only employed the AXB categorial discrimination test in her study here, but she had also used the same test in some of her previous studies. The fact that the predictions in this study were supported by the results indicates that AXB categorial discrimination test is potentially a useful and reliable test. That is why it was used in the perception part of this dissertation as well.

Tsao, Liu & Kuhl (2006) investigated the ability of adult and infant Mandarin and English speakers to discriminate Mandarin fricative-affricates /ɕ/-/tɕ/, /ɕ/-/tɕh/ through three experiments. The experiments were designed to test the effects of language experience on the perception of these contrasts. This can also help us when trying to find out the sources of confusion for native Mandarin speakers in discriminating English fricatives. There are actually important differences between the way native Mandarin and English speakers recognize fricatives and affricates of their L1 and/or discriminate between the two is in their native language. One of the most important differences is the fact that manner of articulation itself is phonemic in English but not so in Mandarin; in other words, native Mandarin speakers do not rely on manner of articulation to distinguish Mandarin fricatives and affricates. Another difference is voicing which is important in distinguishing affricates from fricatives (as well as fricatives from each) in English but which is not very important in the discrimination of such sounds in Mandarin. Another important difference, more specific to the affricates and the fricative in this study, is Mandarin having an alveolo-palatal place of articulation which does not exist in English.

In the first experiment, English and Mandarin speaking adults tried to discriminate three pairs of Mandarin alveolo-palatal consonants comprising of two affricates and an affricative, /ɕ/-/tɕ/, /ɕ/-/tɕ^{h}/, and /tɕ/- /tɕ^{h}/, through an AX discrimination task. As was expected by the authors, the results of this experiment indicated that native English speakers had more difficulty discriminating the consonants which they contributed to the effects of prior language experience on the discrimination of non-native contrasts. More specifically they pointed to distinguishing factors such as aspiration (especially for /tɕ/- /tɕ^{h}/) as well as amplitude rise time and frication duration which the authors believed had a significant effect in the ability of the Mandarin speakers (but not the English speakers) to discriminate the sounds in the three pairs.

Experiment 2 investigated the developmental change in the perception of /tɕ/- /tɕh/ by American and Taiwanese infants in two different age groups of 6-8 and 10-12-month old through a discrimination task. Very different results were obtained for the two language groups indicating that Taiwanese infants' ability to discriminate the two consonants increased considerably from the younger infants to the older infants while the opposite held true for American infants; that is, their ability to discriminate the two sounds decreased with increasing age. However, the younger infants in the two language groups did not show significant differences in their performance and it was only the older infants in the two language groups that significantly differed in their ability to perceptually discriminate the consonants. This shows the difference in the perception of native versus non-native phonetic segments early in life.

The third experiment was similar to the second one. However, in this experiment, the perception of a pair of English affricate and fricative consonants, /tʃ/-/ʃ/, by American infants only was investigated. Also, the infants were not the same as the ones in experiment two but they were divided into the same two age groups as before. The results indicated that the older American infants in this experiment, just like the older Taiwanese infants in the second experiment, performed better than the younger ones and they also performed better in comparison to the older American infants in the second experiment. However a comparison between the performance of the younger infants in this experiment and the performances of the younger American and Taiwanese infants in the second experiment do not reveal much difference.

Finally, McGuire (2007) investigated the perception of Polish alveopalatal and retroflex voiceless sibilants by native speakers of English through a number of brief experiments. These were training experiments in which a two-dimensional stimulus set was used. Later, the effectiveness of the stimulus set was evaluated. The author also explored the process of category formation in adults and studied the perceptual cues which helped the participants in the study to identifying categories.

For native speakers of English the distinction between Polish alveopalatal /ɕ/ and retroflex /ʂ/ sibilants normally falls under the variations of one single English sound which is /ʃ/. This Polish contrast is seen in Mandarin Chinese too as will be discussed shortly. Between English /ʃ/ and Polish /ʂ/, the tongue is much flatter (but with a curled tip) for the latter; however, there are shared features between the two including lip-rounding and width and location of constriction which make them sound similar. /ʃ/ and /ɕ/ also share some features such as lip rounding and very similar places of articulation but /ɕ/ exhibits higher tongue blade and body. The two Polish sibilant fricatives were also compared with the two very similar Mandarin sibilants /ɕ/ and /ʂ/. It was found that for Mandarin /ʂ/ there is no lip rounding, but that the sublingual cavity was larger compared with Polish /ʂ/; other than that, the articulation of the two fricatives in the two languages were very similar (Ladefoged and Maddieson, 1996, cited in McGuire, 2007, p. 394). This is of importance to this dissertation as one of the target sounds, English /ʃ/, was often produced as /ʂ/ by the Chinese participants.

The stimuli for the study were the Polish sibilants /ʂ/ and /ɕ/ preceding the vowel sound [a]. The authors looked at two aspects of the fricatives, the first being fricative noise and the other vocalic transition. They were able to do this by taking [ʂa] and [ɕa] syllables, produced by a native Polish speaker, and breaking them in two at the onset of voicing. The separated fricatives and vowels were then separately put together to create vowel-fricative pairs. Overall, there were 100 vowel-fricative pairs.

Next, five native speakers of American English listened to 16 of the 100 stimuli and attempted to transcribe them using English orthography. Each token was presented five times to the listeners in random order, so, overall, they listened to 80 trials of the stimuli. The results revealed that almost all the participants transcribed the initial consonant in the stimuli as ‘*sh*’. The only exceptions were a participant who transcribed the consonant as ‘ssh’ and another who left it blank and did not write anything for the sound.

Based on the results, the author concludes that the English listeners were not able to discriminate the two fricatives in turn because they could not differentiate the friction noise in them. McGuire admits that it would probably be difficult for the participants to show a distinction in the friction noise orthographically.

In order to solve this problem, the author designed another experiment in which three of the participants in the previous experiment and seven other native speakers of English underwent brief training in labeling the two Polish sibilant fricative sounds to see whether they could identify the fricatives correctly. The subjects then transcribed five repetitions of the 16 stimuli for a total of 80 tokens using the labels they learned in the training. As a result of using these labels, the subjects were not restricted to English orthography to label the stimuli, and so the author was able to see if they could discern the differences in the stimuli. The labels used in the training were *sz* and *ś*. They were told that the two Polish sounds they were about to learn were very similar to English /ʃ/. The participants then tried labeling the tokens.

The results of this experiment indicate that the majority of the listeners could discriminate the two fricatives; however, this discrimination was based on vocalic cues rather than frication noise in almost all the cases. Native Polish speakers use friction noise to discriminate the two fricatives. What is more, the effect of the training for most of the listeners was that they actually paid less attention to frication noise or even stopped using it. Only a few participants had improved discrimination based the frication noise as a result of the training the training.

This study demonstrates that the listeners' responses would probably follow a CG pattern according to Best's PAM (as was mentioned earlier). However, this study goes one step further in that it tries to discover the source of the discrimination ability among those listeners who could distinguish the two fricatives. This indicates that even though the majority of the listeners could discriminate the two fricatives, they still relied, for the most part, on the information that native Polish speakers do not normally use as a primary source to discriminate /ʂ/ and /ɕ/.

2.5 Studies on the Production and Perception of Fricatives

In a study depicting the general characteristics of the English spoken by speakers from China, Deterding (2006) looked at the most salient deviations in the production of 13 college students from China. Among the English sibilant fricatives, he investigated the errors the participants made when producing English /ʒ/ and /z/. There was, however, no discussion of the way they produced English /s/ and /ʃ/ most probably because he did not find them very different from the normal English pronunciation. The participants, 10 male and 3 female, aged 18 to 21, were recorded reading the North Wind and the Sun passage followed by a short interview. The participants reported studying English at school in China, on average, for 8.4 years prior to coming to Singapore. Deterding (2006) also pointed out that there is only one voiced fricative in Standard Mandarin which is represented by IPA character /ʐ/ and by the letter 'r' in pinyin. Pinyin 'r' has another realization in Mandarin phonetics as an approximant but the difference between the two realizations is phonetic and not phonemic. Three of the speakers in his study produced /ʒ/ in the word 'usually' as /ʐ/ and three others as /ʃ/. The speakers who produced /ʐ/ in place of /ʒ/ were all from northern parts of China. In fact, speakers from northern parts of China also pronounce pinyin 'r' as a /ʐ/. It should be noted that Deterding does not use the IPA symbol /ʐ/ and instead uses [ɹ] but as was mentioned earlier, the IPA symbol for the voiced retroflex fricative is /ʐ/; the other realization of this sound is /ɻ/ which is a voiced retroflex approximant. Regarding /z/, there is one instance of /z/ in the word ‘as’ being replaced with /d/ and a few instances of /z/ being substituted with /ð/; there were also several instances of a glottal stop in the place of /z/. And finally there were some speakers who omitted /z/ altogether. In conclusion, this study found that among the four English sibilant fricatives /ʒ/ and then /z/ caused the most difficulty for speakers from China.

Jing and Yanyan (2011) looked at the production and perception of English fricatives by two groups of college students from China. The first group consisted of 32 non-English major and the second of 26 English major students. They employed a listening discrimination test and a reading exercise. There were seven pairs of English fricatives placed in minimal pairs in the listening discrimination test. The contrasting pairs were: /f/-/v/, /w/-/v/, /s/-/z/, /θ/-/s/, /ð/-/z/, /ʃ/-/ʒ/, and /θ/-/ð/. The listeners’ task was to listen to the recording of only one of the two words in each minimal pair (differing only in the fricative sound) and then decide which word they heard. There were six minimal pairs for each fricative pair. To facilitate the process, the listeners were presented with both words in each minimal pair on the test paper and were asked to mark the one they heard. For the last two contrasting fricative pairs (“/ʃ/ and /ʒ/” and “/θ/ and /ð/”), however, there were not enough minimal pairs so the listeners were presented with the phonetic symbols of these fricatives and then listened to a word with either the first or the second fricative and were asked to pick the symbol of the sound they heard.

The same seven contrasting fricative pairs appeared in the reading discrimination test too. This was designed to see if the participants could discriminate the pairs orally. The fricative pairs were placed in word pairs and there were three pairs of words for each fricative pair. All the words were then randomly put into a list and given to the speakers to read while being recorded.

In the listening test, one point was given to the listeners for every correct choice they made. For the reading test, the recordings were judged by two English teachers (the authors did not state whether they were native speakers or not) but just mentioned that they had experience teaching English pronunciation. The criterion for judging the tokens produced by the speakers was comprehensibility of the fricative sounds. So every comprehensible token would get one point as well. They then analyzed the data using SPSS and determined that the accuracy of performance in the listening test for the English major and non-English major students stood at 84.33% and 72.55% respectively while this was 86.45% and 80.07% for the reading test respectively. In the listening test specifically, both groups had difficulty discriminating /ð/ from /z/ and /w/ from /v/. In the reading test, on the other hand, none of the participants could produce /ʃ/ and /ʒ/ and also /θ/ and /ð/ distinctively from each other. They also concluded that the participants had a "good performance" when it came to discriminating /f/ from /v/ and /s/ from /z/.

In explaining this good performance they pointed to positive transfer from Mandarin where we can find /f/, /s/, and the affricate sound /ts/. /f/ occurs in Mandarin and so does /s/ but /s/ is an alveolar sibilant fricative in English while it is a dental one in Mandarin; dental and alveolar /s/, however, are allophones of the same sound in both languages and therefore replacing them with one another will probably not cause any difficulty for the native speakers of either English or Mandarin. But this is not true for /z/ and /ts/ (as Jing and Yanyan claim) and that is because /ts/ is not an allophone of /z/. They acknowledge that /ts/ is an affricate but go on to say that this similarity between /z/ and /ts/ (along with the similarity between English and Mandarin /s/) is "assumed to be beneficial for learners". However, in the previous study, Chinese participants had considerable difficulty producing English /z/. It seems that Jing and Yanyan mostly considered the intelligibility of these sounds in the reading task than strictly their accuracy. One thing that is common between this study and the previous one is the fact that some participants confused /ʃ/ and /ʒ/ in their productions.

2.6 Summary

In conclusion, a variety of theories on L2 production and perception were offered. As was seen, perception theories don not only account for the way learners perceive L2 sounds but also the way they produce them. One of those theories, the SLM, was discussed in detail. It is a theory that is particularly relevant to this current research as it specifically looks at L2 acquisition. Mandarin and English sibilant fricatives were then described along with the similarities and differences between the two. According to the principles of SLM, these similarities and differences certainly affect the way L1 speakers of Mandarin produce and perceive English sibilants. When it comes to actual differences between the two groups, most Mandarin sibilant fricatives differ from English sibilant fricatives in tongue posture (tongue shape) as well as tongue position (place of articulation). Some studies in the literature found that English /ʒ/ and /z/ are the most problematic English sibilant fricatives for L1 speakers of Mandarin. This study looks at all the four English sibilant fricatives produced and perceived by L1 Mandarin speakers and tries to draw a comparison between the two as there are not a lot of studies that have done so.

CHAPTER 3

RESEARCH METHODOLOGY

This chapter has eight main sections. In the first one, the participants of the study are described; the data is described in the second one; and the third and fourth discuss the process of data selection, collection, and analysis. Finally, sections five, six and seven focus on the procedures carried out to test the accuracy of the author’s productions which were used in the perception test of the study.

3.1 Participants

The participants were 16 students from China, eight males and eight females, aged 19 to 21, enrolled in an English proficiency program at Technology Park Malaysia College (TPM College), Kula Lumpur, Malaysia at the time of the study (second semester of the 2010/2011 academic year). They were taking an English proficiency class because they did not pass the English placement test at the college. Right after the proficiency program, they would begin their twinning undergraduate program which involved the students taking the first half of their courses at TPM College and moving to California, United States, for the rest of their programs. All the participants reported speaking Mandarin Chinese at home and no other Chinese dialects. The participants had all entered Malaysia at the same time and had been staying there for six months at the time of the study. The participants normally spoke Mandarin outside of the class, and since their proficiency program was an intensive one, they did not have the chance to spend a lot of time outside of their dorms communicating with Malaysians; it is, therefore, highly unlikely that their English was influenced by the variety of English spoken in Malaysia.

The participants had all studied English at school in China for six years prior to coming to Malaysia. The 16 participants were part of a larger group of students from China at TPM College; these 16 students were chosen because they were all from northern parts of China and reported speaking Mandarin at home and no other Chinese dialects.

3.2 Data

There are two main sources of data in this study: a production task and a perception test.

3.2.1 Production

The production task involved the production of 31 frequently occurring English words containing the English sibilant fricatives /s/, /z/, /ʃ/, /ʒ/ in the following environments: (i) initial and final positions of one-syllable words, (ii) initial and final consonant clusters of one or multi-syllable words, and (iii) word medial position of two-syllable words; however, not all these four sibilants occur in all the positions mentioned as will be seen. There were ten words containing /s/ in all the above-mentioned five positions (two words for each position). Similarly, there were ten words containing /ʃ/ in all the five positions (two words for each position). The sound /z/ occurred in all but the initial consonant cluster position as there are no words in English language with an initial consonant cluster beginning with /z/. As a result, there were eight words containing this sibilant (two words for each position). /ʒ/ is a sound that does not frequently occur in English words and so it was only used in three words in this study (*usually*, *television*, *pleasure*).

3.2.2 Perception

The second source of data, the perception test, came from an AXB categorial discrimination test in which X is an exemplar of either A or B, though not acoustically the same (because the three tokens have been separately produced by the same speaker); for example, if X were /kləʊzd/, either A or B would also be /kləʊzd/ and the other one would differ in the sibilant fricative sound (e.g. /kləʊsd/). Each trial of AXB, therefore, contained three words which included one odd item out. The words used in the perception test were the same 31 words used in the production test. Best, McRoberts & Goodell (2001) also used the AXB categorical discrimination test in their study, as was explained in literature review, section 2.3.

3.3 Procedure

3.3.1 Production Task

The recordings were done in a quiet room at the College (prior permission had been obtained from the college). The participants had to fill out a form stating their age, length of stay in Malaysia, their first language, and how long they had been studying English. They were then informed about the study and what they were expected to do. They were also asked to pronounce the words clearly and avoid reading through the words very quickly. They were also told that their identities would not be revealed and it was their choice to take part in this study.

The recording started with a short practice session in which one of the speakers was given a list of 10 words (other than the 31 words); she was then asked to read the words and was recorded in front of the other speakers; feedback was given as to how clearly she produced the words and if her productions were too fast or too slow. Once the speakers indicated their full understanding of what was expected of them, the recording of the original 31 words, which had been randomized and put onto flash cards, commenced. The speakers were recorded using Marantz PMD661 digital recorder. They were recorded reading the list of words twice with a short break of two minutes in between in order to relieve any possible fatigue or anxiety which could have compromise the quality of their productions.

The recordings were then rated by five native speakers of General American (GA), three men, two women, who had at least some tertiary education. They indicated how close they thought the students' productions of the sibilant fricatives were to those of American English. The rating was done on a five-level Likert scale as follows:

1. Native like
2. Very similar
3. Similar
4. Different
5. Very different

A very similar version of this scale was also used by Johnson and Babel (2010), as mentioned in chapter two, section 2.3.

It should be mentioned that this study was not interested in the interaction between trials 1 and trials 2 of the test as they are not pre- post-tests. The two trials were simply there to ensure that there were sufficient data so that the analyses could produce reliable results and also to increase the probability that their production and perception of the fricatives did not happen by chance.

3.3.2 Perception Test

The words in the AXB perception test were produced by the author of this study. The only thing that made either A or B different from X was the sibilant fricatives in them; that is, each word in every single trial of the perception test was first produced and recorded twice as it is normally pronounced in GA whereas, for the third token, the same word was produced with the incorrect sound in place of the correct sibilant fricative. The choice of which sound to replace the target sibilant fricatives with came from the non-native productions of the sibilant fricative by the Chinese speakers made in the production task; many of these non-native productions involved the participants replacing Mandarin Chinese sounds with the target English sibilant fricatives (as will be discussed). This was done to see whether the participants could later recognize the words with the incorrect consonant. As an example, if the word "ship" (used in the production test) were wrongly pronounced by one or more of the participants with the Mandarin retroflex fricative sound /ʂ/, two of the tokens in the trial would be pronounced /ʃɪp/ and the third one /ʂɪp/ by the author. The participants took the same perception test twice with a short break of two minutes in between; thus there were two trials here as well.

3.3.3 Instrumentation

A Marantz PMD661 Professional Solid State Recorder was used in this study to record all the productions through an Audio Technica ATM73a cardiod condenser headworn microphone. The microphone was placed at the recommended 3 centimeters from the speakers' mouth (Plichta, 2010). In order to ensure high quality recording suitable for acoustic analysis, the recordings were sampled at 44.1 kHz, 16-bit rate. For the perception part, Sony's HT-CT550W 40" 3D Sound Bar System connected to a laptop was used to ensure the quality of playback.

3.4 Data Analysis

After the production data was rated by the GA native speakers, the ratings were analyzed with SPSS Version 21.0. At first, however, frequency charts of the number of tokens receiving a rating of one all the way to a rating of five were made individually for each of the four sibilants in the participants' productions. There were four charts for each sibilant sound showing the frequencies for each gender by trial. The ratings for each sibilant were then analyzed using Levene's test (in SPSS) to see if the four groups (the four ratings in each chart) had equal variances. Next, an independent samples test was carried out to see if there were any significant differences between the ratings for the two trials and gender groups for each sibilant.

After analyzing the ratings for each sibilant separately, the mean ratings for the four sibilants indicating the degree of non-nativeness were compared for the two trials as well as for the mean ratings of trials one and two to see which fricatives caused the most and the least difficulty for the participants in production and if these were same across the two trials. A mixed design ANOVA was, then, carried out to check for the effects of Sibilant and Gender. Next, a comparison of the judges' ratings (degree of non-nativeness) was made for each sibilant sound based on the position of the sibilant in the target words (except for /ʒ/ which occurred in three words in the medial position only). This was done to see if the position of the sibilants had an effect on how accurate they produced the sounds. An ANOVA was then carried out to see if there were any significant differences among the ratings of all the positions for each sibilant sound. This was followed by a Tukey post-hoc test to see exactly which two positions had significantly different ratings.

Finally, the author also listened to the data and phonetically transcribed the sounds for which the Chinese speakers were rated. This was done to see what rating the native speakers gave to the sibilants that were either identified by the author as correct target sibilants or incorrect consonants and if the ratings and the authors transcriptions were comparable.

In the perception test, the number of wrong answers for each gender and sibilant sound in each trial was calculated and compared. This was done to see which gender in which trial had the highest number of wrong answers for each sound. A repeated measures ANOVA was then carried out to see if the gender and trial effects were significant. Next, the percentages of the wrong answers for each sibilant sound (across both trials and genders) was obtained and compared first for each trial separately and then for the two trials combined. The reason percentages rather than number of wrong answers were used here was because the total number of perception test instances varied for each sibilant. A mixed design ANOVA was carried out here as well to check for gender and gender effects across both trials.

The number of incorrect answers in all the word positions for every sibilant (except for /ʒ/ which occurred in three words in the medial position only) was compared and was followed by a one-way ANOVA. This was done to see if the position of the sibilants had a significant effect on whether the listeners could correctly perceive the sounds. The results of the production task and the perception test were then compared for each of the 31 words to see how closely they corresponded with each other.

As there was a wide gap between the production and perception results for /s/, a follow-up perception test was done. This second perception test was the same as the original one in that they could pick any one of the three tokens as the odd item but they also had the additional fourth option of “the same” in case they perceived the /s/ in all the three tokens as the same. The participants were told that the productions were not the same as the ones they listened to in the first perception test as not to make them think that there must have necessarily been an odd item out while in fact there *was* an odd item out in each trial. This second perception test enabled the author to see if the discrepancy in the production and perception results for /s/ was due to the fact that the participants perceived the alveolar and dental /s/ as the same.

3.5 Rating of Author's Productions of English Sibilants

Even though I, as the author of this study, did not grow up in an English speaking country, I started my English education at an early age (around age 5) and had American teachers. I was constantly exposed to American English at home too through American programming on Television and American children's stories on tape and American English audio books later on. Having said that, I decided it would be best if my productions of the sibilants were rated by the native speakers too. Two of the native speaker raters (a man and a woman) rated my productions the same way they did the Chinese speakers'. Figure 3.1 shows the two native speakers' perception of the degree of non-nativeness (%) for the four fricatives in my productions. As can be seen, /ʃ/ and /ʒ/ were rated by both raters as 100% native-like. Only two tokens containing /s/ were given a value of "2" which means that they were perceived by the raters to be "very similar" to American English /s/. The rested of the tokens containing /s/ were perceived as "native-like". Similarly, /z/ was perceived to be "very similar" to American English /z/ only once and the rest of the tokens containing this consonant were judged to be "native-like".

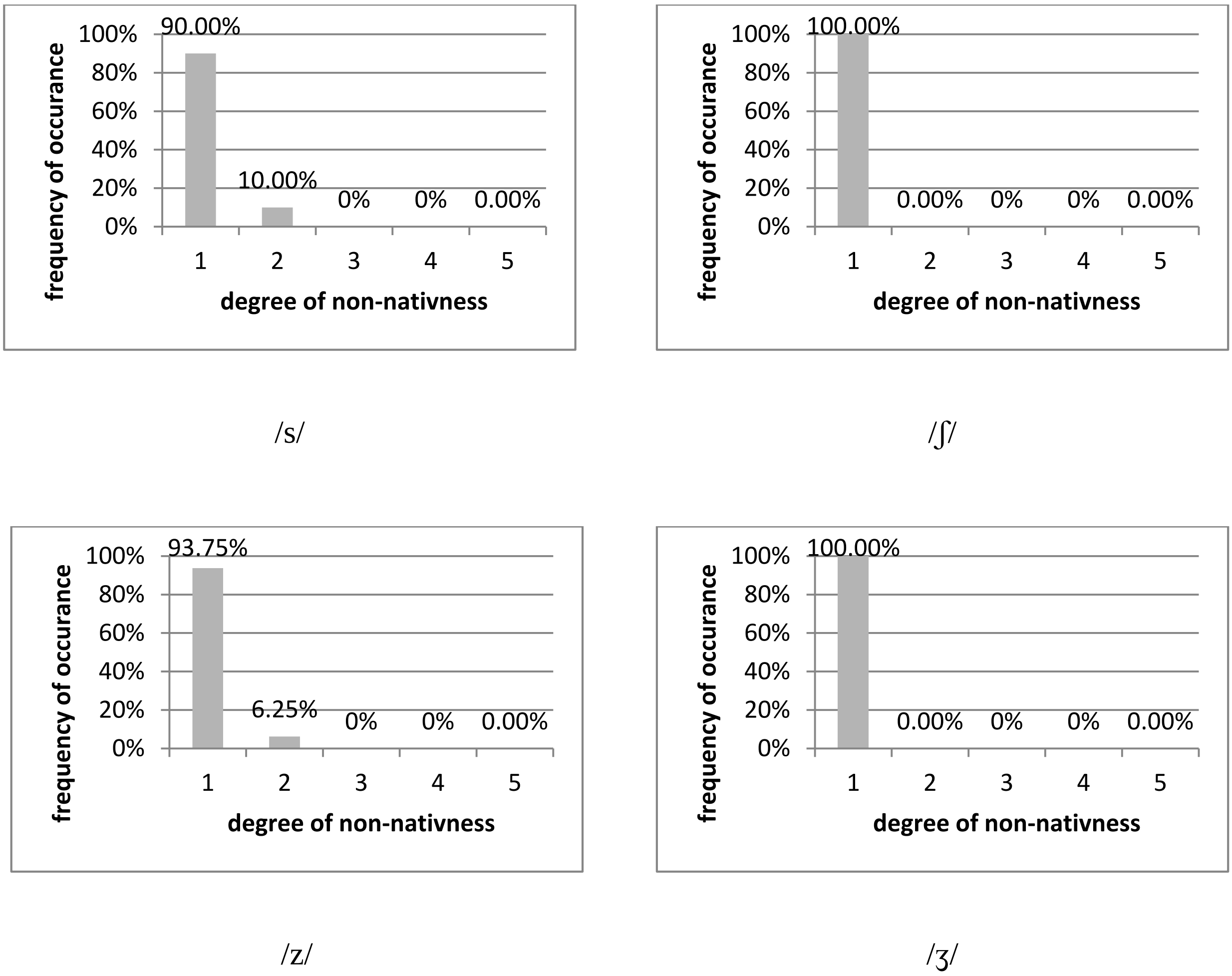


Figure 3.1 Native Speakers’ Perception of the Degree of Non-Nativeness (%)
(1 = native-like, 2 = very similar, 3 = similar, 4 = different, 5 = very different)

Figure 3.2 shows the degree of non-nativeness, as perceived by two of the native speakers, for the four fricatives in the author’s production of the 31 words. Here, the degree of non-nativeness (%) for /ʃ/ and /ʒ/ is zero, and is very small for /z/ and /s/ (1.56% and 2.50% respectively). It can, therefore, be concluded that the author’s productions were suitable for the perception test.

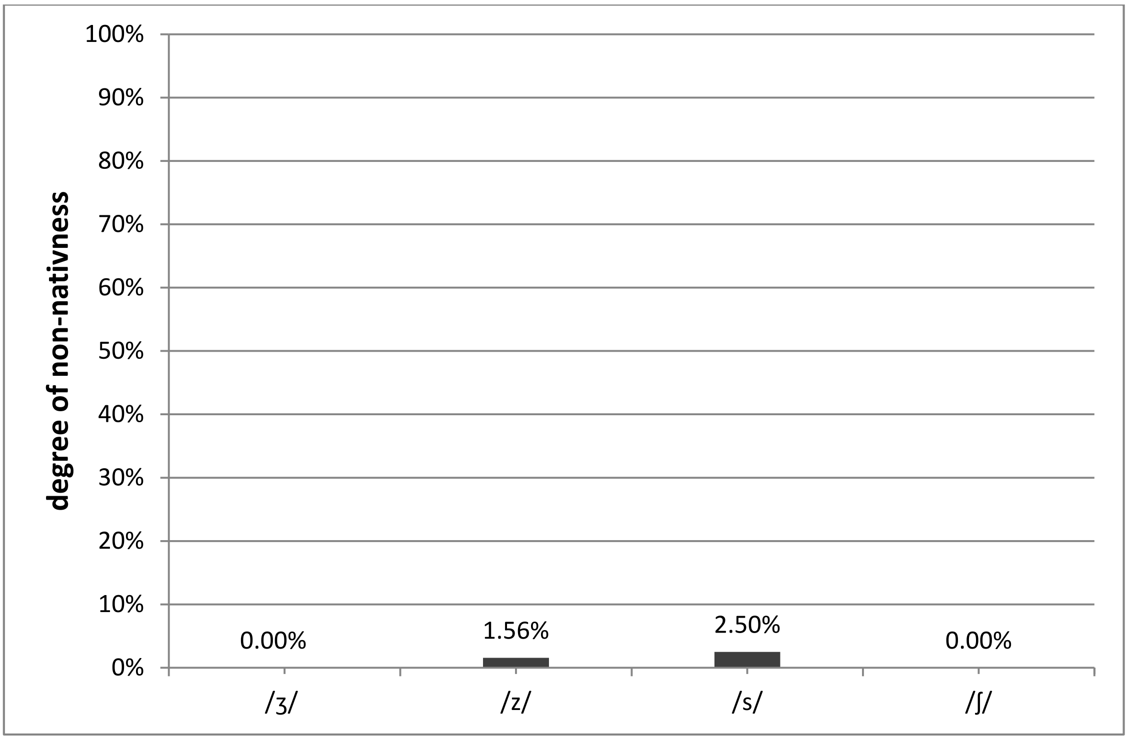


Figure 3.2 Degree of Non-Nativeness in the Author's Productions

3.6 Raters Taking the Perception Test

The two native speakers who rated the author's productions also took the same AXB categorial perception test that the 16 participants took in order to see if they could pick the odd items out. Figure 3.3 shows the perception test percentages of incorrect answers for the four fricatives for the two native speakers. As can be seen in Figure 3.3, the native speakers did not have any difficulty picking the odd item out for /z/, /ʒ/, and /ʃ/. This shows that the native speaker raters could clearly hear the difference between native and non-native sibilants produced by the author. However, the situation was a little different for tokens containing /s/ as the two native speakers did not seem to be certain about which items to pick as odd and ended up picking the wrong items in 35% of cases. However, these were the same productions that they had earlier rated as native-like or very similar to native-like (in two tokens only, see section 3.5). This last point will be elaborated upon in the next chapter.

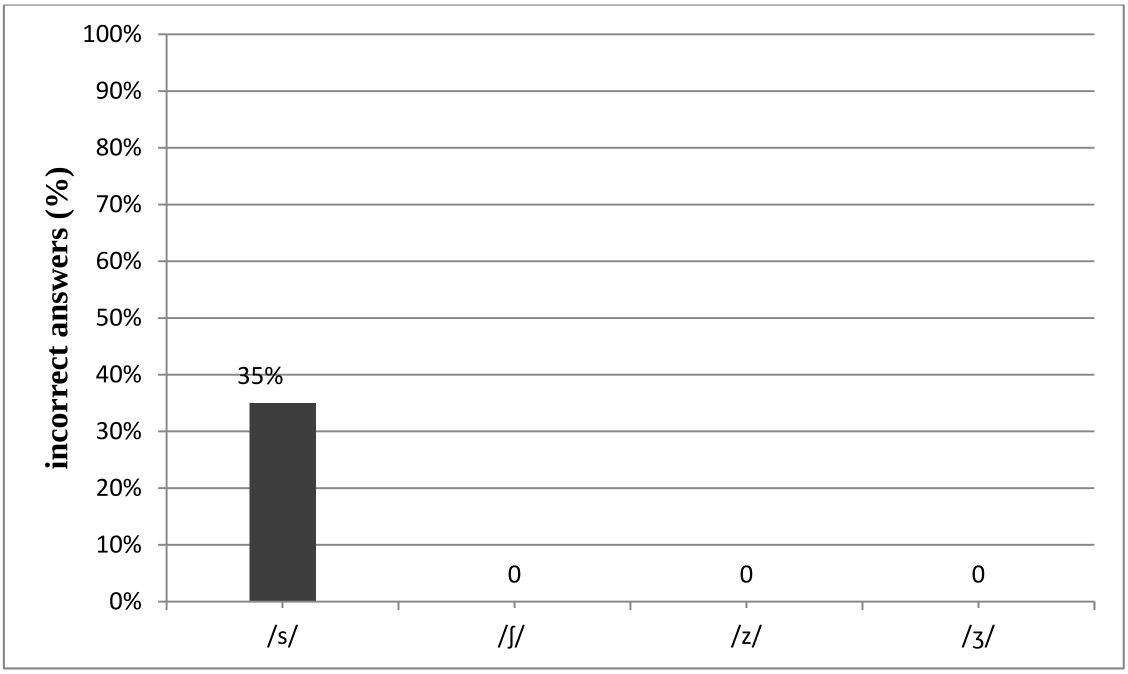


Figure 3.3 Perception Test Percentages of Incorrect Answers for Two of the Native Speakers

3.7 Rating the Author's Production of the Mandarin Consonants

In order to see if the author's productions of the four target Mandarin consonants (used in the AXB categorial test) matched those of the native Mandarin speakers, two tasks were carried out. In the first one, the author produced 12 Mandarin words containing the four Mandarin sounds /ʐ/, /s/, /ts/, and /ʂ/ (voiced retroflex fricative, voiceless dental fricative, voiceless alveo-dental affricate, and voiceless retroflex fricative respectively) which were rated by a native Mandarin speaker from China exactly the same way the 31 words were rated in the production task. These four Mandarin fricatives were the ones that the participants often used to substitute the four target English fricative sibilants with during the production task. Three of the four Mandarin sounds are actually represented by letters 's', 'z', and 'sh' in pinyin; so, it is not very difficult to see why the participants substituted many of the English target sounds, /s/, /z/, and /ʃ/, with the Mandarin sounds, /s/, /ts/, and /ʂ/. Table 3.1 shows the 12 Mandarin words produced by the author along with their ratings, the target sounds, and their transcription in Pinyin.

As can be seen here, the author's productions of the four Mandarin sounds were all rated as 'native-like' by the native Mandarin speaker.

Table 3.1 Mandarin Words Produced by the Author Along with Their Ratings

	Pinying	meaning	Initial sound	rating
肉	Rou	meat	/ʐ/	1
人	Ren	people	/ʐ/	1
让	Rang	allow	/ʐ/	1
三	San	three	/s/	1
四	Si	four	/s/	1
搜	Sou	search	/s/	1
是	Shi	yes	/ʂ/	1
说	Shuo	talk	/ʂ/	1
水	Shui	water	/ʂ/	1
在	Zai	be	/ts/	1
早	Zao	early	/ts/	1
最	Zui	most	/ts/	1

Whereas the native Mandarin speaker rated the four Chinese sounds in Chinese words in the first task, his job, in the second task, was to rate the four Mandarin sounds in the *English* words produced by the author in the perception test. Here again, he was asked to rate how close the four consonants produced by the author sounded to the four target Mandarin consonants. Table 3.2 shows the target English words containing the 4 Mandarin sounds produced by the author which were rated by the native Mandarin speaker. As can be seen, with the exception of /ʐ/ in 'usually' which was rated as 'very similar' to native pronunciation, all the target sounds in the rest of the words were rated as 'native-like' indicated by 1 in the rating column.

Table 3.2 English Words With Mandarin Consonants Rated by the Native Mandarin Speaker

no.	Word	Target sound	rating
1	Zip	/ts/	1
2	Kiss	/s̪/	1
3	Closed	/ts/	1
4	Washed	/ʂ/	1
5	Push	/ʂ/	1
6	Fast	/s̪/	1
7	Cause	/ts/	1
8	Shrink	/ʂ/	1
9	Missing	/s̪/	1
10	Miss	/s̪/	1
11	Usually	/ʐ/	2
12	Bishop	/ʂ/	1
13	Spark	/s̪/	1
14	Zeal	/ts/	1
15	Facing	/s̪/	1
16	Television	/ʐ/	1
17	Wasp	/s̪/	1
18	Shark	/ʂ/	1
19	Fashion	/ʂ/	1
20	Razor	/ts/	1
21	Ship	/ʂ/	1
22	Step	/s̪/	1
23	Maze	/ts/	1
24	Fish	/ʂ/	1
25	Crazed	/ts/	1
26	Sad	/s̪/	1
27	Cashed	/ʂ/	1
28	Busy	/ts/	1
29	Sick	/s̪/	1
30	Shrimp	/ʂ/	1
31	Pleasure	/ʐ/	1

3.8 Conclusion

This chapter discussed procedures related to selecting participants and data collection all the way to the methodology used in order to examine the English sibilant fricatives as produced by speakers from China. Four Chinese consonants were concisely discussed as well because of the important role they play in the study as the Chinese speakers often replaced the English sibilant fricative with them. The results of the study will be presented in the next chapter.

CHAPTER 4
FINDINGS AND DISCUSSION

In this chapter, the findings of the current research, coming mainly from a production task and a perception test, are presented and discussed. In the production task, the results of the participants' production of the 31 words containing English sibilant fricatives are discussed. The 31 words contain the four English sibilants in different positions. In the perception test, a discussion of the participants' perception of the author's recordings will be presented. The author's recordings include the same 31 words in the production task and contain the four English sibilant fricatives as well as incorrect sounds in place of the sibilants (produced by the participants in the production task) in order to see if the participants could pick the words with the correct English sibilants.

4.1 Production Task

As explained in Chapter 3, the 16 participants were recorded producing the 31 target words. Their production was then rated by five native speakers of American English on a five-level scale.

4.1.1 Production of /s/

Figure 4.1 shows the number of times the tokens containing /s/ produced by the speakers in trial one and two were rated by the native speakers on the scale from 1 (native-like) to 5 (very different). There were 10 words containing /s/ and 8 speakers in each gender group rated by five native speakers; so overall, the tokens were rated 400 times for each gender: $10 \times 8 \times 5 = 400$. Statistical analyses were conducted to see if there were any significant differences between the two trials as well as between the two genders. Levene's Test for Equality of Variances indicate that the groups have equal variances, $t(30) = -.073$, $p > .05$. Next an independent samples test was carried out to see if there were any significant differences between the ratings of the two gender groups. According to the results, it can be concluded that the difference is due to chance and not significant, $t(30) = -.073$, $p > .05$. The same thing was true for the trials as well where a paired sample t-test was conducted indicating that the difference among the ratings of the two trials was non-significant, $t(15) = 2.134$, $p = .05$.

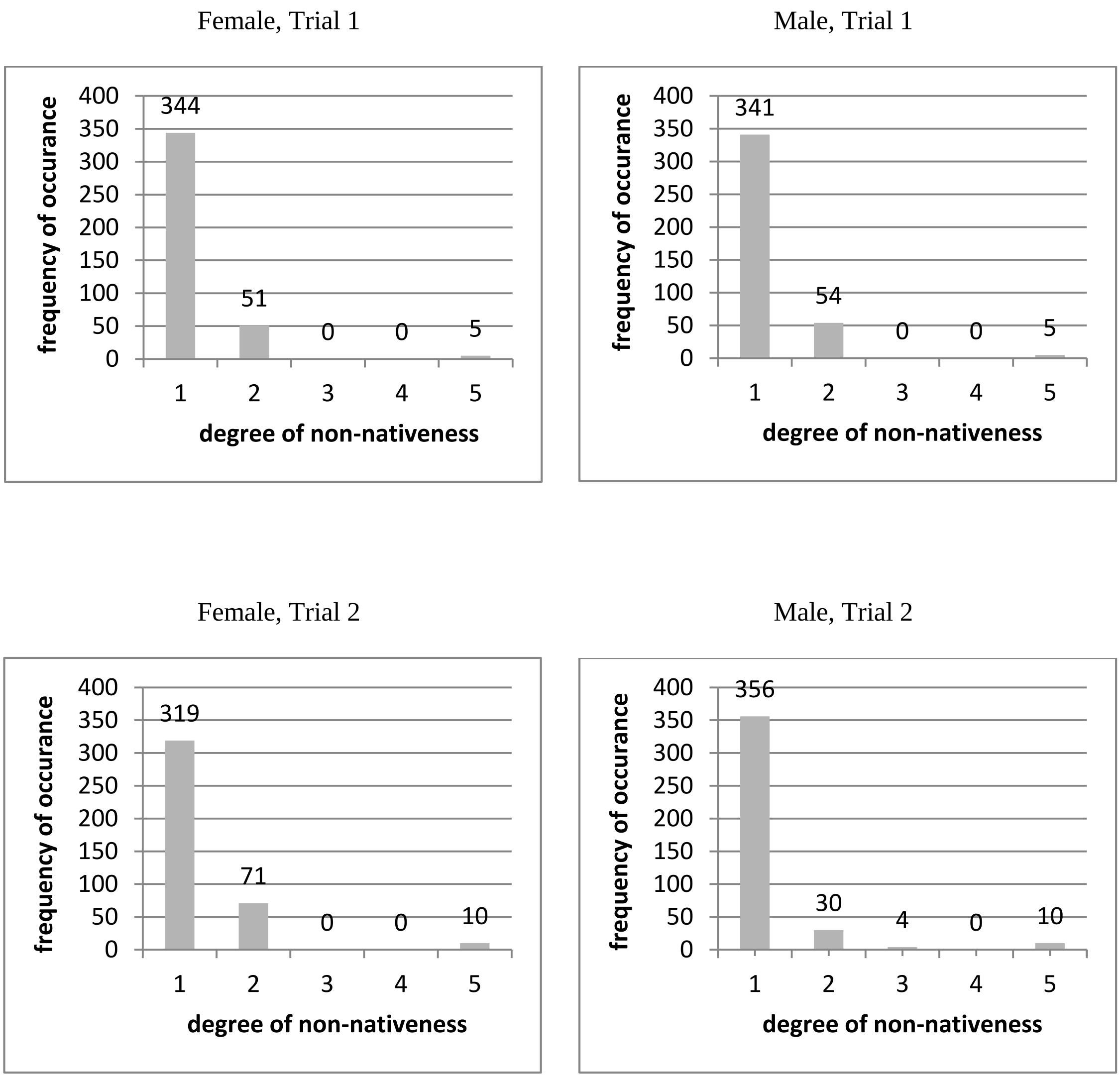


Figure 4.1 Degree of Nativeness as Perceived by Native Speakers for /s/

4.1.2 Production of /z/

Figure 4.2 shows the number of times the tokens containing /z/ in both trials were rated by the native speakers on the scale. There were 8 words containing /z/ and 8 speakers in each gender group. Overall, the tokens were rated 320 times for each gender: 8 × 8 × 5 = 320. Based on the Levene's test, equal variances were assumed here as well, $t(30) = .786$, $p > .05$. To see if gender had a significant effect, an independent samples test was carried which revealed no significant differences between the ratings of the two gender groups, $t(30) = .786$, $p > .05$. Next a paired sample t-test was carried out to see if the two trials differed significantly from each other which indicated that the ratings were significantly different for the two trials, $t(15) = -2.527$, $p < .05$.

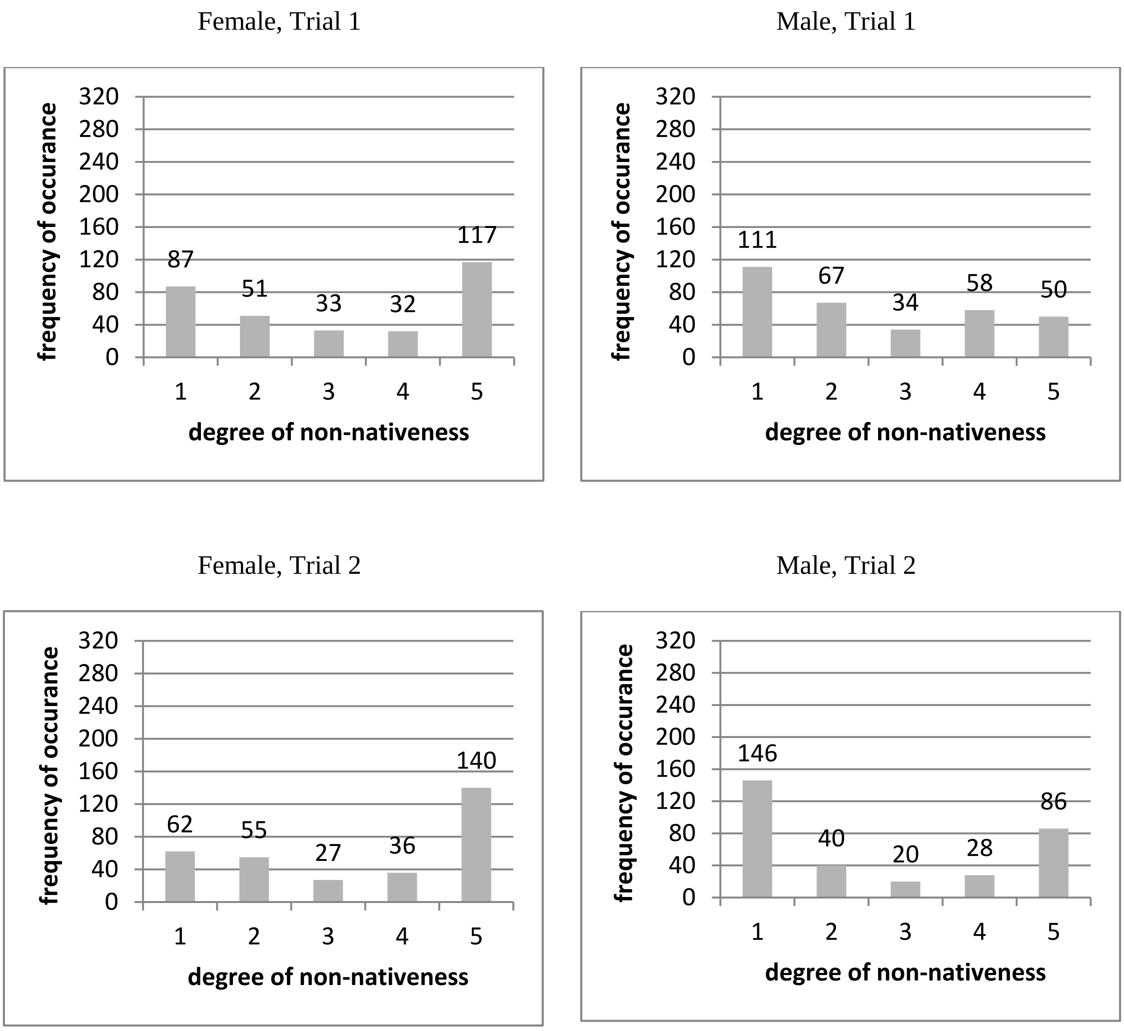


Figure 4.2 Degree of Nativeness as Perceived by Native Speakers for /z/

4.1.3 Native speakers' perception of /ʃ/

Male speakers were perceived as slightly more native-like in tokens containing /ʃ/. Figure 4.3 shows the number of times the tokens containing /ʃ/ in both trials were rated by the native speakers on the scale. There were 10 words containing /ʃ/ and 8 speakers in each gender group; so overall, the tokens were rated 400 times for each gender: 10 × 8 × 5 = 400. Running a Levene's test here as well, equal variances were assumed, $t(30) = .518$, $p > .05$. The independent samples test carried out shows no significant differences between the ratings of the two gender groups, $t(30) = .518$, $p > .05$. A paired sample t-test was also carried out which revealed no significant statistical differences between the two trials, $t(15) = -.754$, $p > .05$.

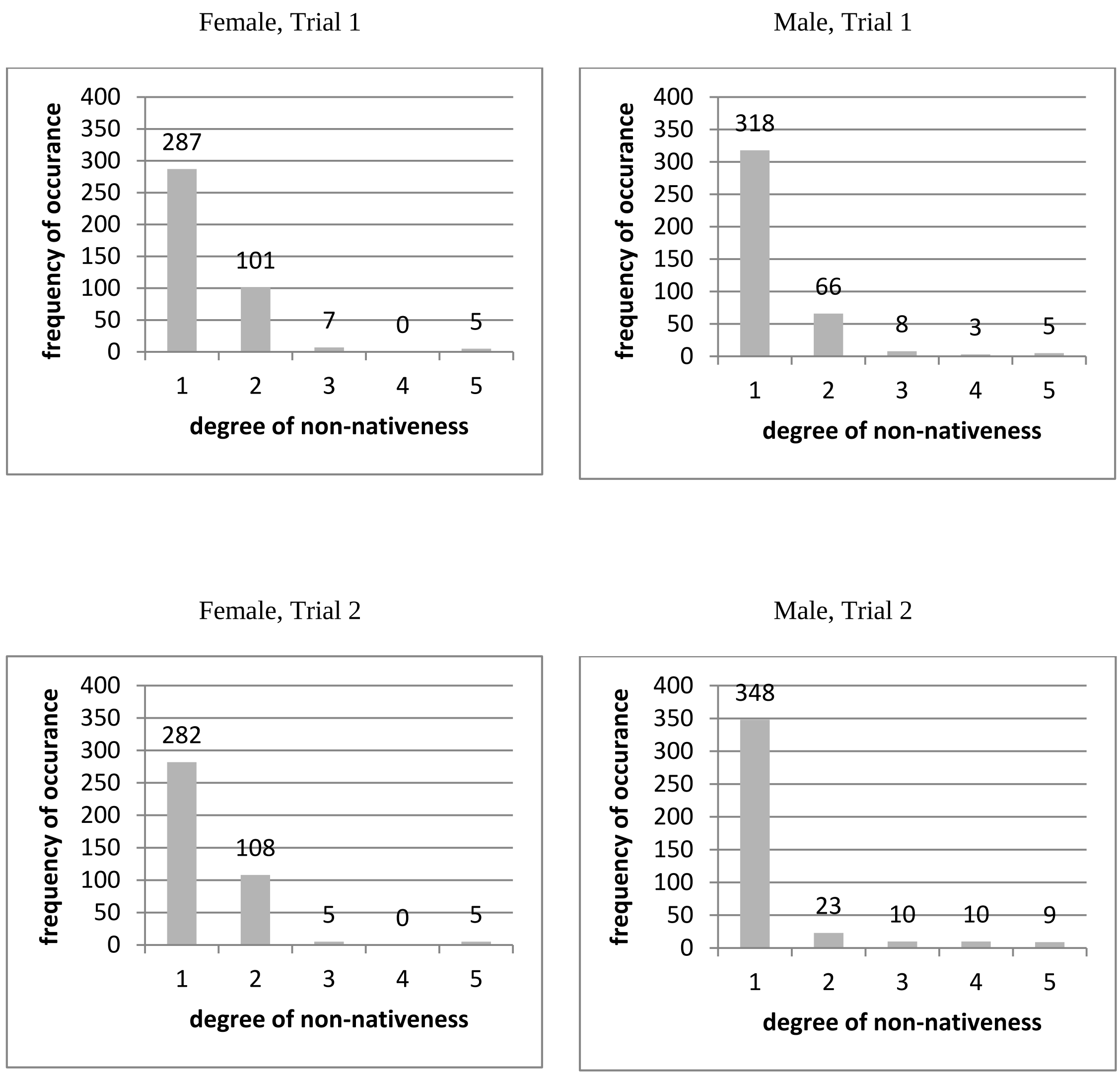


Figure 4.3 Degree of Nativeness as Perceived by Native Speakers for /ʃ/

4.1.4 Production of /ʒ/

Figure 4.4 shows the number of times the tokens containing /ʒ/ were rated by the native speakers on the scale. There were only 3 words containing /ʒ/ and 8 speakers in each gender group; so overall, the tokens were rated 320 times for each gender: 3 × 8 × 5 = 120. In the first trial, female speakers appeared to be slightly more native-like while, in the second, it was the male speakers who were perceived to be slightly more native-like; however, an independent samples test done (after equal variances were assumed though the Levene's test, $t(30) = .395$, $p > .05$) revealed no significant differences between the ratings of the two gender groups, $t(30) = .395$, $p > .05$. The paired sample t-test carried out showed no significant differences between the two trials either, $t(15) = .0$, $p > .05$.

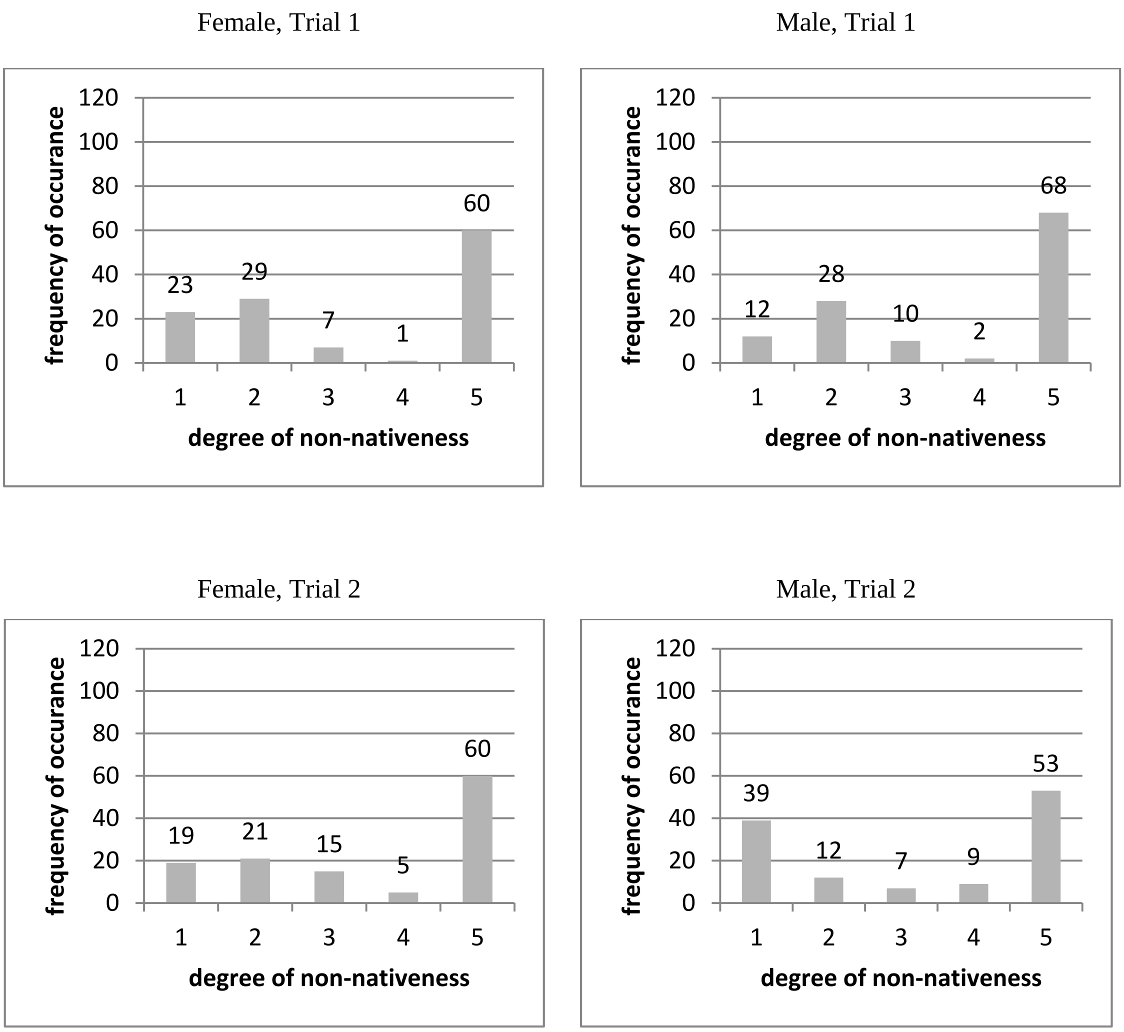


Figure 4.4 Degree of Nativeness as Perceived by Native Speakers for /ʒ/

As can be seen in the ratings for the degree of non-nativeness so far, there are no significant differences between the two trials and also between the two genders for each sibilant fricative. The only exception is the significant difference between the two trials of the production of /s/. This indicates that there are no wide gaps between speakers' performance across the two trials as well as the two genders for each fricative.

4.1.5 Comparison of the Ratings

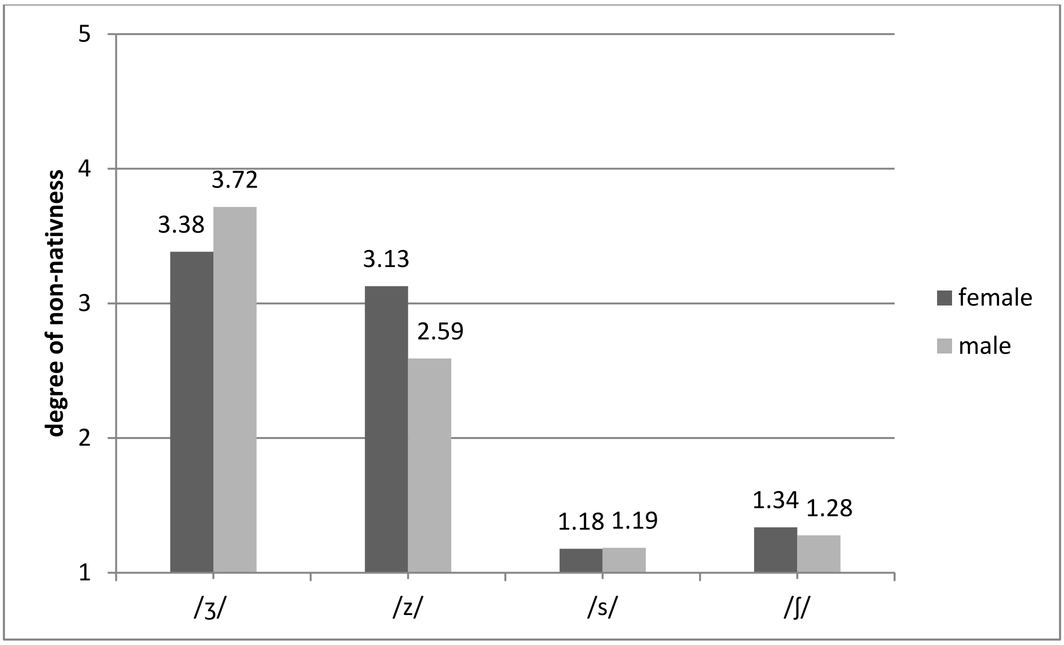


Figure 4.5 Mean Degree of Nativeness Ratings for the Four Fricatives in Trial One

As can be seen in Figure 4.5, male speakers were perceived to be slightly more native-like for /z/ and /ʃ/ in trial one; this is more evident in the case of /z/. However, female speakers were rated as closer to native-like for tokens containing /ʒ/ and just slightly so for /s/. In figure 4.6, Male speakers were perceived as more native-like for all the fricatives in the second trials and as with the first trial, this difference is the most evident in tokens containing /z/.

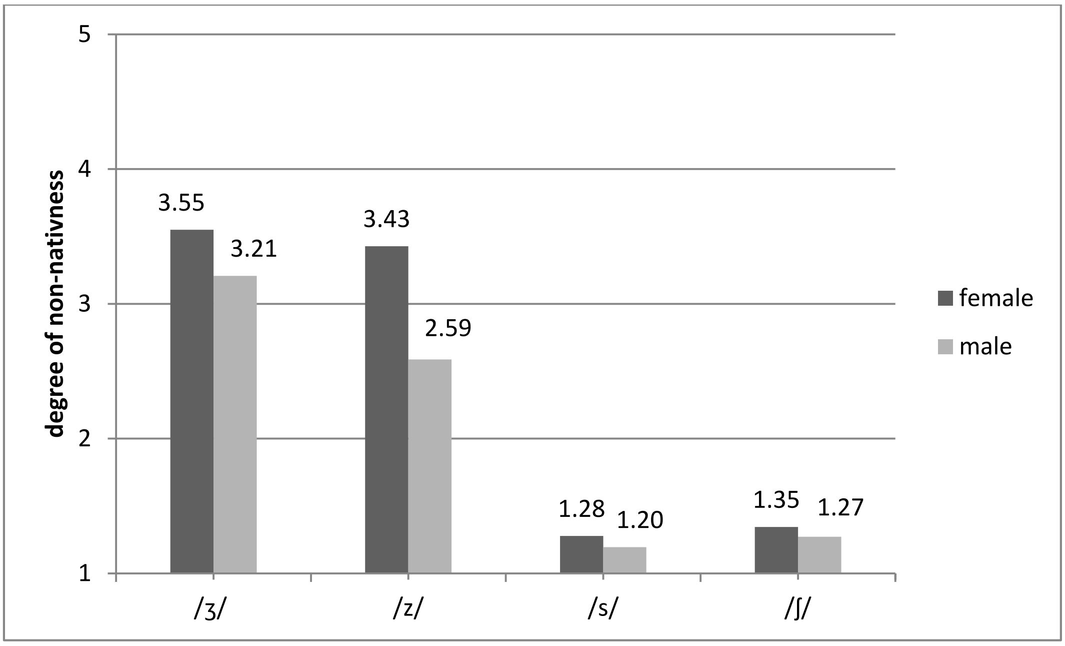


Figure 4.6 Mean Degree of Nativeness Ratings for the Four Fricatives in Trial Two

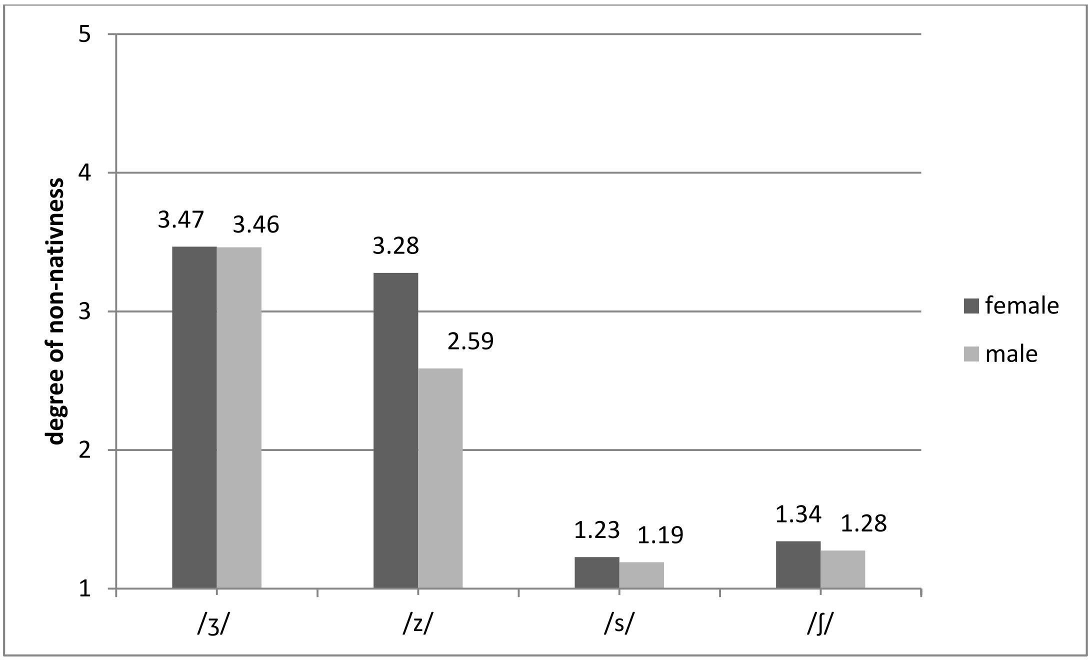


Figure 4.7 Overall Mean Degree of Nativeness Ratings for the Four Fricatives

Figure 4.7 shows mean degree of non-nativeness for the four fricatives in both trials. /ʒ/ has the highest degree of non-nativeness for both male and female speakers followed by /z/. The means for /s/ and /ʃ/ were not very different from each other because many tokens containing /s/ and /ʃ/ were rated as native-like or ‘very similar’ to native pronunciation in both trials and for both genders.

According to what has been presented so far in this chapter, it is evident (based on the native speaker raters’ perception) that there are significant differences in the degree of non-nativeness among the four fricatives; the overall effect of gender seemed to be significant as well. A mixed design ANOVA was carried out with native speakers’ perception of the degree of non-nativeness (across both trials) as a within-subjects factor and Sibilant and Gender as between-subjects factors. Both Gender and Fricative effects as well as their interaction were highly significant [Fricative: $F(3, 2472) = 644.156$, $p < .05$, Gender: $F(1, 2472) = 20.661$, $p < .05$, Gender × Fricative $F(3, 2472) = 16.00$, $p < .05$]. A Tukey post-hoc test then revealed that the fricatives significantly different from each other were /ʒ/ and /z/, /ʒ/ and /s/, /ʒ/ and /ʃ/, /z/ and /s/ and finally /z/ and /ʃ/, $p < .05$. As was mentioned earlier in this chapter at the end of section 4.1.4, there were almost no significant differences between the performances of the speakers across the two trials and genders for *each* fricative. However, based on the ANOVA results just mentioned, the overall effect of gender (across all the fricatives and trials) was significant.

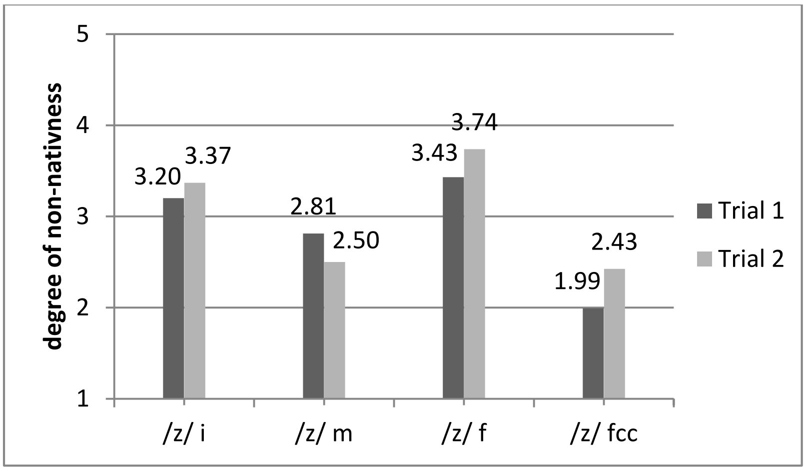
As can be seen in Figure 4.7, /s/ had the lowest degree of non-nativeness which means that the tokens containing this sibilant were perceived as the most native-like by the raters. As was mentioned in Chapter 2, section 2.3, English and Mandarin /s/ differ only in their phonetic details (Li et al., 2007). This means that there is a high probability that regardless of which /s/ the participants used in the production of the English words, the majority of the tokens were perceived as native-like. For the purposes of this study, however, native speaker ratings suffice. In another similar study, also mentioned in

Chapter 2, section 2.3, Chang et al. (2009) emphasized that pairs such as English and Mandarin /s/ can undergo “equivalence classification” and therefore “become indistinguishable from each other” (Chang et al., 2009, p. 38). Applying acoustic analysis measurements, they found that about half of their participants could distinguish English and Mandarin /s/ in their production with the majority of the distinguishers being among the Mandarin heritage speakers and Native English speaking late learners of Mandarin while few monolingual speakers of Mandarin were able to do so.

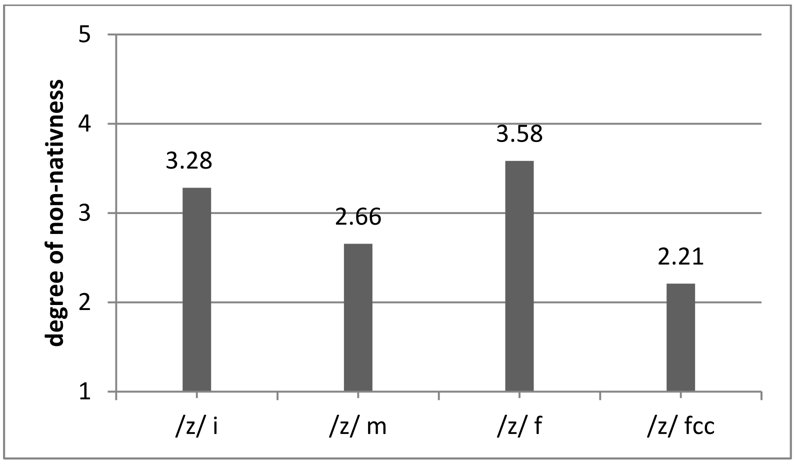
The sound /z/ does not occur in Mandarin and so it was expected to be challenging for the participants to produce. As can be seen in Figure 4.7, there was a rather high degree of non-nativeness associated with /z/ and it is the second most problematic English sibilant after /ʒ/ for the participants; this is consistent with Deterding’s findings (2006) that among the English sibilant fricatives the production of /ʒ/ followed by /z/ caused the most difficulty for his participants (Chapter 2, section 2.4). As for /ʃ/, the degree of non-nativeness was relatively low for this sibilant in this study; however, Chang et al. (2009) comparing the production of /ʃ/ in English words and /ʂ/ in Mandarin Chinese words by heritage speakers, late learners of Mandarin, and monolingual Mandarin speakers, found that only about half the participants – mostly heritage speakers – could distinguish the two. The current study, in contrast, only includes L1 Mandarin speakers who are late L2 learners of English. The speakers’ productions will be further discussed later in this chapter (section 4.3.5) along with a discussion of the perception results. This will ensure a more complete picture of the nature of the speakers’ production and perception. In the literature on L2 perception and production too (as was mentioned in Chapter 2, section 2.1.1) there is not always a clear distinction between production and perception.

4.1.6 Comparison of Ratings Based on the Fricative Positions

The final thing that was considered in the production task was the comparison of the native speakers' ratings based on the position of the sibilants in the target words. Three of the fricatives were considered here: /s/, /z/, and /ʃ/. /ʒ/, however, only occurred in three words in medial position and therefore was not considered here because it could not be compared against /ʒ/ in other positions.



Trials 1&2



Trials 1&2 combined

Figure 4.8 Comparison of Rates Based on Position of /z/

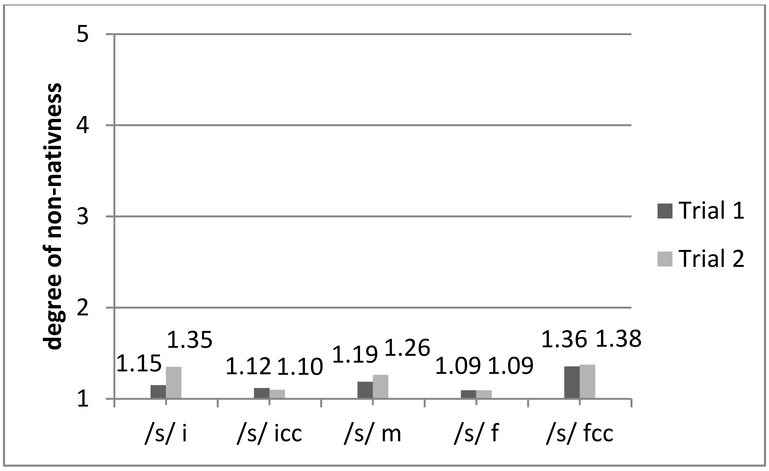
The sound /z/ occurred in eight of the 31 target words in four different positions of initial (*i*), medial (*m*), final (*f*), and final consonant cluster (*fcc*). As can be seen in

Figure 4.8, there is a similar pattern in terms of the degree of non-nativeness rating across the two trials in the sense that for both genders final /z/ was rated as the most non-native, followed by initial /z/, then medial /z/, and finally the ‘final consonant cluster’ (*fcc*) /z/.

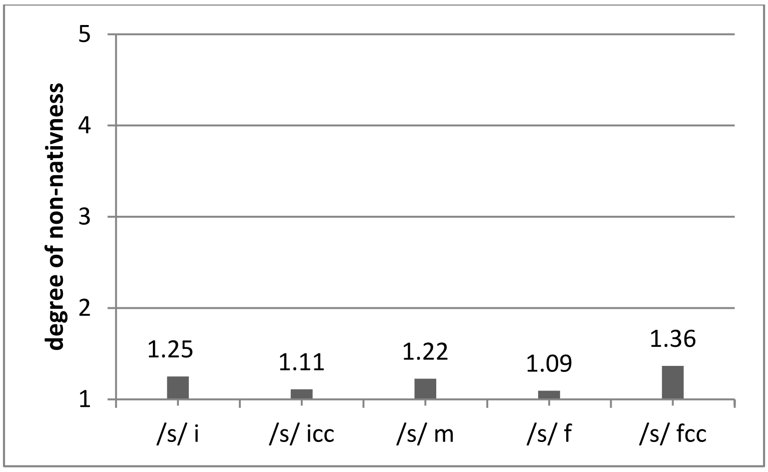
It should, however, be mentioned that many of the speakers inserted a vowel in final consonant clusters containing /z/ either at the end or between the two consonants of the cluster. There were two words containing /z/ in final consonant clusters, “closed” and “crazed”. About half the speakers had vowel insertions at the end of the word “closed” in either trial one or two or both with the majority of them inserting a vowel at the end of the cluster, rather than between the two consonants. About half the speakers also had vowel insertions at the end of the word “crazed” with the majority of them again inserting a vowel at the end of the cluster rather than inside. The vowel inserted at the end of the clusters was mostly a schwa. Finally, the /z/ produced by the speakers in the final consonant cluster position was rated as the most native-like compared with /z/ in other positions. The results of a one-way between groups ANOVA reveal that the means for the four different positions are significantly different from each other: F (3, 124) = 5.06, $p < .05$. More specifically, a Tukey post-hoc test indicates that only the means for final and final consonant cluster were significantly different from each other, $p < .05$.

The consonant /s/ occurred in 10 of the 31 target words in five positions of initial (*i*), initial consonant cluster (*icc*), medial (*m*), final (*f*), and final consonant cluster (*fcc*). As can be seen in Figure 4.9, the pattern across the two trials is relatively similar; that is, in trial one the order of positions from the most non-native to the least is: *fcc, m, I, icc, f* , and in the second trial it is: *fcc, I, m, icc, f*. Here the speakers inserted a vowel in final as well as initial consonant clusters containing /s/. About half the speakers produced the word “fast” as /fæstə/ (with a schwa at the end). A similar pattern was observed with “wasp” but only a couple of speakers had insertions at the end of this consonant cluster. None of the speakers inserted a vowel inside the clusters.

Three of the speakers produced the word “spark” with an initial schwa; similarly, two of the speakers produced the word “step” with an initial schwa, but there were no insertions between the two consonants of each cluster. A one-way between groups ANOVA carried out here indicates that overall the means for the five different positions



Trials 1&2

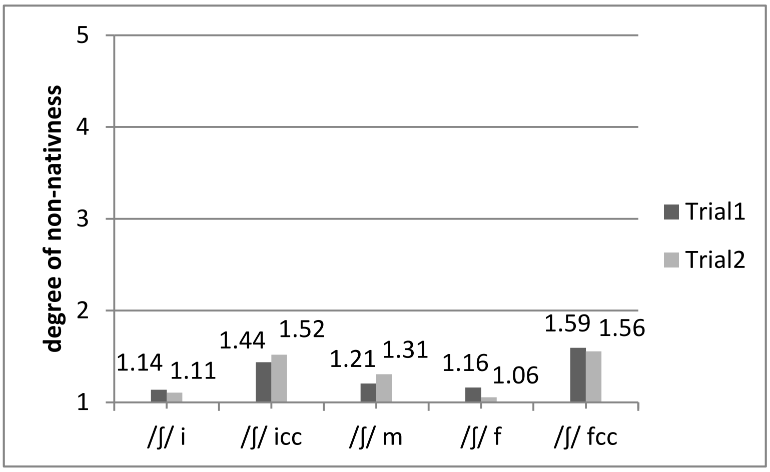


Trials 1&2 combined

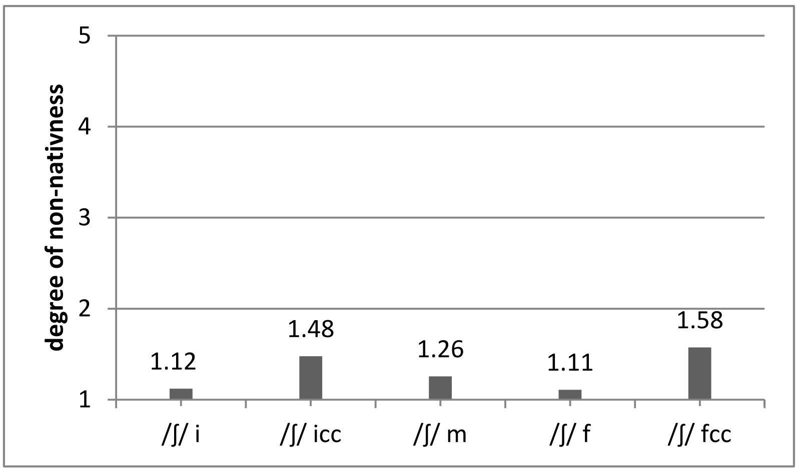
Figure 4.9 Comparison of Rates Based on Position of /s/

are significantly different from each other: F (4, 155) = 2.742, p < .05. However, the results of a Tukey post-hoc test reveal no significant differences between the means of any two positions.

The sound /ʃ/ occurred in ten of the 31 target words in five different positions of initial (*i*), initial consonant cluster (*icc*), medial (*m*), final (*f*), and final consonant cluster (*fcc*). Here too, the pattern across the two trials is relatively similar with the *fcc* /ʃ/ being the most non-native in trial one followed by *icc*, *m*, *f*, and finally *I* (least non-native); the order for the second trial (from the most non-native to the least) is: *fcc*, *icc*, *m*, *i*, and finally *f*, as can be seen in Figure 4.10.



Trial 1



Trial 2

Figure 4.10 Comparison of Rates Based on Position of /ʃ/

There were insertions in consonant clusters containing /ʃ/ as well. About half the speakers inserted a schwa at the end of the word "washed". The same pattern was observed with "cashed" except that one speaker inserted a vowel inside as well as at the end of the consonant cluster. A one-way between groups ANOVA results show the means for the five different positions are significantly different from each other: F (4, 155) = 11.977, $p < 0.05$. A post-hoc Tukey test further reveals that the pairs 'initial and initial consonant cluster', 'initial and final consonant cluster', 'initial consonant cluster and medial', 'initial consonant cluster and final', 'medial and final consonant cluster', and 'final and final consonant cluster' are all significantly different from each other, $p < 0.05$.

4.2 Phonetic Transcription of the Sibilants

All the sounds that got a rating of 'one' by native speaker raters were judged by the author to be the target sound (the correct sibilant). As an example, the sound at the beginning of the word 'zip' (the voiced alveolar, /z/) was judged by all five native speakers to be the correct sound in the production of two Chinese female speakers, f5 and f7, and two male Chinese speakers, m7 and m8 in trial 1. These four speakers, therefore, got a rating of 'one' by all the native speaker raters. Similarly, the author also perceived the sound as /z/ in the production of the four speakers. Three other Chinese speakers were judged by the author to have produced /z/ but not all the five raters gave them a rating of 'one'; they are f1, m1, and m6 but they had the lowest ratings right after the four speakers who received a rating of one (Table 4.1). This last point is a general pattern observed in the data across the speakers and trials; that is, other than the speakers who got a rating of one, the rest of the speakers who were also perceived by the author to have produced the target sound had in most cases received the lowest ratings from the five raters.

Table 4.1: Mean Rating and Initial Sounds as Perceived by Raters and Author for 'zip'

f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08
/z/	/ð/	/ts/	/ð/	/z/	/ð/	/z/	/ts/
1.8	5	3.4	5	1	5	1	5

m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
/z/	/ts/	/ts/	/ð/	/ts/	/z/	/z/	/z/
2.2	4	3	4.8	4.8	1.4	1	1

Table 4.2: Mean Rating for Target Sounds, Non-Target Sounds, and all the Sounds

	Trial 1			Trial 2		
	Mean rate for target sounds	Mean rate for non-target sounds	mean rate for all sounds	Mean rate for target sounds	Mean rate for non-target sounds	mean rate for all sounds
zip	1.2	4.145	3.225	1.343	4.444	3.088
kiss	1.057	1.2	1.075	1.036	1.16	1.075
closed	1.156	3.8	2.313	1.2	3.371	2.15
washed	1.229	2.4	1.888	1.543	2.489	2.075
push	1.038	0[1]	1.038	1.123	1.867	1.263
fast	1.057	2.1	1.188	1.071	1.3	1.1
cause	4.8	4.587	4.6	0	3.988	3.988
shrink	1.071	3.5	1.375	1.092	2.8	1.413
missing	1.077	2.467	1.338	1.129	1.6	1.188
usually	1.3	4.971	4.513	2.133	4.831	4.325
miss	1.031	1.467	1.113	1.086	1.3	1.113
bishop	1.077	3.333	1.5	1.123	1.533	1.2
spark	1.057	1.5	1.113	1.062	1.267	1.1
zeal	1.25	4.267	3.513	1.52	4.127	3.313
facing	1.092	1.6	1.188	1.117	1.4	1.188
television	1.311	4.257	2.6	1.556	3.714	2.5
wasp	1	3.25	1.563	1.022	2.371	1.613
shark	1.075	0	1.075	1.147	1.4	1.163
fashion	1.014	1.8	1.113	1.129	1.8	1.213
razor	1.267	3.92	2.925	1.1	4.367	3.55
ship	1.107	1.6	1.138	1.067	1.8	1.113
step	1.031	1.333	1.088	1.138	0	1.138
maze	1.04	3.709	2.875	1.4	2.973	2.875
fish	1.027	1.8	1.075	1.04	1.4	1.063
crazed	1.125	3.95	2.538	1.156	2.714	1.838
sad	1.26	2.033	1.55	1.127	1.32	1.188
cashed	1.129	1.9	1.225	1.033	1.35	1.113
busy	1.044	3.4	2.075	1.133	3.286	2.075
sick	1.083	1.35	1.15	1.062	1.333	1.113
shrimp	1.145	2.8	1.663	1.14	2	1.463
pleasure	1.2	3.855	3.025	1.333	4.4	3.825

[1] Zero indicates that the author found all the target sounds in those tokens as native-like.

Table 4.2 shows mean native speakers' ratings for target sounds (the sounds perceived by the author to be the correct English sibilant) and non-target sounds (the sounds perceived by the author to be different from the correct English sibilant) as well as mean native speakers' rating for all the sounds. As can be seen, in all the 31 words, the mean rate for target sounds is always less than the overall mean and the mean rate for non-target sounds is more than the overall mean which means the author's perception of the participants' production is close to that of the five raters.

4.3 Perception Test

The participants also took an AXB categorial discrimination test where recordings of each of the 31 words produced by the author were played twice (two trials) to the listeners. Each word was produced three times: twice with the 'correct' fricative and once with the 'wrong' fricative. The participants' job was to choose the word with the wrong fricative (odd item).

4.3.1 Participants' Perception of /s/

Here the listeners had to distinguish English alveolar /s/ form Mandarin dental /s/. Figure 4.11 shows the number of wrong answers the listeners gave when choosing the odd items in words containing /s/. 10 of the 31 words contained /s/ and there were 8 people in each gender group; hence, the participants in each group in each trial listened to 80 instances of the AXB categorical test containing /s/ which adds up to a total of 320 instances for all the listeners across both trials. As can be seen in Figure 4.11, in both trials, female participants picked the odd items (correct answers) more often than did the male participants and a repeated measures analysis of variance results confirms the gender effect was significant; however, there are no effects for 'trial' nor for 'gender × trial' [$F(1, 28) = 4.287$, $p < .05$, $F(1, 28) = .603$, $p > .05$, $F(1, 28) = 067$, $p > .05$].

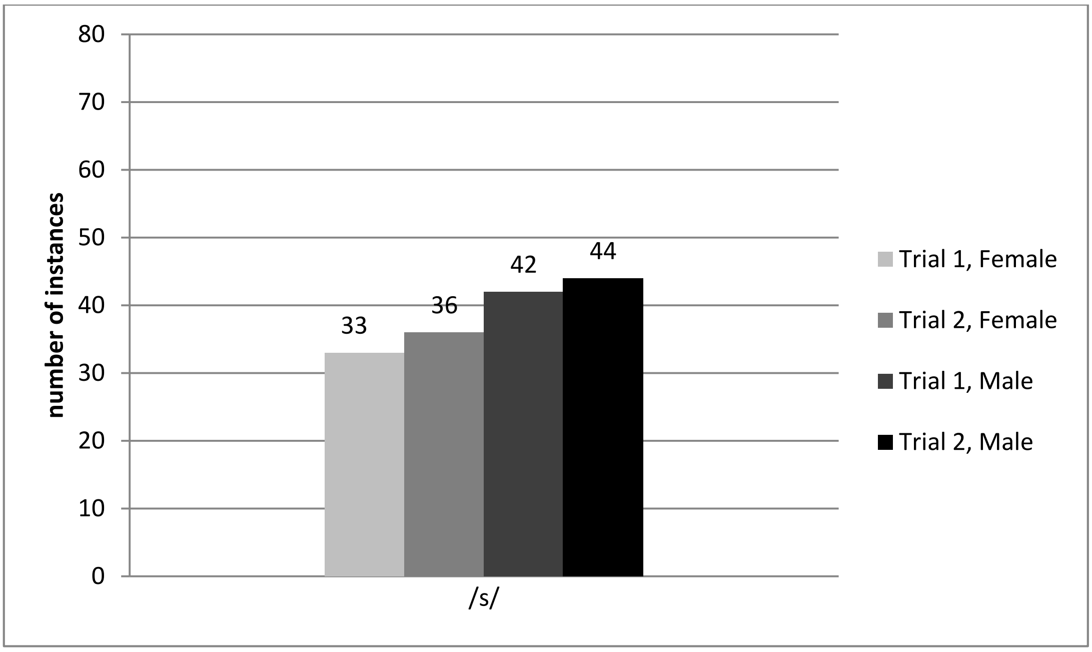


Figure 4.11 Number of Incorrect Answers to Words Containing /s/

4.3.2 Participants' Perception of /ʃ/

Figure 4.12 shows the number of wrong answers the listeners gave when choosing the odd items in words containing /ʃ/. The odd items were the words in which /ʃ/ had been

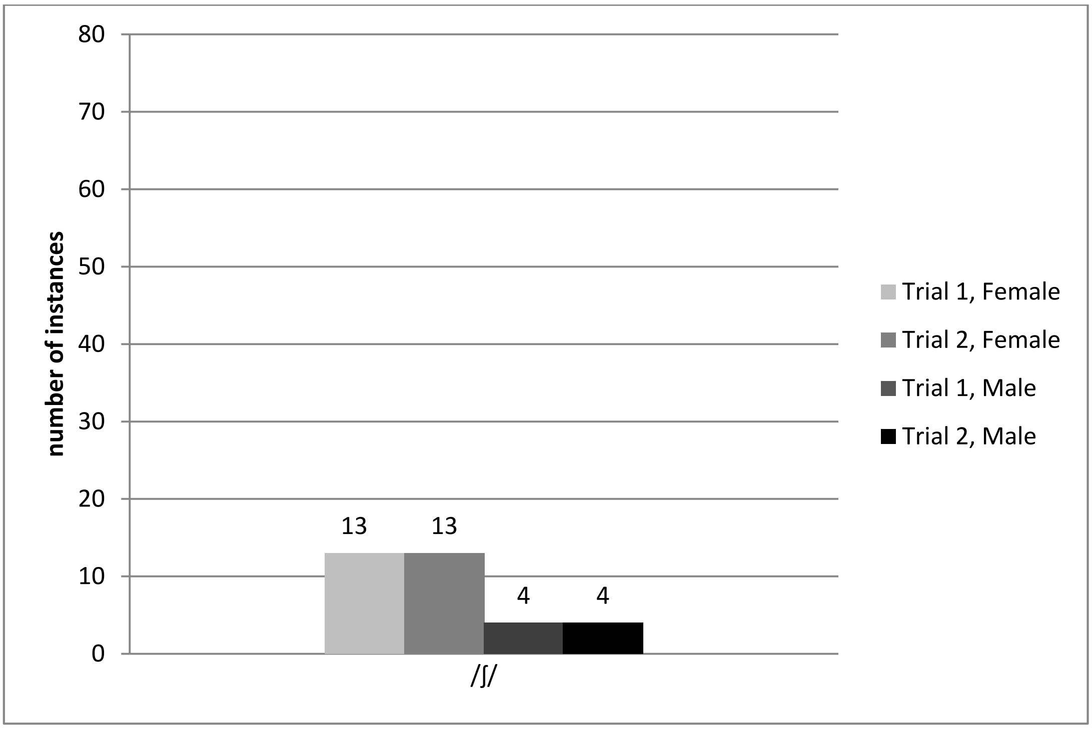


Figure 4.12 Number of Incorrect Answers to Words Containing /ʃ/

replaced by voiceless retroflex /ʂ/. The participants listened to 80 instances of the AXB categorical test in each trial for each gender group (10 words containing /ʃ/, 8 listeners in each gender group). The male participants, in both trials, picked the odd items (correct answers) more often than did the female speakers (Figure 4.12). A repeated measures analysis of variance results shows the significance of this gender effect; however, just like the previous test, there are no effects for 'trial' nor for 'gender × trial' [$F(1, 28) = 2$, $p < .05$, $F(1, 28) = 0$, $p > .05$, $F(1, 28) = 0$, $p > .05$].

4.3.3 Participants' Perception of /z/

Figure 4.13 shows the number of wrong answers the listeners gave when choosing the odd items in words containing the fricative /z/. Whereas there was only one odd sound replacing the English sibilants in each of the two previous tests, here there were three: /ð/ (voiced dental non-sibilant fricative), /s/ (voiceless dental sibilant fricative), and /ts/ (voiceless alveolar sibilant affricate). These three consonants were the ones that the participants often replaced /z/ with in the production task. The listeners then had to distinguish /z/ from /ð/, /z/ from /s/, and /z/ from /ts/ in two trials for each pair; there were also 8 words containing /z/ and the same participants in each gender group which means that there were 192 instances of the AXB categorical test for this fricative in each trial for each group. As can be seen in Figure 4.13, the female participants picked the odd items (correct answers) a couple of times more than did the male participants in each trial; however, this was not statistically significant. A repeated measures analysis of variance results reveal that there are no effects for 'gender', 'trial', or 'gender × trial' [$F(1, 28) = 1$, $p > .05$, $F(1, 28) = 1$, $p > .05$, $F(1, 28) = 1$, $p > .05$].

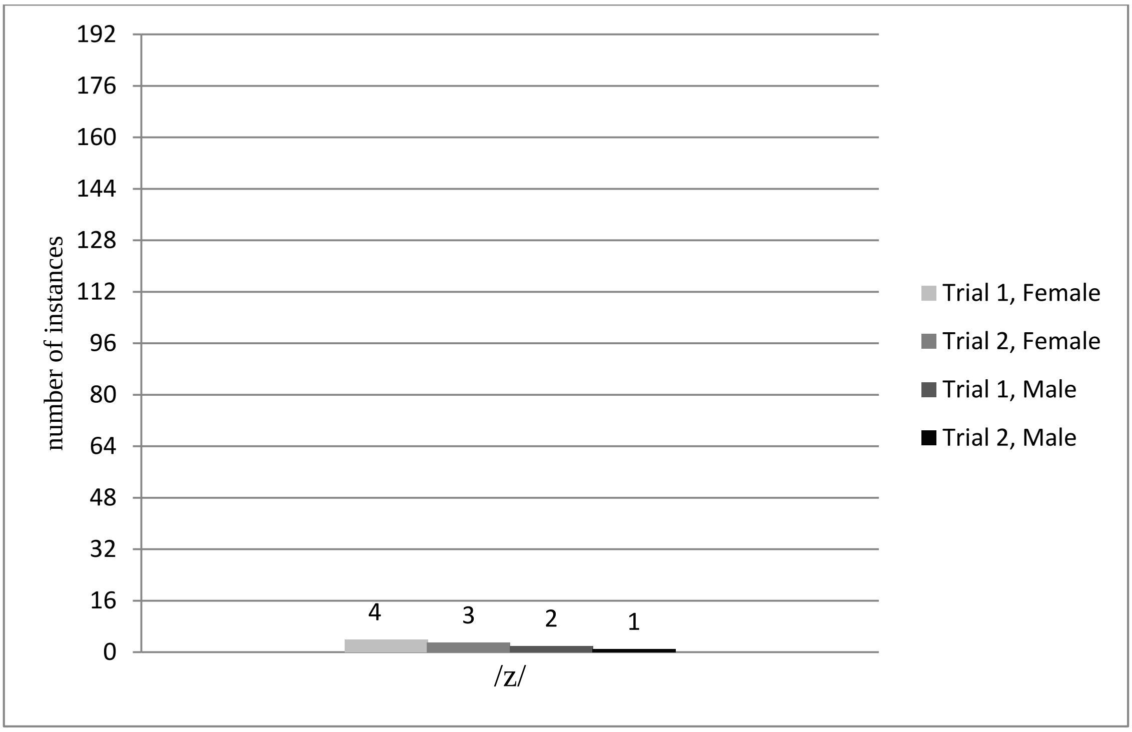


Figure 4.13 Number of Incorrect Answers to Words Containing /z/

4.3.4 Participants' Perception of /ʒ/

Figure 4.14 shows the number of wrong answers the listeners gave when choosing the odd items in words containing the fricative /ʒ/. There were three odd consonants – /ʂ/ (voiceless retroflex sibilant fricative), /tʂ/ (voiceless retroflex sibilant affricate), and /ʐ/ (voiced retroflex sibilant fricative) – which the participants used to replace /ʒ/ with in the production task; this means that there were three contrasting pairs here: '/ʒ/ /ʂ/', '/ʒ/ /tʂ/', and '/ʒ/ /ʐ/'. There were also three words containing /ʒ/ and the same participants in each gender group. This adds up to 72 instances of the AXB categorical test for /ʒ/. As can be seen in Figure 4.14, the number of wrong answers is the same across the two trials and gender groups except for male speakers in trial one who had one additional wrong answer. A repeated measures analysis of variance results shows that there are no 'gender', 'trial', or 'gender × trial' effects here either [$F(1, 28) = 1$, $p > .05$, $F(1, 28) = 0$, $p > .05$, $F(1, 28) = 0$, $p > .05$].

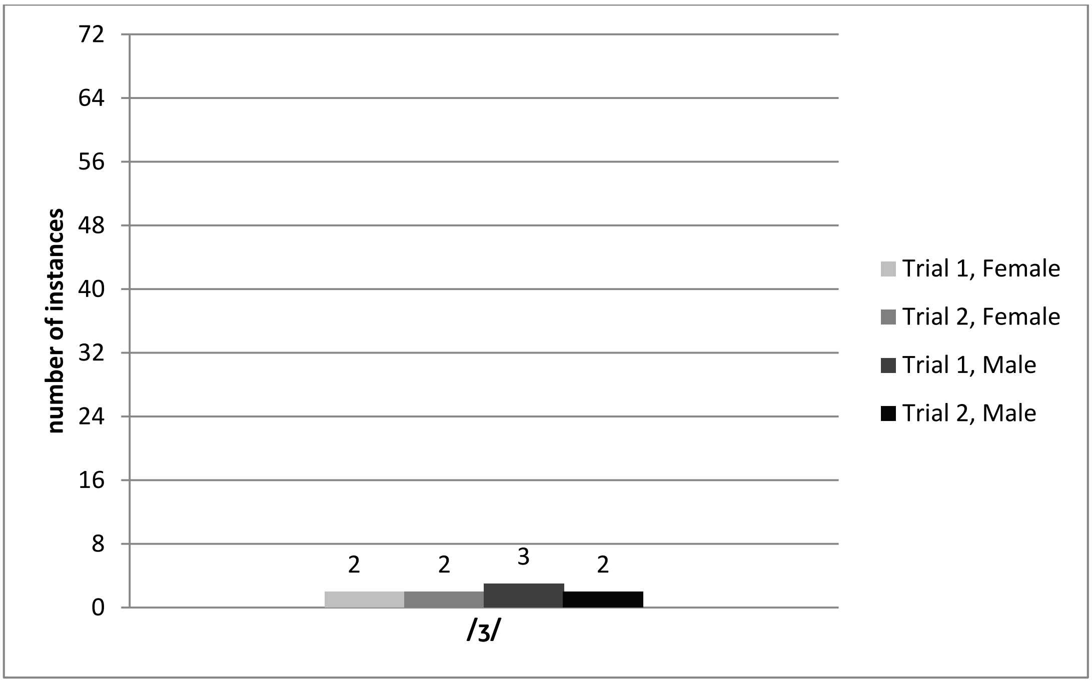


Figure 4.14 Number of Incorrect Answers to Words Containing /ʒ/

4.3.5 Comparison of the Ratings

Figure 4.15 shows mean perception test percentages of wrong answers for each of the four sibilants in trial one for male and female participants. What is striking here is that the participants did not have much difficulty perceiving the odd items for /z/ and /ʒ/; only about 2.78% of the instances the tokens were incorrectly identified by female participants as being odd items for /ʒ/ in the first trial; this was 4.17% for the males. The percentages of incorrect answers were even lower for /z/ in trial one with 2.08% of tokens wrongly identified as odd items by females and only 1.04% by males. Tokens containing /ʃ/ were misperceived by the participants as odd items more often than were the ones containing /ʒ/ and /z/; this was especially evident in the case of the females whose percentage of wrong answers was 16.25 as compared to only 5% for the males. /s/, however, was different from the other three fricatives in that about 41% and 52% of the tokens containing this sound were wrongly perceived by the females and males respectively.

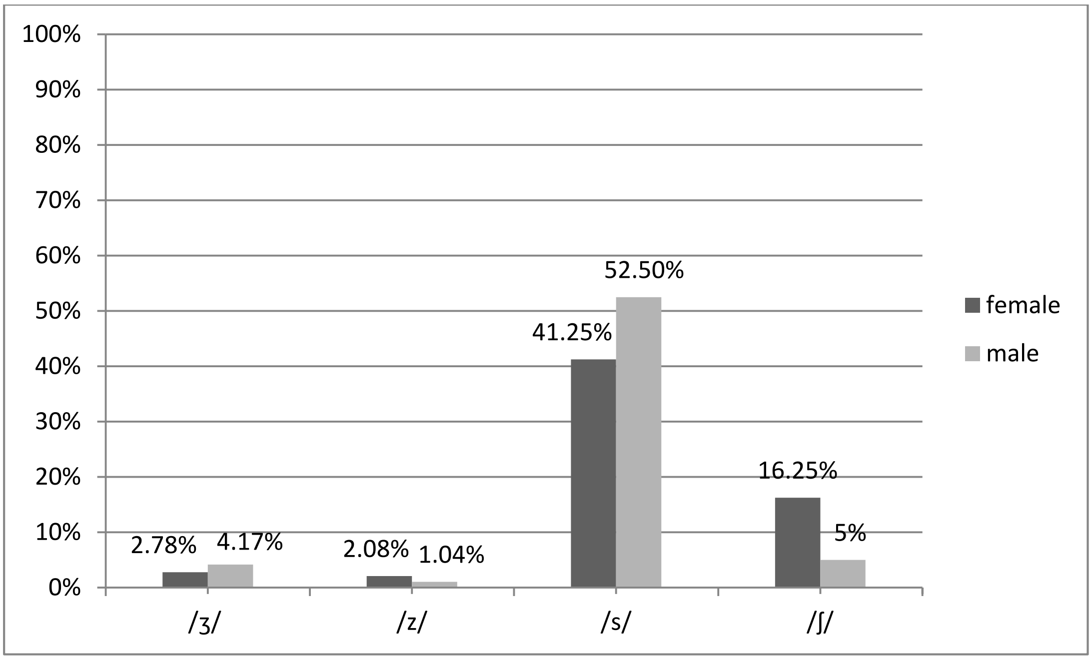


Figure 4.15 Mean Perception Test Percentages of Incorrect Answers in Trial 1

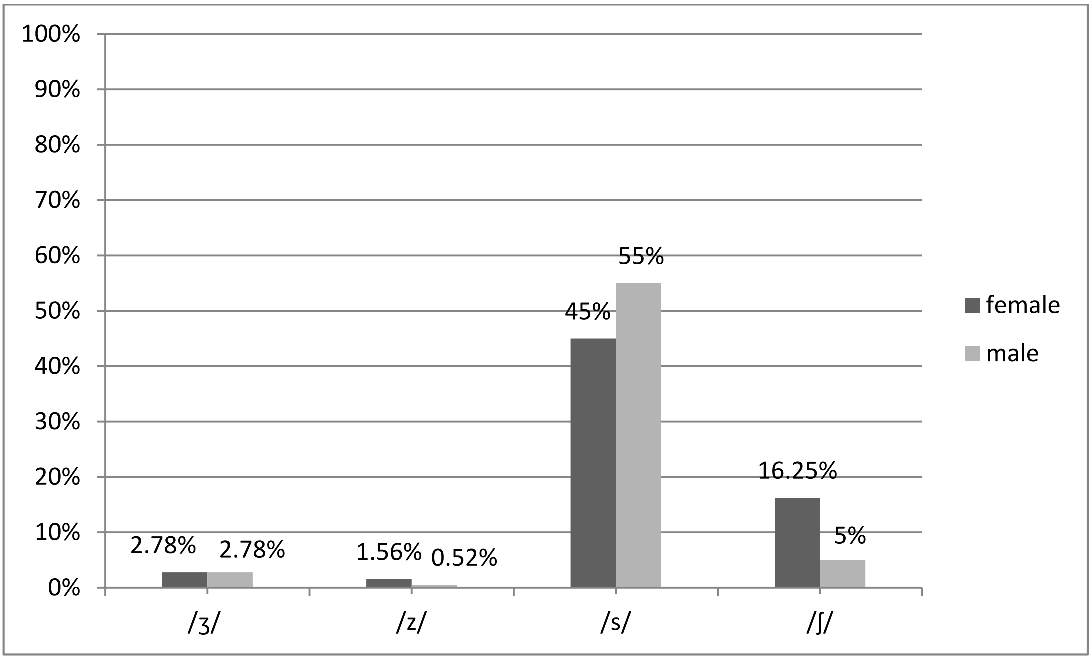


Figure 4.16 Mean Perception Test Percentages of Incorrect Answers in Trial 2

In the second trial of the perception test (Figure 4.16), similar patterns to the first trial's were observed; that is, the tokens that the participants had the least difficulty with were the ones containing /z/ and /ʒ/ with 2.78% incorrect answers for both; that is, they were able to pick the odd items about 97.22% of the time. The percentages of the wrong answers for tokens containing /ʃ/ for male and female participants were the same as the ones in trial one. And finally, the percentages of wrong answers for tokens containing the fricative /s/ for male and female too were similar to the ones in trial one with male speakers picking the odd items more often than did the female participants.

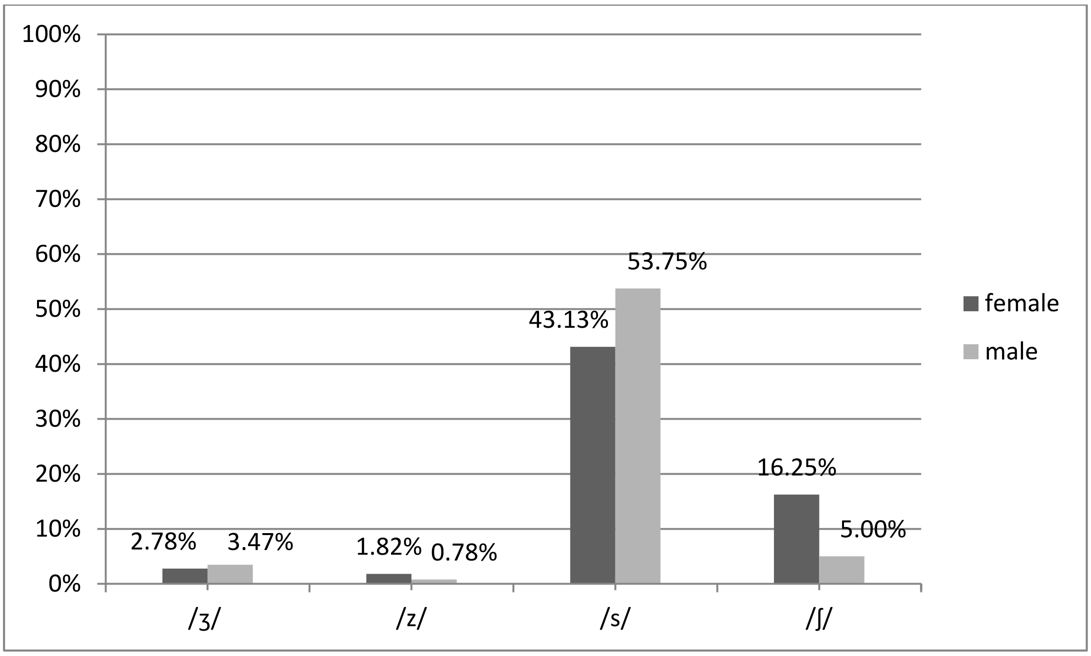


Figure 4.17 Overall Mean Perception Test Percentages of Incorrect Answers

Figure 4.17 shows mean perception test percentages of incorrect answers for the four fricatives in trials 1 and 2 for male and female participants. Overall, the percentages of incorrect answers for /z/ and /ʒ/ were not very different from the ones for trial one and trial two separately. More tokens containing /s/ were correctly picked (as being odd items) by the females than were by males; however, for /ʃ/, it was the males who more often picked the right tokens containing /ʃ/.

A mixed design ANOVA was carried out here too with the within-subjects factor being the number of incorrect answers and the between-subjects factors being Fricative and Gender. The effect of Gender was not significant here, $F(1, 840) = .066$, $p > .05$, but the Fricative effect, just like in the production test (native speakers' perception of the participants' production), was highly significant: $F(3, 840) = 206.677$, $p < .05$; the interaction of the two was significant as well, $F(3, 840) = 6.933$, $p < .05$. So even though the effect of gender was separately significant in the perception of /s/ and /ʃ/ (see sections 4.3.1 & 4.3.2), the overall effect of gender across all the fricatives and trials was not significant. Furthermore, a Tukey post-hoc test done on the fricatives revealed that the fricatives /ʃ/ and /s/, /ʃ/ and /z/, /ʃ/ and /ʒ/, /s/ and /z/ and finally /s/ and /ʒ/ significantly differed from each, $p < .05$.

What is striking when comparing the participants' production and perception results is that while the fricatives /ʒ/ and /z/ were rated by the native speakers as much less native-like in the production task (as compared with /s/ and /ʃ/), the participants correctly identified the odd items in most of the tokens containing /ʒ/ and /z/, in the perception test. This can perhaps be explained through Best's Perceptual Assimilation Model (PAM), as mentioned in Chapter 2, section 2.1.2.1; Best's perceptual assimilation pattern number five, uncategorized versus categorized assimilation, states that when an L2 learner is faced with a situation in which there are two L2 phonemes, one assimilating into an L1 sound and the other not, discrimination is normally very good. The participants in this study were facing a similar situation when it came to /ʒ/ and /z/ in the perception test. That is, there were three pairs of consonants involving each of these two English sibilants and all the pairs contained either /z/ or /ʒ/ plus other consonants that the participants made in the production test in place of the two sibilants; these other consonants were all Mandarin sounds except one, /ð/. So almost all the pairs had a consonant that matched a consonant in Mandarin and one that was uncategorized; hence the good perception of /z/ and /ʒ/. The fact that the participants produced different consonant sounds in the place of each of these two English sibilants in the production task (according to the author's phonetic transcription of the productions) further supports that they were not sure which phoneme they could perceptually categorize them under (see Appendices D & E).

In addition, Jing and Yanyan (2011), as was mentioned in Chapter 2, section 2.4, reached a somewhat similar conclusion. They had a group of Mandarin L1 students and employed a reading task and a listening discrimination test with seven pairs of English fricatives placed in minimal pairs: /f/ and /v/, /w/ and /v/, /s/ and /z/, /θ/ and /s/, /ð/ and /z/, /ʃ/ and /ʒ/, and /θ/ and /ð/. They concluded that the participants could easily discriminate, among others, the pairs /s/ /z/ and /ʃ/ /ʒ/ while in the reading task, none of them could produce /ʃ/ and /ʒ/ distinctively from each other.

With /s/, the situation was the opposite of that of /z/ and /ʒ/; the participants did not pick the odd items in the perception test about half the times (the highest number of wrong answers among the four fricatives) while their production of the same sound were rated by native speakers to be close to native-like (closer than the other three fricatives). This may be explained considering that the difference between English alveolar /s/ and Mandarin dental /s/ is phonetic in nature and, as was mentioned in Chapter 2, section 2.1.2.3 (Flege's Speech Learning Model), late L2 learners might, through Ll phonology, filter out L2 sound features that are phonetically important but not phonologically so. It can, therefore, be concluded that the two options of alveolar and dental /s/ were both almost equally good examples of Chinese dental /s/. According to PAM, when this happens, discrimination is usually not good (section 2.1.2.1). This may explain the inability of the speakers to correctly discriminate the two types of /s/ (dental versus alveolar) in about half the instances.

As for /ʃ/, the participants could perceive most of the tokens containing this fricative correctly. This good discrimination between /ʃ/ and /ʂ/ indicates that, according to PAM's category-goodness difference assimilation pattern (as mentioned in chapter two section, 2.1.2.1), the two phonemes are assimilated into a single L1 category but one of them (in this case /ʂ/) is perceived closer to the L1 sound than the other and discrimination is normally moderate to very good. The evidence for this comes partly from the production test where most of the non-native productions of this fricative by the Chinese speakers involved /ʂ/ (according to the author's phonetic transcription of the productions). It is also worth noting that the participants' productions of /ʃ/ were perceived by the American English speakers as close to native-like in the production task (as seen earlier in this chapter, section 4.1.5) which means their production as well as their perception of this sibilant sound were very good (close to the native level). We can conclude then, that the closest Mandarin sound to English /ʃ/ is most probably Mandarin /ʂ/. Further support for this comes from Chang et al. (2009) who also found that some of the native speakers of Mandarin from China in their study could not distinguish English /ʃ/ from Mandarin /ʂ/ (Chapter 2, section 2.3).

4.3.6 Comparison Based on Fricative Position

The incorrect responses that the participants gave in the perception of the target fricatives in each position were compared for the three fricatives of /s/, /z/, and /ʃ/. As was discussed earlier, /z/ occurred in eight of the 31 target words in four different positions of initial (*i*), medial (*m*), final (*f*), and final consonant cluster (*fcc*). The participants had no incorrect answers in the perception of medial and final consonant cluster /z/, but they picked the highest number of wrong items (wrong option) in initial /z/ in both trials (Figure 4.18). However, it was concluded, through a one-way ANOVA, that the number of wrong items picked by the listeners in different positions were not significantly different from each other, *F*(3, 124) = 1, *p* > .05.

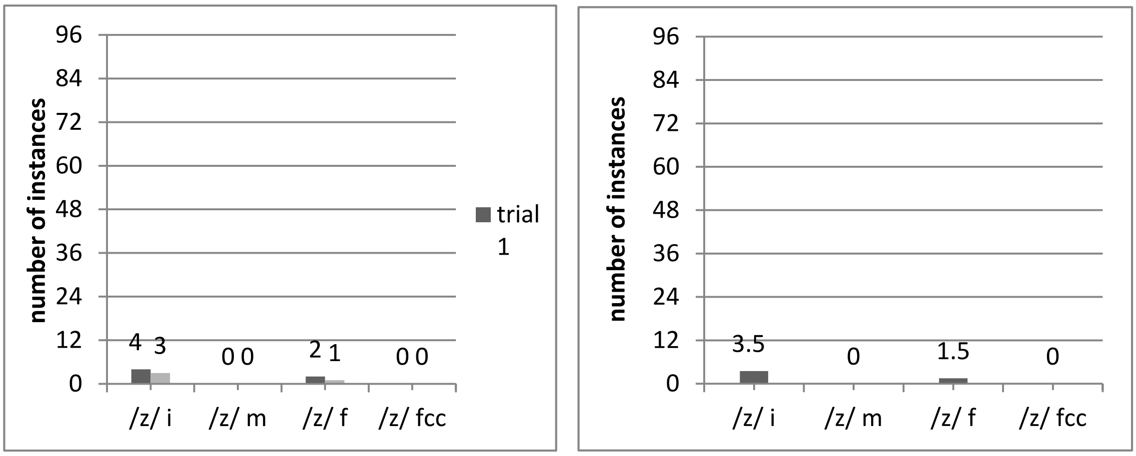


Figure 4.18 Comparison of Incorrect Answers in Perception of /z/

The sound /s/ occurred in 10 of the 31 target words in five positions. Here the participants had the least incorrect responses perceiving medial /s/ in both trials (Figure 19). In the first trial, they had the most incorrect answers in initial, initial consonant cluster, and final /s/. In the second trial, they picked the highest number of wrong items in initial /s/. So overall, they had the highest number of wrong answers in initial /s/; however, just like the results for /z/, the results of a one-way ANOVA here indicates that the number of wrong items picked in different positions for /s/ was not significantly different from each other, *F*(4, 155) = 1, *p* > .05.

The sound /ʃ/ occurred in ten of the 31 target words in five different positions. In trial one, as can be seen in Figure 4.20, the most incorrect responses were given when perceiving initial consonant cluster and final consonant cluster /ʃ/, and the least number of incorrect responses were made in the case of medial and final /ʃ/. In trial two, the participants had the least and the most number of wrong answers in final and final consonant cluster /ʃ/ respectively, but here too, a one-way ANOVA indicates that the number of wrong items picked in each position was not significantly different from each other, $F(4, 155) = 1$, $p > .05$.

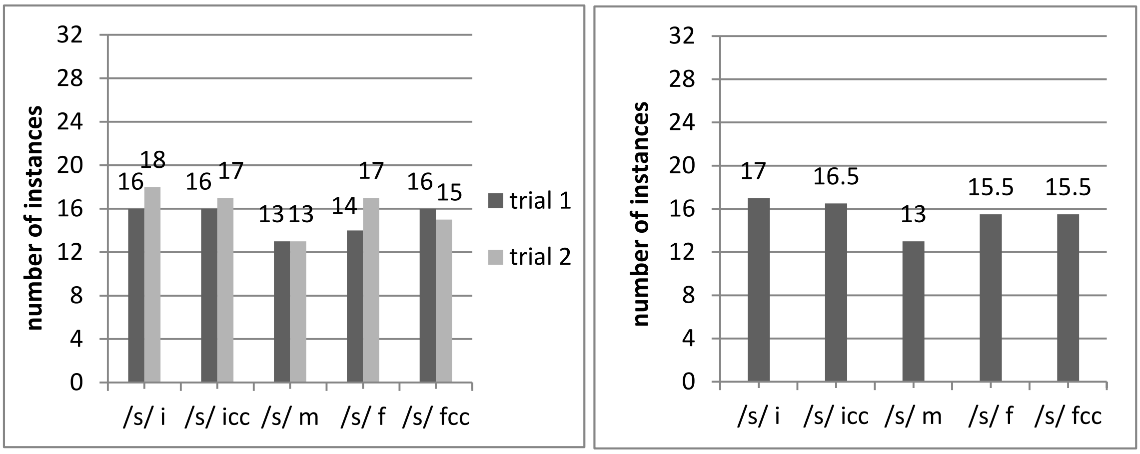


Figure 4.19 Comparison of Incorrect Answers in Perception of /s/

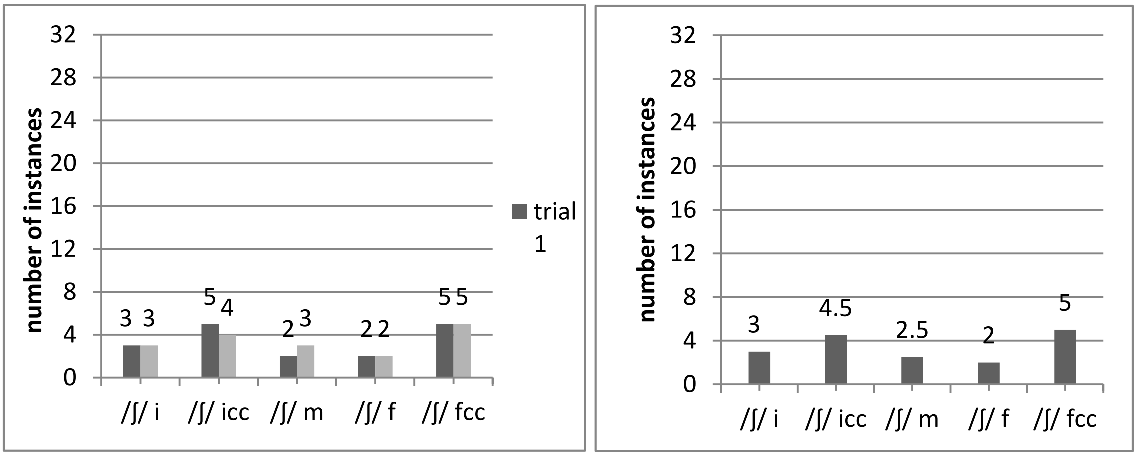


Figure 4.20 Comparison of Incorrect Answers in Perception of /ʃ/

4.4 Comparison of the Participants' Production and Perception

Table 4.3 shows mean percentages of the degree of non-nativeness for each of the 31 English words in the production task and the perception test. As was seen earlier, the production of /s/ was rated as close to native-like while the perception of this fricative was incorrect about half the time. For production involving /z/ and /ʒ/, the opposite holds true; that is, without exception, the production of every word containing /z/ or /ʒ/ was rated as not close to native-like while the perception of the same sounds was correct in many cases. This, however, was not exactly the case for /ʃ/. Some words containing

/ʃ/ were perceived better while others were produced better (according to the raters) by the participants, but the overall difference in production and perception (as was shown under production task and perception test earlier in this chapter) was not as big as in the other three fricatives. This is further demonstrated in Table 4.3.

Table 4.3 Mean Production and Perception Percentages of Degree of Non-Nativeness for the 31 Words

Word	Production	Perception		Word	Production	Perception
zip	53.91	4.17		wasp	14.69	53.13
kiss	1.88	46.88		shark	2.97	15.63
closed	30.78	0		fashion	4.06	15.63
washed	24.53	18.75		razor	55.94	0
push	3.75	3.13		ship	3.13	3.13
fast	3.59	59.38		step	2.81	59.38
cause	82.34	4.17		maze	46.88	0
shrink	9.84	18.75		fish	1.72	9.38
missing	6.56	50		crazed	29.69	0
usually	85.47	4.17		sad	9.22	53.13
miss	2.81	50		cashed	4.22	12.5
bishop	8.75	3.13		busy	26.88	0
spark	5.78	62.5		sick	3.28	65.63
zeal	60.31	3.13		shrimp	14.06	12.5
facing	4.69	50		pleasure	60.63	4.17
television	38.75	2.08				

As was mentioned in chapter two, section 2.1.1, Flege (1995) claimed that the accuracy of L2 learners' productions is constrained by how accurately they perceive L2 sounds. He further argued that perceptual representation of L2 phonetic segments correspond with their productions which means L2 production is usually not more native-like than its perception and can even be less native-like. This can potentially explain why the perception of /z/ or /ʒ/ in this study is better than their production and also why there is not a big gap between production and perception of /ʃ/. The only exception here is /s/; however, as was mentioned in this chapter, section 4.1.5, it is entirely possible that the native speaker raters have rated many dental /s/ productions as native-like. As Jing and Yanyan (2011) also pointed out in their study (Chapter two, section 2.4), dental and alveolar /s/ are allophones of the same sound in both English and Mandarin Chinese and therefore replacing them with one another will probably not cause any difficulty for the native speakers of either languages. The fact that many of the listeners in this study could not discriminate dental and alveolar /s/ is further was confirmed by a second perception test in the next section.

4.5 Second Perception Test

Looking at the results, it was obvious that most of the participants could correctly perceive the three fricatives /z/, /ʒ/, and /ʃ/ in most of the words; however, this was not the case for /s/. It was, therefore, assumed that most of the participants perceived the tokens containing dental and alveolar /s/ as the same. Therefore, a follow-up perception test was done in which the participants were asked, a month later, to listen to the same 10 words containing /s/ which had been produced by the author for the first perception test in two trials. Therefore, there were 320 instances of the perception test here: 10 (words) × 16 (participants) × 2 (trials) = 320. This time, however, in addition to being able to pick one of the words as the odd item, the participants had the additional fourth option of “the same”. The participants were told that the tokens in the second test were not the same as the ones they listened to in the first perception test so as not to make them think that there must have necessarily been an odd item out while in fact there *was* an odd item out in each trial. Figure 4.21 shows the second perception test means (%) for the 10 words containing /s/ in which the participants, in each trial of every word, chose an odd item out (correct choice), the wrong item out, or perceived them as the same. The participants were tested twice. 70% of the time the participants perceived the tokens in each trial as the same even though there was indeed an odd item out.

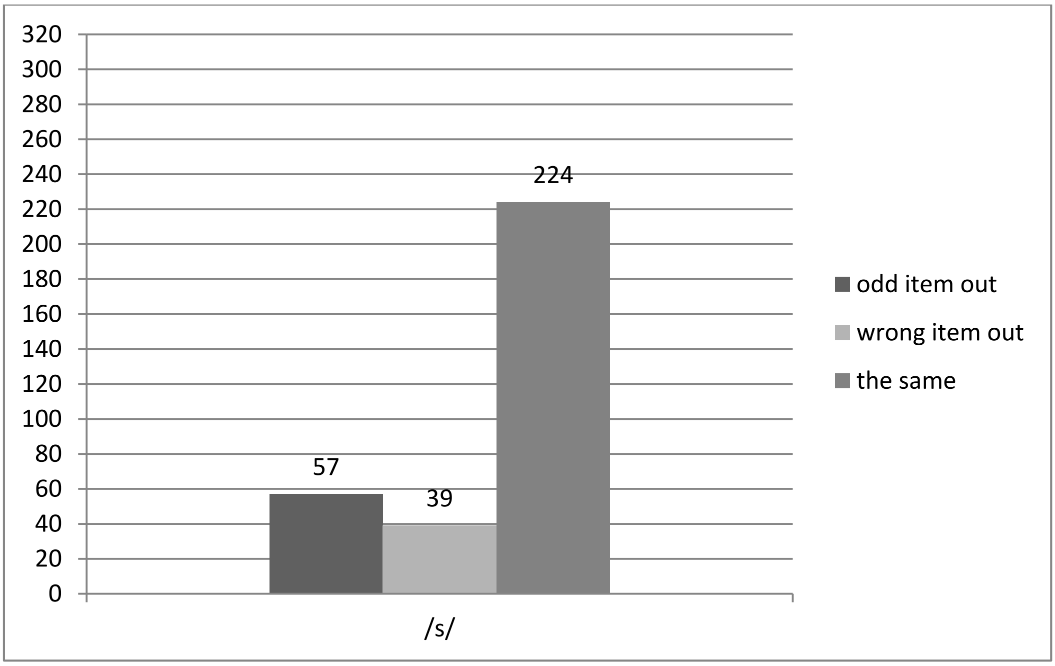


Figure 4.21 Second Perception Test Results for the 10 Words Containing /s/

4.6 Summary

The results of the production task involving native Mandarin speaker participants indicate that their production of /s/ and /ʃ/ was much more native like than their production of /ʒ/ and /z/. These were the first and second most difficult English sibilants for the participants to produce respectively. The difference between English and Mandarin /s/ are phonetic and as Chang et al. (2009) pointed out, the two sibilants can “undergo equivalence classification” and “become indistinguishable from each other” (Chang et al., 2009, p. 38). The same thing can be said about English /ʃ/ and Mandarin /ʂ/ as they are categorized as ‘similar’ according to Flege (1987). About /ʒ/ and /z/, Deterding (2006) reached a similar conclusion indicating that these two sibilants cause the most difficulty in production for Native speakers of Mandarin from China. In the perception test, however, most of the participants could perceive /ʒ/ and /z/ correctly. Jing and Yanyan (2011) also reached a somewhat similar conclusion (Chapter two, section 2.4). Finally, a second perception test confirmed the fact that many of the participants could not distinguish Chinese /s/ from English /s/ and perceived them as the same.

CHAPTER 5

CONCLUSION

5.1 Summary

In this research, an attempt was made to study native Mandarin speakers' production and perception of English sibilant fricatives. In doing so, the overriding goal was to see how close to General American English the participants produced the sibilants, and to what extent they could perceive the same sounds and also to see if there were any differences between their production and perception of these sounds in terms of the degree of nativeness. Five native American English speakers rated the productions in relation to the degree of 'nativeness' of the pronunciation. In the perception part, the same participants listened to and selected the odd item in each instance of the perception test. Now to summarize the main findings, it is appropriate to look at the research questions once again and attempt to provide a summary through answering these questions.

5.1.1 Research question 1: How close to General American English sounds do the participants produce English sibilant fricatives?

The results of the production task show that /s/ had the lowest degree of non-nativeness among the four sibilants with a mean rate of 1.21 on the 5-level scale (as perceived by the raters) across both trials and genders. This means that the Chinese speakers' productions of this sibilant were very close to native-like. However, it is believed by the author that the Chinese speakers had dental /s/ in their productions, but the native speaker raters did not generally discriminate dental /s/ from alveolar /s/. It may be possible that they were perceived as the allophones of the same sound. The reason why it is believed the Chinese speakers had dental /s/ in their productions is the fact that their production of this sound was much more native-like than their perception because as was mentioned earlier in Chapter 2, section 2.1.1, Flege argued that L2 production cannot normally be more native-like than its perception which is the case for the three other English sibilants here (this will be further discussed in the answer to the third research questions). The results of the second perception test (Chapter 2, section 4.5) further demonstrates the participants' lack of discrimination between dental and alveolar /s/. /ʃ/ was perceived by the raters as being close to native-speaker pronunciation with a mean rate of 1.31. While /s/ and /ʃ/ were perceived to be very close to native-like pronunciation, /z/ and /ʒ/ were rated as rather far from it. /z/ had a rather high degree of non-nativeness, 2.93, as perceived by the judges and was the second most problematic English sibilant here. /ʒ/ had the highest degree of non-nativeness associated with it at 3.46 on the scale (as perceived by the raters). The result of the author's transcription of the sibilants in the speakers' productions is consistent with the judges' ratings as well. (this will also be further elaborated on in the answer to the third research questions).

5.1.2 Research question 2: How do the participants' perceive English sibilant fricatives?

In the perception test, /z/ had the lowest percentage of wrong answers (1.3%); this means that the participants could correctly identify the odd items in most of the tokens containing this sibilant. The sibilant with the second lowest percentage of wrong answers was /ʒ/ with only 3.47% of the answers containing this fricative being perceived incorrectly by the participants. The results of the perception test for /z/ and /ʒ/ is in accordance with Best's Perceptual Assimilation Model, uncategorized versus categorized pattern. In the perception test involving these two sibilants, mostly native Mandarin phonemes were paired with non-native categories (/z/ and /ʒ/) in the instances of the AXB categorial discrimination test. The Chinese listeners, therefore, could easily recognize one of the sounds in each pair as a native phoneme but the other sound could not fit into a native category. According to PAM discrimination is usually very good under such situations (as was explained in Chapter 2, section 2.1.2.1). The participants did not have a lot of difficulty perceiving /ʃ/ either. The percentage of wrong answers for this fricative was 10.62 which is higher than the percentage for /s/ but still relatively low (compared with the ones for /z/ and /ʒ/). The perception of this sibilant followed PAM's category-goodness difference pattern (also mentioned in Chapter two section, 2.1.2.1) in which the two categories involving this sibilant were both assimilated into a single L1 category but one of them, /ʂ/, was perceived closer to the native sound than the other. According to PAM, discrimination is normally moderate to very good in such situations. The sibilant /s/, however, was incorrectly perceived about half the times (48.44%). The discrimination between dental /s/ versus alveolar /s/ by the Chinese listeners followed a single-category pattern in PAM which means both choices could almost equally be identified as the native category. Discrimination in these situations is normally not very good (not close to the native level). As a result, it can be concluded that the listeners' choices involving this sibilant were mostly due to chance. In conclusion, while the participants could correctly perceive most of the tokens containing /z/, /ʒ/, and /ʃ/, they could not do so when it came to /s/.

5.1.3 Research question 3: To what extent is there a difference between the students' perception and production of English sibilant fricatives?

When it came to the comparison of the results of the production task and the perception test, some striking differences were observed. The production of /s/, for example, was rated as close to native-like while it was incorrectly perceived about half the time. As was explained in Chapter 4, both dental and alveolar /s/ were perceived as belonging to the same native category by the majority of Chinese listeners. A second perception test involving these two phonemes confirmed this. The results of the study also lead us to believe that the native speaker raters too had a difficult time discriminating the two types of /s/ in the Chinese participants' productions especially considering the huge gap between the participants' productions and perceptions and bearing in mind that, according to Flege (Chapter 2, section 2.1.1), non-native segmental production cannot normally be better than its perception.

For the sibilants /z/ and /ʒ/, however, the opposite was true where most of the productions of these sibilants in the token words were not rated by the judges as native-like or even very similar while the same target sounds were perceived correctly by the participants in most instances of the perception test. These two sibilants do not exist in Mandarin Chinese and so they generally followed an uncategorized versus categorized pattern in the perception test (according to PAM) which means their perception can indeed be very good but not necessarily their production. This is supported by the fact that those Chinese speakers who did not correctly produce /z/ and /ʒ/ in the production task were not consistent in their choice of phoneme in the place of each of these two English sibilants. This can be an indication that they were not sure which phoneme they could perceptually categorize /z/ and /ʒ/ under.

Finally /ʃ/ was correctly produced (according to the raters) and perceived by the participants in many instances. So /ʃ/ was the only sibilant here whose production and perception by the participants corresponded in terms of the degree of nativeness. As was explained in the answer to the second research question, the perception of this sibilant matched PAM's category-goodness difference pattern. This means the two categories involving this sibilant in the perception test (/ʃ/-/ʂ/) were both assimilated into a single L1 sound, but one of them, /ʂ/, was perceived closer to the native sound than the other.

According to this pattern perception is moderate to very good. Indeed, the closest Mandarin sound to English /ʃ/ is /ʂ/. The incorrect productions of /ʃ/ by the Chinese participants mostly involved /ʂ/ (unlike /z/ and /ʒ/ for each of which several different incorrect sounds were produced). This implies that they either perceived /ʃ/ correctly or categorize it under /ʂ/ (with just a few exceptions as shown in appendix D). In conclusion, we can say that the participants could perceive at least two of the sibilants, /z/ and /ʒ/, more accurately than they could produce them and for /ʃ/ this seemed to be about the same. For /s/ we can tentatively say that, based on the results, the productions were, in fact, not any better than the perception (if not worse).

5.2 Implications

The present study provides a description of the production and perception of English sibilant fricatives by people from China as well as a comparison of the results of the two. It is based on a production task rated by native speakers of American English and a perception test evaluated by the author indicating how close to the target sounds did the participants produce and perceive the sibilants. The findings of this study will contribute to the scarce body of knowledge on how English language learners from China produce and perceive English sounds; this, in turn, highlights the problematic areas that need to be improved when it comes to the teaching and learning of English pronunciation in China. Some of these areas include retroflexisation and africatisation of English sibilants as well as devoicing of voiced English sibilants. This study also demonstrates, though on a small scale, that not all the non-native productions can be problematic in terms of ineligibility or even contribute to a very noticeable foreign accent. The implication of this would be the identification of and paying more attention on the part of EFL course designers and teachers to those sounds that can cause unintelligibility and breakdowns in conversation.

5.3 Recommendations for Further Research

The present research has resulted in some useful results and conclusions on the production and perception English sibilant fricatives by people from China and at the same time it has revealed some areas/issues that need further research/improvement. One of the issues in this study has been a small sample size. Further research on the same area, therefore, should include a bigger sample size and ideally a wider range of participants (probably those who speak other Chinese dialect as their first language in China, for example Cantonese). Another area that needs improvement is the study of the sibilants in conversation rather than in a word list only in order to ensure that the data resemble the natural occurrence of these sounds. Another potential area of improvement in future studies of the subject is the instrumental analysis of the sibilants including inspecting the spectral shape of the productions. For example, as was discussed in Chapter 4, the participants as well as the five raters did not seem to be able to differentiate dental versus alveolar /s/ from each other in the majority of cases; therefore more sophisticated acoustic analysis measurements, such as the ones used by Li et al. (2007), are needed. Finally, the production and perception of English fricatives as a whole, as opposed to only sibilants, can be investigated in subsequent research.

5.4 Concluding Remarks

This study depicted certain characteristics and trends in the production and perception of English sibilant fricatives by a group of Chinese students through the ratings of five native speaker judges and also by employing a perception test. As well as facilitating the teaching and learning of English pronunciation in China, the findings of this study can potentially encourage subsequent research on similar areas of difficulty for such learners too.

References

Abramson, A., & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. In B. Hála, M. Romportl & P. Janota (Eds.), *Proceedings of the Sixth International Congress of Phonetic Sciences* (pp. 569–573). Prague, Czech Republic: Academia.

Adamson, B. (2001). English with Chinese characteristics: China's new curriculum. *Asia Pacific Journal of Education*, *21*, 19-33.

Altmann, H., & Kabak, B. (2011). Second Language Phonology. In N. C. Kula, B. Botma & K. Nasukawa (Eds.), *Continuum Companion to Phonology* (pp. 298-320). New York, NY: Continuum International Publishing Groups.

Aslin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. H. Yeni-Komshian, J. F. Kavanagh & C. A. Ferguson (Eds.), *Child Phonology: Perception* (pp. 67-96). New York, NY: Academic.

Baker, W., Trofimovich, P., Mack, M., & Flege. J. E. (2002). The effect of perceived phonetic similarity on non-native sound learning by children and adults. In: B. Skarabela, S. Fish & A. Do (Eds.), *Proceedings of the 26th annual Boston University Conference on Language Development* (36-47). Somerville, MA: Cascadilla Press.

Best, C. T. (1994a). Learning to perceive the sound pattern of English. In C. Rovee-Collier & L. P. Lip-sitt (Eds.), *Advances in Infancy Research*, *vol. 9* (pp. 217-304). Norwood, NJ: Ablex Publishing.

Best C. T. (1994b). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In H. C. Nusbaum & J. Goodman (Eds.), *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words* (pp. 167-224). Cambridge, England: MIT Press.

Best C. T. (1995). A direct realist perspective on cross-language speech perception. In: W. Strange (Ed.), Speech *Perception and Linguistic Experience: Theoretical and Methodological Issues in Cross-language Speech Research* (pp. 171-206). Timonium, MD: York Press.

Best, C. T., McRoberts, G. W. & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345–360.

Best, C. T. & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, *20*, 305-330.

Best, C. T., McRoberts, G. & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America, 109*, 775-794.

Bever, T. (1981). Normal acquisition processes explain the critical period for language learning. In K. Diller (Ed.), *Individual Differences and Universal in Language Learning Aptitude* (pp. 176-198). Rowley, England: Newbury House.

Broselow, E. (1984). An investigation of transfer in second language phonology. *International Review of Applied Linguistics, 22*, 253-269.

Chang, C. B., Haynes, E. F., Yao, Y. & Rhodes, R. (2009). A Tale of Five Fricatives: Consonantal Contrast in Heritage Speakers of Mandarin. *University of Pennsylvania Working Papers in Linguistics, 15*, 37-43.

Churchland, P. (1986). *Neurophilosophy*. Cambridge, England: MIT Press.

Da´vila, A., Bohara, A. & Saenz, R. (1993). Accent penalties and the earnings of Mexican Americans. *Social Science Quarterly, 74*, 902–915.

DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499-534.

Deterding, D. (2006). The pronunciation of English by speakers from China. *English World-Wide, 27*, 175-198.

Edelman, G. (1989). *The Remembered Present: A Biological Theory of Consciousness*. New York, NY: Basic Books.

Eimas, P. D. (1991). Comments: Some effects of language acquisition on speech perception. In I. G. Mattingly & M. Studdert-Kennedy (Eds.), *Modularity and the Motor Theory of Speech Perception*. Hillsdale, NJ: Erlbaum.

Erdmann, P. H. (1973). Patterns of stress-transfer in English and German. *International Review of Applied Linguistics in Language Teaching, 11*, 229–242.

Finegan, E. (2012). *Language: Its Structure and Use*. Boston, MA: Cengage Learning.

Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America, 76*, 692-707.

Flege, J. E. (1986). The production and perception of foreign language speech sounds. In H. Winitz (Ed.), *Human Communication and its Disorders, Vol. II*, (pp. 224-401). Norwood, NJ: Ablex Publishing.

Flege, J. E. (1987). The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics*, *15*, 47-65.

Flege J. E. (1990). Perception and production: The relevance of phonetic input to L2 language learning. In: C. Ferguson & T. Heubner (Eds.), *Crosscurrents in Second Language Acquisition and Linguistic Theories* (pp. 249-289). Philadelphia, PA: John Benjamins.

Flege, J. E. (1991). Age of learning affects the authenticity of voice onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America, 89*, 395-411.

Flege, J. E. (1995). Second-language speech learning: Theory, findings, and problems. In: W. Strange (Ed.), *Speech Perception and Linguistic Experience: Issues in Cross-Language Research* (pp. 229-273). Timonium, MD: York Press.

Flege, J. E. & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, *15*, 67-83.

Flege, J. E. & Liu, S. (2001). The effect of experience on adults. acquisition of a second language. *Studies in Second Language Acquisition, 23*, 527-552.

Flege, J. E. & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language and Speech*, *24,* 125-146.

Francis, A., & Nusbaum, H. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 349-366.

Gillette, S. (1980). Contextual variation in the perception of /r/ and /l/ by Japanese and Korean speakers. *Minnesota Papers in Linguistics and the Philosophy of Language, 6*, 59–72.

Gottfried, T. & Beddor, P. (1988). Perception of temporal and spectral information in French vowels. *Language and Speech*, *31*, 57-75.

He, D. & Li, D. C. S. (2009). Language attitudes and linguistic features in the China English debate. *World Englishes 28*, 70-89.

Hoke, Z. (2013, October, 2013). American English becoming more popular in former British colony. *Voice of America*. Retrived from http://www.voanews.com/content/american-english-becoming-more-popular-former-british-colony-hongkong/1779711.- html

Hu, G. W. (2002). English language teaching in the People's Republic of China. In R. E. Silver, G. W. Hu & M. Iino (Eds.), *English language education in China, Japan, and Singapore* (PP. 1-77). Singapore: National Institute of Education.

Hutchison, P. (2013, July 2). Hong Kong's boom in American accent schools. *Agence France-Presse*. Retrieved from http://www.google.com/hostednews/afp/article/-ALeqM5ibdBNVFC9tglK9HrxuK2rIog4GUA?docId=8aeb51ad7bff43f2b5ee74ef2e236def

Ingram, J. C. & Park, S. G. (1998). Language, context, and speaker effects in the identification and discrimination of English /ɹ/ and /l/ by Japanese and Korean listeners. *Journal of the Acoustical Society of America*, *103*, 1161–1174.

Iverson, P. & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of Acoustic Society of America, 99,* 1130–1140

Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.

Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an international language. *Applied Linguistics, 23*, 83-103.

Jia, C. (2013, July 24). US still top choice for students. *China Daily*. Retrieved from http://usa.chinadaily.com.cn/epaper/2013-07/24/content_16822768.htm

Jing, X. & Yanyan, Z. (2011, August). *A study of Chinese EFL learners' acquisition of English fricatives*. Paper presented at the 16th Conference of the Pan-Pacific Association of Applied Linguistics, Hong Kong. Retrieved from http://paaljapan.org/conference2011/ProcNewest2011/pdf/oral/2B-3.pdf

Johnson, C. (2000). Children's phoneme identification in reverberation and noise. *Journal of Speech, Language, and Hearing Research*, *43*, 129- 143.

Johnson, K. & Babel, M. (2010). On the perceptual basis of distinctive features: Evidence from the perception of fricatives by Dutch and English speakers. *Journal of Phonetics, 38*, 127–136.

Klein, W. (1986). *Second Language Acquisition*. Cambridge, England: Cambridge University Press.

Kuhl, P. K. (1991). Human adults and human infants show a 'perceptual magnet effect' for the prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50*, 93–107.

Kuhl, P. K. (1992). Speech prototypes: Studies on the nature, function, ontogeny and phylogeny of the "centers" of speech categories. In: Y. Tohkura, E. Vatikiotis-Bateson & Y. Sagisaka (Eds.), *Speech Perception, Production and Linguistic Structure* (pp. 239–264). Tokyo, Japan: Ohmsha.

Kuhl, P. K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Science, 97*, 11850-11857.

Kuhl, P. K. & Iverson, P. (1996). Linguistic experience and the "perceptual magnet effect." In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 121-154). Baltimore, MD: York Press.

Kuhl, P. K. & Meltzoff, A. (1996). Infant vocalizations in response to speech: Vocal imitation and developmental change. *Journal of the Acoustical Society of America, 100,* 2425-2438.

Kuhl, P. K., Williams, K.A., Lacerda, F., Stevens, K.N. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606–608.

Ladefoged, P. & Maddieson, I. (1998). *The Sounds of the World's Languages*. Malden, MA: Blackwell Publishers.

Ladefoged, P. & Wu, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricatives. *Journal of Phonetics, 12*, 267-278.

Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers.* Ann Arbor, MI: University of Michigan Press.

Lam, C. H. (2007). *Englishes in Hong Kong: Students' awareness of, attitudes towards accent differences and the intelligibility of accents of English*. (Unpublished master's thesis). University of Hong Kong, Hong Kong.

Leather, J., & James, A. (1996). The acquisition of second language speech. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 269-316). San Diego, CAL: Academic Press.

Lehn, W. & Slager, W. R. (1959). A contrastive study of Egyptian Arabic and American English: The segmental phonemes. *Language Learning, 9*, 25–33.

Li, F. (2008). *The phonetic development of voiceless sibilant fricatives in English, Japanese and Mandarin Chinese. (Unpublished PhD thesis).* The Ohio State University, Ohio.

Li, F., Edwards, J. & Beckman, M. E. (2007, August). *Spectral measures for sibilant fricatives of English, Japanese, and Mandarin Chinese*. Paper presented at the the 16th International Congress of Phonetic Sciences, Saarbruecken, Germany. Retrieved from http://www.icphs2007.de/conference/Papers/1446/1446.pdf

Lively, S. E., Logan, J. S. & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. *Journal of the Acoustical Society of America, 94*, 1242–1255.

Lodge, K. (2009). *A Critical Introduction to Phonetics*. London, England: Continuum International Publishing Group.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251-285.

MacKain, K. S., Best, C. T. & Strange, W. (1981). Categorical perception of English /ɹ/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, *2*, 369–390.

McAllister, R., Flege, J. E. & Piske, T. (2002, August). *The influence of the L1 on the acquisition of Swedish vowel quantity by native speakers of Spanish, English and Estonian*. Paper presented at the 14th International Congress of Phonetic Sciences, San Francisco. Retrieved from http://www2.ling.su.se/fon/perilus/1999_16.pdf

McGuire, G. (2007). English listeners' perception of Polish alveopalatal and retroflex voiceless sibilants: A pilot study. *UC Berkeley Phonology Lab Annual Report*. Retrieved from http://linguistics.berkeley.edu/phonlab/annual_report/documents/2007/McGuire.pdf

McLaughlin, B. (1977). Second-language learning in children. *Psychological Bulletin, 84*, 438-459.

Meyer, C. F. (2009). *Introducing English Linguistics*. Cambridge, England: Cambridge University Press.

Michaels, D. (1974). Some replacements and phonological systems. *Linguistics, 126*, 69-81.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J. J. & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception and Psychophysics, 18*, 331–340.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, *36*, 39-66.

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly, 37*, 589-613.

Patkowski, M. (1989). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics, 11*, 73-89.

Pisoni, D. (1995). Some thoughts on "normalization" in speech perception. Research on Spoken Language Processing, Progress Report. *Department of Psychology, Indiana University, 20*, 3-30.

Plichta, B. (2010). Using proper microphone placement. Retrieved from http://bartus.org/akustyk/mic_placement.php, 5 March 2010.

Polivanov, E. (1931). La perception des sons d'une langue étrangère. *Travaux du Cercle Linguistique de Prague*, *4*, 79–96.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America, 89,* 2961–2977.

Rajadurai, J. (2007). Intelligibility studies: a consideration of empirical and ideological issues. *World Englishes*, *26*, 87-98.

Rochet, B. (1995). Perception and production of L2 speech sounds by adults. In: W. Strange (Ed.), *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues* (pp. 379-410). Timonium, MD: York Press.

Scovel, T. (1988). *A Time to Speak. A Psycholinguistic Inquiry into the Critical Period for Human Speech*. Cambridge, England: Newbury House.

Sebastián-Gallés, N. & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, *72*, 111-123.

Sheldon, A. & Strange, W. (1982). The acquisition of /ɹ/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3,* 243–261.

Song, D. (2013, November 24). U.S. colleges draw record foreigners. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2013/nov/11/local/la-me-foreign-students-20131111

Strange, W. & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Percept Psychophys, 36*, 131–145.

Takagi, N. & Mann, V. (1995). The limits of extended naturalistic exposure on the perceptual mastery of English /ɹ/ and /l/ by adult Japanese learners of English. *Applied Psycholinguistics, 16*, 379–405.

Tottie, G. (2002). *An Introduction to American English*. Oxford, England: Blackwell Publishers.

Trubetzkoy, N. (1939/1969). *Grundzuge der phonology*. Travaux du Cercle Linguistique de Prague, 7 (English translation by C. Baltaxe). Berkeley, MO: University of California Press.

Trudgill, P. & Hannah, J. (2002). *International English: A guide to varieties of standard English*. London: Arnold; New York: Co-published in the United States of America by Oxford University Press.

Tsao, F.-M., Liu, H.-M. & Kuhl, P. K. (2006). Perception of native and non- native affricate-fricative contrasts: Cross-language tests on adults and infants. *The Journal of the Acoustical Society of America, 120*, 2285-2294.

Walley, A. & Flege J. E. (2000). Effects of lexical status on children's and adults' perception of native and non-native vowels. *Journal of Phonetics, 27,* 307-332.

Wang, W. F., & Lam, A. (2009). The English language curriculum for senior secondary school in China: Its evolution from 1949. *RELC Journal, 40(1)*, 65-82.

Wang, J. (2006). Difficulties and countermeasures in the implementation of quality oriented education in rural primary and secondary schools. *Jiaoyu Yanjiu (Educational Research), 11*, 41-46.

Ward, C., & Francis, E. (2010, November 15). Ganbei!! China Embraces English Language. *abcnews.com*. Retrieved from http://abcnews.go.com/WN/China/china-pushes-english language/story?id=12154435.

Wei, Z. & Ding, L. (2011, August). *A Preliminary study on Chinese EFL learners' attitude towards their accent.* Paper presented at the 17th International Congress of Phonetic Sciences, Hong Kong. Retrieved from http://www.icphs2011.hk/resources/OnlineProceedings/RegularSession/Zhang,%20Wei/Zhang,%20Wei.pdf

Werker, J. F., Gilbert, J. H., Humphrey, K. & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*, 349–355.

Werker, J. F. & Tees, R. C. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, *75*, 1866–1878.

Xu, S. & Connelly, M. F. (2009). Narrative inquiry for teacher education and development: Focus on English as a foreign language in China. *Teaching and Teacher Education*, *25,* 219-227.

Yamada, R. A. & Tohkura, Y. (1992). The effects of experimental variables on the perception of American English /ɹ/ and /l/ by Japanese listeners. *Perception & Psychophysics, 52*, 376–392.

Zou, W.C., & Zhang, S.L. (2011). Family background and English learning at compulsory stage in Shanghai. In A.W. Feng (Ed.), *English language education across Greater China* (pp. 189-211). New York, NY: Multilingual Matters.

Appendix A: List of words containing the four sibilants used in the production task

/s/	/ʃ/	/z/	/ʒ/
Kiss	washed	zip	Usually
Fast	push	closed	Television
Missing	shrink	cause	pleasure
Miss	bishop	zeal	
Spark	shark	razor	
Facing	fashion	maze	
Wasp	ship	crazed	
Step	fish	busy	
Sad	cashed		
Sick	shrimp		

Appendix B: Sample Questionnaire

Participant's number :
1. Full name:
2. Gender:
3. Age:
4. Which Grade are you in and which course are you studying?
5. Where are you from in China?
6. How long you have been living in Malaysia?
7. What Chinese dialect did you grow up speaking at home?
8. How long have you been studying English?

Appendix C: The ratings of the participants' productions by the five raters

Target Words	Trial 1, Rater 1															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	3	5	1	5	1	5	2	4	3	5	5	2	1	1
kiss	2	2	2	1	1	1	1	1	1	1	1	1	1	1	2	1
closed	1	3	1	5	1	2	5	3	2	2	4	2	1	1	1	5
washed	2	3	2	2	3	2	1	2	3	3	2	2	2	1	2	5
push	1	2	2	1	1	1	1	2	3	2	1	1	1	1	1	1
fast	1	1	2	2	1	1	1	1	1	1	2	1	1	1	1	1
cause	5	3	4	4	3	5	5	3	5	5	3	5	4	4	4	5
shrink	2	1	1	1	1	1	1	5	1	2	2	1	1	1	1	2
missing	2	2	1	2	1	1	1	2	1	1	2	2	2	1	1	1
usually	5	4	5	5	5	5	2	5	5	5	5	2	5	5	3	5
miss	2	1	1	1	1	1	1	1	2	1	1	1	1	1	1	2
bishop	2	1	2	2	1	1	1	2	1	2	2	1	2	1	1	1
spark	1	1	1	1	1	1	1	1	2	2	1	1	1	1	2	2
zeal	2	5	4	5	1	3	1	5	5	5	4	5	4	2	1	2
facing	2	1	2	1	1	1	1	2	1	1	2	1	2	1	1	1
television	2	2	2	5	2	2	2	5	2	2	5	2	2	5	1	2
wasp	1	1	2	1	1	2	5	2	5	1	1	1	1	2	1	1
shark	2	1	2	1	1	1	1	1	1	2	1	1	1	1	1	1
fashion	2	2	2	2	1	1	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	2	5	1	5	4	1	1	5	4	4	1	5
ship	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
step	1	1	1	1	1	1	1	1	2	2	2	1	1	2	1	2
maze	2	5	4	2	1	5	2	4	2	3	2	2	4	2	4	5
fish	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1
crazed	1	2	1	1	1	2	5	5	2	2	2	1	2	2	1	2
sad	2	2	1	1	1	1	2	1	2	2	1	1	1	1	2	2
cashed	2	1	2	2	1	1	2	1	1	1	1	1	1	2	1	1
busy	2	1	4	5	1	5	1	5	1	2	2	2	1	2	1	2
sick	1	1	1	1	1	1	1	2	1	2	1	2	1	2	1	1
shrimp	2	1	1	1	1	2	1	2	1	2	1	1	1	2	3	2
pleasure	1	3	5	5	2	5	2	5	5	5	5	3	5	5	2	5

Target Words	Trial 1, Rater 2															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	4	5	1	5	1	5	2	4	3	4	5	1	1	1
kiss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	1	3	1	3	1	1	5	2	1	1	4	1	1	1	1	4
washed	2	2	1	2	3	2	2	3	2	2	1	2	2	1	2	5
push	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1
fast	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
cause	5	4	4	4	2	5	5	2	4	4	3	4	4	3	4	4
shrink	2	1	1	2	1	1	1	5	1	1	2	1	1	1	2	1
missing	1	1	1	1	1	1	1	1	1	1	2	2	1	1	1	1
usually	5	3	5	5	5	5	1	5	5	4	5	2	5	5	3	5
miss	1	2	1	1	1	1	1	1	1	1	2	1	1	1	1	1
bishop	1	2	2	1	1	1	1	1	1	2	1	1	1	1	1	1
spark	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
zeal	2	5	4	5	2	2	1	5	5	4	3	5	4	2	1	1
facing	1	1	1	2	1	2	1	2	1	2	1	1	2	1	1	1
television	2	2	2	5	1	1	1	5	2	1	5	2	1	5	1	2
wasp	1	1	1	1	1	2	5	1	5	1	1	1	1	1	1	1
shark	2	2	1	1	1	2	2	1	1	2	1	1	1	1	2	1
fashion	1	1	2	1	2	1	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	5	3	1	1	5	5	3	1	5
ship	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
step	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
maze	2	5	3	3	2	4	2	3	2	2	1	1	4	2	3	5
fish	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1
crazed	1	2	1	1	1	2	4	5	1	2	1	1	1	2	1	3
sad	1	1	1	1	2	2	1	2	1	2	1	1	1	1	1	1
cashed	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
busy	1	1	3	4	1	5	1	5	1	1	1	1	1	2	1	2
sick	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1
shrimp	2	1	1	1	1	1	1	2	1	1	1	1	1	2	4	1
pleasure	1	2	5	5	1	5	1	5	5	5	5	2	5	5	3	5

Target Words	Trial 1, Rater 3															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	1	5	3	5	1	5	1	5	2	4	3	5	4	1	1	1
kiss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	1	2	1	4	1	1	5	2	2	2	3	2	1	1	1	4
washed	1	2	2	2	2	1	1	2	2	2	1	2	2	1	2	5
push	1	2	2	2	1	1	1	1	2	2	1	1	1	1	1	1
fast	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
cause	5	3	3	5	2	5	5	3	4	4	2	4	4	4	3	5
shrink	2	1	1	1	1	1	1	5	1	1	1	1	1	1	1	2
missing	2	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1
usually	5	3	5	5	5	5	1	5	5	4	5	2	5	5	3	5
miss	1	1	1	1	2	1	1	1	1	1	1	2	1	1	1	1
bishop	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
spark	1	2	1	1	1	1	1	1	1	1	1	1	1	1	2	1
zeal	2	4	5	5	1	3	3	5	5	5	3	5	4	2	1	2
facing	1	1	1	1	2	2	1	1	1	1	2	1	1	1	1	1
television	2	1	2	5	2	1	1	5	1	2	5	2	1	5	1	2
wasp	1	1	1	1	1	1	5	1	5	1	1	1	1	1	1	1
shark	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
fashion	1	1	2	2	1	1	1	1	2	2	1	1	1	1	1	1
razor	5	5	5	5	2	5	1	5	4	1	2	5	4	2	1	5
ship	2	2	1	1	1	1	1	2	1	1	1	1	1	1	1	1
step	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
maze	2	4	4	2	1	4	2	3	2	2	2	1	4	2	4	5
fish	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
crazed	1	2	1	1	1	1	5	5	2	2	1	1	1	2	1	2
sad	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
cashed	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
busy	2	1	4	5	1	5	1	4	1	1	1	1	1	2	1	1
sick	1	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1
shrimp	2	2	1	2	1	1	1	1	1	2	1	1	1	2	3	1
pleasure	1	2	5	5	1	5	2	5	5	5	5	2	5	5	2	5

Target Words	Trial 1, Rater 4															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	3	5	1	5	1	5	3	4	3	5	5	2	1	1
kiss	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	2	2	1	5	1	1	5	3	2	2	4	2	1	1	1	5
washed	1	2	3	2	2	1	1	2	3	2	2	2	2	2	3	5
push	2	1	2	2	1	1	1	1	2	2	1	1	1	1	1	1
fast	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1
cause	5	4	5	5	3	5	5	2	4	5	3	5	5	4	5	5
shrink	2	1	1	1	1	1	1	5	1	1	1	1	1	2	1	1
missing	1	1	1	1	2	1	1	1	2	1	2	2	1	1	1	1
usually	5	3	5	5	5	5	2	5	5	5	5	2	5	5	3	5
miss	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	2
bishop	1	1	2	2	1	1	1	2	1	1	1	1	2	1	1	1
spark	1	1	1	1	2	1	1	1	1	1	1	1	1	2	1	1
zeal	2	5	5	5	2	3	1	5	5	4	3	5	4	3	1	2
facing	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1
television	2	2	3	5	2	2	1	5	2	2	5	3	2	5	1	2
wasp	1	1	2	2	1	1	5	1	5	1	1	1	1	2	1	1
shark	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1
fashion	1	1	2	2	1	1	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	5	3	1	2	5	5	3	1	5
ship	2	1	1	1	1	1	2	1	1	1	1	1	1	2	1	1
step	2	1	1	1	2	1	1	1	1	1	2	1	1	2	1	1
maze	2	5	5	3	1	5	2	3	2	3	1	1	4	3	4	5
fish	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
crazed	1	2	1	1	1	1	5	5	2	2	1	1	1	2	1	3
sad	1	1	1	1	1	1	1	1	2	1	1	1	1	1	2	2
cashed	2	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1
busy	2	1	4	5	1	5	1	5	1	1	1	2	1	3	1	1
sick	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	2
shrimp	3	1	2	1	1	1	1	1	1	2	1	1	1	3	4	2
pleasure	2	2	5	5	1	5	1	5	5	5	5	3	5	5	3	5

Target Words	Trial 1, Rater 5															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	4	5	1	5	1	5	2	4	3	5	5	1	1	1
kiss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	1	3	2	4	1	1	4	2	1	1	3	1	1	1	1	4
washed	1	2	1	2	3	2	1	2	2	2	1	2	2	1	1	5
push	1	2	2	2	1	1	1	1	2	1	1	1	1	1	1	1
fast	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1
cause	4	3	3	4	3	5	5	3	4	4	2	4	4	4	4	4
shrink	2	1	1	1	1	1	1	5	1	1	1	1	1	1	1	1
missing	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
usually	5	3	5	5	5	5	1	5	5	5	5	2	5	5	3	5
miss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
bishop	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
spark	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
zeal	2	5	4	5	1	2	1	5	5	4	3	5	4	2	1	1
facing	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1
television	2	2	2	5	1	1	1	5	1	1	5	2	1	5	1	2
wasp	1	1	2	1	1	1	5	1	5	1	1	1	1	1	1	1
shark	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1
fashion	1	1	2	1	1	1	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	5	3	1	1	5	5	3	1	5
ship	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
step	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1
maze	1	5	3	2	2	4	2	4	2	2	1	1	4	3	3	5
fish	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
crazed	1	2	1	1	1	1	4	5	1	1	1	1	1	2	1	3
sad	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
cashed	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
busy	1	1	3	5	1	5	1	5	1	1	1	1	1	2	1	1
sick	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
shrimp	2	1	1	1	1	1	1	1	1	1	1	1	1	2	4	2
pleasure	1	3	5	5	2	5	1	5	5	5	5	2	5	5	2	5

Target Words	Trial 2, Rater 1															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	4	5	1	5	3	5	1	5	3	5	5	2	1	1
kiss	2	1	2	1	1	1	1	1	1	2	1	1	1	1	1	2
closed	3	4	3	5	1	1	2	3	1	1	4	1	1	1	1	5
washed	2	2	2	2	3	2	1	2	2	3	2	1	2	1	2	5
push	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1
fast	1	2	1	2	1	1	1	1	2	3	1	1	1	2	1	1
cause	4	5	5	5	5	4	2	5	5	5	4	5	4	5	5	5
shrink	1	2	2	1	1	2	1	5	1	1	1	1	1	1	1	1
missing	1	5	2	1	2	1	2	2	1	1	1	1	2	1	1	1
usually	5	5	5	5	5	5	1	5	5	5	5	5	5	5	2	5
miss	2	1	2	1	1	2	2	1	1	2	1	1	1	1	1	1
bishop	1	1	1	2	2	1	1	1	4	1	1	1	1	4	1	1
spark	2	2	1	2	1	1	1	1	2	1	2	1	1	2	1	1
zeal	1	5	3	5	1	5	3	5	5	5	5	5	4	3	1	1
facing	2	2	2	2	2	1	2	2	1	2	1	1	1	1	1	1
television	2	2	2	5	5	2	2	5	2	2	5	4	1	4	1	1
wasp	1	1	1	1	2	1	5	1	5	2	1	1	1	1	1	1
shark	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1
fashion	1	1	1	2	1	2	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	5	2	1	1	2	4	5	1	2
ship	1	2	1	2	2	1	1	2	1	1	1	1	1	1	1	2
step	1	1	2	2	2	2	1	2	1	1	1	1	1	1	1	1
maze	3	5	5	5	3	3	1	5	2	4	3	2	1	2	1	1
fish	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2
crazed	2	4	5	2	1	2	5	5	1	1	1	1	5	1	1	4
sad	2	2	2	2	2	2	2	2	1	2	2	5	2	1	1	2
cashed	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1
busy	2	5	3	1	1	5	1	4	2	1	1	1	1	1	1	2
sick	2	2	1	2	2	2	1	1	2	2	1	1	1	1	1	1
shrimp	2	1	2	1	2	2	2	1	2	4	1	1	1	1	1	4
pleasure	3	3	4	5	3	2	1	5	3	5	1	2	5	5	2	1

Target Words	Trial 2, Rater 2															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	3	5	1	5	2	4	1	5	3	5	5	2	1	1
kiss	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	2	5	4	4	1	1	2	2	1	1	4	1	2	1	1	4
washed	1	2	3	2	2	1	1	2	2	2	1	1	1	1	2	5
push	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fast	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1
cause	4	5	4	5	5	4	1	5	4	5	4	5	5	4	5	5
shrink	1	2	1	1	1	1	1	5	1	1	1	1	1	1	1	1
missing	1	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1
usually	5	4	5	5	5	5	1	5	5	4	5	5	5	5	1	5
miss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
bishop	1	1	1	2	2	1	1	1	5	1	1	1	1	3	1	1
spark	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
zeal	1	5	2	5	1	5	2	4	5	5	4	5	5	2	1	2
facing	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1
television	1	2	2	5	5	1	2	5	1	1	5	4	1	3	1	1
wasp	1	1	1	1	1	1	5	1	5	1	1	1	1	1	1	1
shark	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fashion	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	1	2	1	1	2	4	5	1	1
ship	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
step	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
maze	3	5	4	4	2	3	1	5	3	4	3	1	1	2	1	1
fish	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
crazed	2	4	4	2	1	1	5	5	1	1	1	1	5	1	2	3
sad	1	1	1	2	2	1	1	1	1	1	1	5	1	1	1	1
cashed	2	2	1	2	1	1	1	1	1	1	1	1	1	1	1	1
busy	2	5	4	2	1	5	1	3	2	1	1	1	1	1	1	2
sick	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
shrimp	2	1	1	1	1	2	1	1	2	5	1	1	1	1	1	3
pleasure	3	3	4	5	3	2	1	5	3	4	1	1	5	5	1	1

Target Words	Trial 2, Rater 3															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	4	5	1	5	2	5	1	4	3	5	5	3	1	1
kiss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	3	5	3	5	1	1	2	3	1	1	5	1	1	1	1	5
washed	2	2	2	2	2	2	1	2	2	2	1	1	1	1	1	5
push	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fast	1	1	1	2	1	1	1	1	1	3	1	1	1	1	1	1
cause	5	5	5	5	5	5	2	5	5	5	5	5	5	4	4	5
shrink	1	2	2	1	1	1	1	5	1	1	1	1	1	1	1	1
missing	1	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1
usually	5	5	5	5	5	5	1	5	5	5	5	5	5	5	1	5
miss	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1
bishop	1	1	1	1	2	1	1	1	4	1	1	1	1	3	1	1
spark	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
zeal	2	5	2	5	1	5	2	5	5	5	5	5	5	2	1	1
facing	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
television	2	2	2	5	5	1	1	5	1	1	5	5	1	3	1	2
wasp	1	1	1	1	1	1	5	1	5	2	1	1	1	1	1	1
shark	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
fashion	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	1	1	1	1	1	5	5	1	2
ship	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	2
step	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
maze	4	5	5	5	3	2	1	5	3	4	3	1	1	3	1	1
fish	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
crazed	2	3	4	2	1	1	5	5	1	1	1	1	5	2	1	4
sad	1	1	2	1	1	1	1	1	1	1	1	5	2	1	1	1
cashed	2	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1
busy	2	5	4	1	1	5	1	4	2	1	1	1	1	1	1	2
sick	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1
shrimp	2	1	1	1	1	1	1	1	3	4	1	1	1	1	1	4
pleasure	3	3	4	5	3	2	1	5	2	5	1	1	5	5	1	2

Target Words	Trial 2, Rater 4															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	4	5	1	5	2	5	1	5	3	5	5	2	1	1
kiss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
closed	2	5	4	5	1	2	2	3	1	1	4	1	1	1	1	5
washed	1	2	3	2	3	2	1	2	2	2	1	1	1	1	3	5
push	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
fast	1	1	1	1	1	1	1	1	1	3	1	1	1	2	1	1
cause	5	5	5	5	5	5	2	5	5	5	5	5	5	5	5	5
shrink	1	2	2	1	1	2	1	5	1	1	1	1	1	1	1	1
missing	1	5	1	2	1	1	2	1	1	1	1	1	1	1	1	1
usually	5	5	5	5	5	5	2	5	5	5	5	5	5	5	1	5
miss	2	1	1	1	1	2	1	1	1	2	1	1	1	1	1	1
bishop	1	1	1	2	3	1	2	1	5	1	1	1	1	4	1	1
spark	1	1	1	1	1	1	1	1	2	1	1	1	1	2	1	1
zeal	2	5	2	5	2	5	3	5	5	5	5	5	5	3	1	2
facing	1	1	2	2	1	1	1	2	1	2	1	1	1	1	1	1
television	1	2	2	5	5	2	1	5	1	2	5	4	1	4	1	1
wasp	1	1	1	1	2	1	5	1	5	2	1	1	1	1	1	1
shark	1	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1
fashion	1	1	1	2	1	1	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	1	2	1	1	2	5	5	1	3
ship	1	2	1	1	2	1	1	2	1	1	1	1	1	1	1	1
step	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1
maze	4	5	5	5	3	3	1	5	3	5	4	1	1	3	1	1
fish	1	1	1	2	2	1	1	1	1	1	1	1	1	1	1	1
crazed	3	5	5	3	1	2	5	5	1	1	1	1	5	2	1	4
sad	2	1	2	2	2	2	1	2	1	2	1	5	1	1	1	2
cashed	2	2	1	2	2	1	1	1	1	1	1	1	1	1	1	1
busy	2	5	5	2	1	5	1	4	2	1	1	1	1	1	1	3
sick	2	2	1	1	1	2	1	1	1	1	1	1	1	1	1	1
shrimp	2	1	2	2	2	2	2	1	3	5	1	1	1	1	1	4
pleasure	3	3	5	5	3	2	1	5	3	5	1	1	5	5	1	2

Target Words	Trial 2, Rater 5															
	Female								Male							
	f 01	f 02	f 03	f 04	f 05	f 06	f 07	f 08	m 01	m 02	m 03	m 04	m 05	m 06	m 07	m 08
zip	2	5	4	5	1	5	2	5	1	5	2	5	5	2	1	1
kiss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
closed	2	4	3	5	1	1	2	2	1	1	3	1	1	1	1	5
washed	1	2	2	2	2	1	1	1	2	2	1	1	1	1	2	5
push	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
fast	1	1	1	1	1	1	1	1	1	3	1	1	1	1	1	1
cause	5	5	5	5	5	4	2	5	4	5	4	5	5	5	4	5
shrink	1	2	1	1	1	1	1	5	1	1	1	1	1	1	1	1
missing	1	5	1	1	1	1	1	1	1	1	1	1	1	1	1	1
usually	5	5	5	5	5	5	1	5	5	5	5	5	5	5	2	5
miss	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
bishop	1	1	1	2	2	1	1	1	4	1	1	1	1	3	1	1
spark	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
zeal	1	5	2	5	1	5	2	4	5	5	5	5	5	2	1	1
facing	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
television	1	2	1	5	5	1	1	5	1	1	5	5	1	3	1	1
wasp	1	1	1	1	1	1	5	1	5	1	1	1	1	1	1	1
shark	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
fashion	1	1	1	2	1	1	1	1	1	2	1	1	1	1	1	1
razor	5	5	5	5	1	5	1	1	2	1	1	2	4	5	1	2
ship	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
step	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
maze	4	5	5	4	2	2	1	5	2	5	3	1	1	2	1	1
fish	1	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1
crazed	2	4	4	2	1	1	5	5	1	1	1	1	5	1	1	4
sad	1	1	1	1	1	1	1	1	1	1	1	5	1	1	1	1
cashed	2	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1
busy	2	5	4	1	1	5	1	3	2	1	1	1	1	1	1	2
sick	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
shrimp	2	1	2	1	1	2	1	1	2	3	1	1	1	1	1	3
pleasure	3	3	4	5	3	2	1	5	3	4	1	2	4	5	1	1

Appendix D: Transcription of the participants' production of target sibilants

Trial 1

Trial 1		Female								Male							
Target words	target sounds	F1	F2	F3	F4	F5	F6	F7	F8	M1	M2	M3	M4	M5	M6	M7	M8
zip	/z/	/z/	/ð/	/ts/	/ð/	/z/	/ð/	/z/	/ts/	/z/	/ts/	/ts/	/ð/	/ts/	/z/	/z/	/z/
		1.8	5	3.4	5	1	5	1	5	2.2	4	3	4.8	4.8	1.4	1	1
kiss	/s/	/s̪/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/
		1.4	1.2	1.4	1	1	1	1	1	1	1	1	1	1	1	1.2	1
closed	/z/	/z/	/s/	/z/	/s/	/z/	/z/	/s/	/s/	/s/	/z/	/s/	/z/	/z/	/z/	/z/	/s/
		1.2	2.6	1.2	4.2	1	1.2	4.8	2.4	1.6	1.6	3.6	1.6	1	1	1	4.4
washed	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʂ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/tʃ/
		1.4	2.2	1.8	2	2.6	1.6	1.2	2.2	2.4	2.2	1.4	2	2	1.2	2	5
push	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1.2	1.6	1.8	1.6	1	1	1	1.2	2.2	1.6	1	1	1	1	1	1
fast	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/
		1.2	1	1.6	1.4	1	1	1	1	1	1.2	1.2	1	1	1	1	1
cause	/z/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/
		4.8	3.4	3.8	4.4	2.6	5	5	2.6	4.2	4.4	2.6	4.4	4.2	3.8	4	4.6
shrink	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/s/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		2	1	1	1.2	1	1	1	5	1	1.2	1.4	1	1	1.2	1.2	1.4
missing	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/
		1.4	1.2	1	1.2	1.2	1	1	1.2	1.2	1	1.6	1.8	1.2	1	1	1
usually	/ʒ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʒ/	/ʐ/	/ʐ/	/ tʂ /	/ʐ/	/ʒ/	/ʐ/	/ʐ/	/ʒ/	/ʐ/
		5	3.2	5	5	5	5	1.4	5	5	4.6	5	2	5	5	3	5
miss	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/
		1.2	1.2	1	1	1.2	1	1	1.2	1.2	1	1.2	1.2	1	1	1	1.4
bishop	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1.2	1.4	1.8	1.4	1	1	1	1.4	1	1.4	1.2	1	1.4	1	1	1
spark	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/
		1	1.2	1	1	1.2	1	1	1	1.2	1.2	1	1	1	1.2	1.4	1.2
zeal	/z/	/ð/	/ð/	/ts/	/ts/	/z/	/ð/	/z/	/ts/	/ts/	/ts/	/ts/	/ð/	/ts/	/z/	/z/	/z/
		2	4.8	4.4	5	1.4	2.6	1.4	5	5	4.4	3.2	5	4	2.2	1	1.6
facing	/s/	/s̪/	/s/	/s̪/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/
		1.2	1	1.2	1.2	1.2	1.8	1	1.4	1	1.2	1.4	1	1.4	1	1	1
television	/ʒ/	/ʒ/	/ʒ/	/ʐ/	/ts/	/ʒ/	/ʒ/	/ʒ/	/ʂ/	/ʒ/	/ʒ/	/ʐ/	/ʐ/	/ʒ/	/ʐ/	/ʒ/	/ʒ/
		2	1.8	2.2	5	1.6	1.4	1.2	5	1.6	1.6	5	2.2	1.4	5	1	2
wasp	/s/	/s/	/s/	/s̪/	/s/	/s/	/s̪/	/ʃ/	/s̪/	/ʃ/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/
		1	1	1.6	1.2	1	1.4	5	1.2	5	1	1	1	1	1.4	1	1
shark	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1.4	1.2	1.4	1.4	1	1.2	1.2	1	1	1.4	1.2	1	1	1	1.2	1
fashion	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1.2	1.2	2	1.6	1.2	1	1	1	1.2	2	1	1	1	1	1	1
razor	/z/	/ð/	/ð/	/ts/	/ts/	/z/	/ð/	/z/	/ð/	/ð/	/z/	/z/	/ð/	/ts/	/ts/	/z/	/ð/
		5	5	5	5	1.4	5	1	5	3.4	1	1.4	5	4.6	3	1	5
ship	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1.8	1.4	1	1	1	1	1.2	1.2	1	1	1	1	1	1.2	1	1
step	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/
		1.2	1	1	1	1.2	1	1	1.2	1.2	1.2	1.4	1	1	1.6	1	1.2
maze	/z/	/s/	/θ/	/ts/	/z̪/	/z/	/s/	/s/	/s/	/s/	/ts/	/s/	/z̪/	/ts/	/ts/	/s/	/θ/
		1.8	4.8	3.8	2.4	1.4	4.4	2	3.4	2	2.4	1.4	1.2	4	2.4	3.6	5
fish	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1	1	1.4	1.4	1	1	1	1	1	1.2	1	1	1	1	1	1
crazed	/z/	/z/	/s/	/z/	/z/	/z/	/z/	/s/	/ʃ/	/z̪/	/z/	/z̪/	/z/	/z/	/z̪/	/z/	/s/
		1	2	1	1	1	1.4	4.6	5	1.6	1.8	1.2	1	1.2	2	1	2.6
sad	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s̪/
		1.2	1.2	1.2	1	1.2	1.2	1.2	1.2	1.4	1.4	1	1	1	1	1.4	1.4
cashed	/ʃ/	/ʂ/	/ʃ/	/ʃ	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/
		1.4	1	1.2	1.4	1	1	1.4	1	1	1	1	1	1	1.4	1	1
busy	/z/	/z/	/z/	/ð/	/ð/	/z/	/ð/	/z/	/ts/	/z/	/z/	/z̪/	/z̪/	/z/	/z̪/	/z/	/z/
		1.6	1	3.6	4.8	1	5	1	4.8	1	1.2	1.2	1.4	1	2.2	1	1.4
sick	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s̪/	/s/	/s/	/s/	/s̪/	/s/	/s/
		1	1	1	1	1	1	1	1.2	1	1.6	1	1.2	1.2	1.2	1.2	1.2
shrimp	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/
		2.2	1.2	1.2	1.2	1	1.2	1	1.4	1	1.6	1	1	1	2.2	3.6	1.6
pleasure	/ʒ/	/ʒ/	/ʐ/	/ʐ/	/ʐ/	/ʒ/	/ʐ/	/ʒ/	/ts/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/
		1.2	2.4	5	5	1.4	5	1.4	5	5	5	5	2.4	5	5	2.4	5

Trial 2

Trial 2		Female								Male							
Target words	target sounds	F1	F2	F3	F4	F5	F6	F7	F8	M1	M2	M3	M4	M5	M6	M7	M8
zip	/z/	/z/	/ð/	/ts/	/ð/	/z/	/ð/	/ts/	/ts/	/z/	/ts/	/ts/	/ð/	/ts/	/ts/	/z/	/z/
		2	5	3.8	5	1	5	2.2	4.8	1	4.8	2.8	5	5	2.2	1	1
kiss	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/
		1.4	1	1.2	1	1	1	1	1	1	1.2	1	1	1	1	1	1.4
closed	/z/	/s/	/s/	/s/	/s/	/z/	/z/	/z/	/s/	/z/	/z/	/s/	/z/	/z/	/z/	/z/	/s/
		2.4	4.6	3.4	4.8	1	1.2	2	2.6	1	1	4	1	1.2	1	1	4.8
washed	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʂ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/tʃ/
		1.4	2	2.4	2	2.4	1.6	1	1.8	2	2.2	1.2	1	1.2	1	2	5
push	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1.2	1	1.2	1	1.2	1	1	1	1	1	1	1	1	1	1	1
fast	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/
		1	1.2	1	1.4	1	1	1	1	1.2	2.8	1	1	1	1.4	1	1
cause	/z/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/ð/	/s/	/s/	/s/	/s/	/z/	/s/	/s/	/s/
		4.6	5	4.8	5	5	4.4	1.8	5	4.6	5	4.4	5	4.8	4.6	4.6	5
shrink	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/s/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1	2	1.6	1	1	1.4	1	5	1	1	1	1	1	1	1	1
missing	/s/	/s/	/θ/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/
		1	5	1.2	1.2	1.2	1	1.4	1.2	1	1	1	1	1.2	1	1	1
usually	/ʒ/	/ʐ/	/ʂ/	/ʐ/	/ʐ/	/ʐ/	/ʐ/	/ʒ/	/ʐ/	/ʐ/	/ʐ/	/ʂ/	/ʐ/	/ʐ/	/ʐ/	/ʒ/	/ʐ/
		5	4.8	5	5	5	5	1.2	5	5	4.8	5	5	5	5	1.4	5
miss	/s/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/
		1.4	1	1.2	1	1	1.6	1.2	1	1	1.4	1	1	1	1	1	1
bishop	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/
		1	1	1	1.8	2.2	1	1.2	1	4.4	1	1	1	1	3.4	1	1
spark	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s̪/	/s/	/s/
		1.2	1.2	1	1.2	1	1	1	1	1.4	1	1.2	1	1	1.6	1	1
zeal	/z/	/z/	/ð/	/ts/	/ð/	/z/	/ð/	/ts/	/ts/	/ð/	/ts/	/ts/	/ð/	/ts/	/ts/	/z/	/z/
		1.4	5	2.2	5	1.2	5	2.4	4.6	5	5	4.8	5	4.8	2.4	1	1.4
facing	/s/	/s/	/s/	/s̪/	/s̪/	/s/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/
		1.2	1.2	1.6	1.8	1.2	1	1.2	1.4	1	1.4	1	1	1	1	1	1
television	/ʒ/	/ʒ/	/ʐ/	/ʒ/	/ʂ/	/ʐ/	/ʒ/	/ʒ/	/ʂ/	/ʒ/	/ʒ/	/ʐ/	/ʂ/	/ʒ/	/ʐ/	/ʒ/	/ʒ/
		1.4	2	1.8	5	5	1.4	1.4	5	1.2	1.4	5	4.4	1	3.4	1	1.2
wasp	/s/	/s/	/s/	/s/	/s/	/s̪/	/s/	/ʃ/	/s/	/ʃ/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/
		1	1	1	1	1.4	1	5	1	5	1.6	1	1	1	1	1	1
shark	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1	1.2	1.8	1	1	1.2	1	1	1	1	1	1	1	1	1	1
fashion	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1	1	1	2	1	1.2	1	1	1	1.6	1	1	1	1	1	1
razor	/z/	/ð/	/ð/	/ts/	/ð/	/z/	/ð/	/z/	/ð/	/z/	/z/	/z/	/z/	/ts/	/ts/	/z/	/ts/
		5	5	5	5	1	5	1	1.8	1.8	1	1	1.8	4.4	5	1	2
ship	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1	1.6	1	1.2	1.6	1	1	1.4	1	1	1	1	1	1	1	1.4
step	/s/	/s/	/s/	/s̪/	/s/	/s̪/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/	/s/	/s/
		1	1	1.4	1.2	1.4	1.2	1	1.2	1	1	1	1	1	1	1	1
maze	/z/	/s/	/θ/	/s/	/s/	/s/	/s/	/z/	/ð/	/s/	/ts/	/ts/	/z/	/z/	/s/	/z/	/z/
		3.6	5	4.8	4.6	2.6	2.6	1	5	2.6	4.4	3.2	1.2	1	2.4	1	1
fish	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		1	1	1	1.8	1.2	1	1	1	1	1	1	1	1	1	1	1.2
crazed	/z/	/ts/	/s/	/s/	/ts/	/z/	/z/	/s/	/ʃ/	/z/	/z/	/z/	/z/	/θ/	/z/	/z/	/s/
		2.2	4	4.4	2.2	1	1.4	5	5	1	1	1	1	5	1.4	1.2	3.8
sad	/s/	/s̪/	/s/	/s̪/	/s̪/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/θ/	/s̪/	/s/	/s/	/s/
		1.4	1.2	1.6	1.6	1.6	1.4	1.2	1.4	1	1.4	1.2	5	1.4	1	1	1.4
cashed	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/
		2	2	1.2	1.6	1.8	1	1	1	1	1	1	1	1	1	1	1
busy	/z/	/ts/	/ð/	/ð/	/z/	/z/	/ð/	/z/	/ð/	/ts/	/z/	/z/	/z/	/z/	/z/	/z/	/ts/
		2	5	4	1.4	1	5	1	3.6	2	1	1	1	1	1	1	2.2
sick	/s/	/s/	/s̪/	/s/	/s̪/	/s/	/s̪/	/s/	/s/	/s̪/	/s/	/s/	/s/	/s/	/s/	/s/	/s/
		1.6	1.6	1	1.2	1.2	1.4	1	1	1.2	1.2	1	1	1	1	1	1
shrimp	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/	/ʃ/	/ʃ/	/ʂ/	/ʂ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʃ/	/ʂ/
		2	1	1.6	1.2	1.4	1.8	1.4	1	2.4	4.2	1	1	1	1	1	3.6
pleasure	/ʒ/	/ʐ/	/ʐ/	/ʂ/	/ʐ/	/ʐ/	/ʐ/	/ʒ/	/ʂ/	/ʐ/	/ʐ/	/ʒ/	/ʒ/	/tʂ/	/ʐ/	/ʒ/	/ʒ/
		3	3	4.2	5	3	2	1	5	2.8	4.6	1	1.4	4.8	5	1.2	1.4

Appendix E: Participants’ responses to the perception test

NOTE: ✓ indicates correct response and ✗ indicates incorrect response.

Alveolar /s/ replaced with dental /s̪/
/ʃ/ replaced with /ʂ/
/ʒ/ replaced with /ʐ/
/z/ replaced with /ð/

Trial 1

Trial 1	Female								Male							
Subjects	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
zip	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
kiss	✓	✗	✓	✓	✗	✗	✓	✓	✗	✓	✓	✗	✓	✓	✗	✓
closed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
washed	✓	✓	✗	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
push	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓
fast	✓	✗	✓	✓	✗	✓	✓	✗	✗	✓	✗	✗	✓	✗	✗	✓
cause	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
shrink	✗	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
missing	✗	✓	✓	✗	✓	✓	✗	✓	✓	✗	✗	✓	✓	✓	✗	✗
usually	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
miss	✓	✓	✓	✗	✗	✗	✗	✗	✓	✗	✓	✗	✓	✓	✓	✗
bishop	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
spark	✗	✓	✗	✗	✓	✓	✗	✓	✗	✓	✗	✓	✗	✗	✓	✗
zeal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
facing	✓	✓	✓	✓	✓	✗	✓	✗	✓	✗	✗	✓	✓	✗	✓	✗
television	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
wasp	✗	✗	✓	✓	✓	✗	✓	✓	✗	✗	✓	✗	✓	✗	✗	✓
shark	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
fashion	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓
razor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ship	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
step	✓	✗	✗	✗	✓	✓	✗	✓	✓	✗	✓	✓	✓	✗	✗	✓
maze	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
fish	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
crazed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
sad	✓	✓	✓	✗	✗	✓	✓	✓	✗	✗	✓	✗	✓	✓	✗	✗
cashed	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓
busy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
sick	✗	✓	✗	✓	✓	✗	✓	✗	✓	✗	✗	✓	✗	✗	✓	✗
shrimp	✓	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
pleasure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓

Trial 2

Trial 2	Female								Male							
Subjects	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
zip	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
kiss	✗	✗	✓	✓	✓	✗	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗
closed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
washed	✗	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓
push	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
fast	✓	✓	✗	✓	✓	✗	✓	✗	✓	✗	✗	✗	✗	✓	✗	✗
cause	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
shrink	✓	✗	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
missing	✗	✗	✓	✗	✓	✓	✓	✓	✓	✓	✗	✗	✓	✓	✓	✗
usually	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
miss	✗	✓	✓	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✗	✗	✗
bishop	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
spark	✗	✓	✗	✓	✗	✓	✗	✓	✗	✗	✗	✓	✗	✓	✓	✗
zeal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
facing	✓	✗	✗	✓	✓	✓	✗	✓	✓	✗	✗	✗	✓	✗	✓	✓
television	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
wasp	✗	✗	✓	✓	✓	✗	✓	✗	✓	✓	✓	✓	✗	✓	✓	✗
shark	✓	✗	✓	✗	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓
fashion	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓
razor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
ship	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
step	✗	✓	✓	✗	✗	✗	✗	✗	✓	✓	✓	✗	✓	✗	✗	✓
maze	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
fish	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓
crazed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
sad	✗	✗	✓	✓	✗	✓	✗	✗	✓	✓	✓	✗	✓	✗	✓	✗
cashed	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓
busy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
sick	✓	✓	✗	✗	✓	✗	✗	✓	✓	✗	✗	✓	✗	✗	✗	✗
shrimp	✓	✓	✗	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓
pleasure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓

/ʒ/ replaced with /ʂ/

/z/ replaced with /s/

Trial 1

	Female								Male							
Subjects	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
zip	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
closed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
cause	✓	✗	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓
usually	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
zeal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
television	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓
razor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
maze	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
crazed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
busy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
pleasure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓

Trial 2

	Female								Male							
Subjects	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
zip	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
closed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
cause	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
usually	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
zeal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
television	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
razor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
maze	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
crazed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
busy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
pleasure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓

/z/ replaced with /ts/

/ʒ/ replaced with /tʂ/

Trial 1

	Female								Male							
Subjects	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
zip	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓
closed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
cause	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
usually	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
zeal	✓	✗	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓
television	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
razor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
maze	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
crazed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
busy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
pleasure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

Trial 2

	Female								Male							
Subjects	**F1**	**F2**	**F3**	**F4**	**F5**	**F6**	**F7**	**F8**	**M1**	**M2**	**M3**	**M4**	**M5**	**M6**	**M7**	**M8**
zip	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✗	✓	✓	✓
closed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
cause	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
usually	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
zeal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
television	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
razor	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
maze	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
crazed	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
busy	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
pleasure	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓