Utrecht University
Department of Modern Languages
Bachelor's Thesis

Dealing with Dragons: A Linguistic Perspective on Intercultural Business Communication in China

Name:	David Franssen
Student Number:	3732606
Address:	Ina Boudier-Bakkerlaan 43, 3582 VH Utrecht
Bachelor:	English Language and Culture
Supervisor:	dr. Aoju Chen
Second Reader:	dr. Rias van den Doel
Course:	Principles and Parameters of Second Language Acquisition
Model:	British English
Date:	1 April 2015
Word Count:	8931 words

Table of Contents

INTELLECTUAL PROPERTY STATEMENT

Utrecht University defines "plagiarism" as follows: "If, in a thesis or some other paper, data or parts of a text produced by someone else are used without the source being identified, this shall be considered plagiarism. Among other things, plagiarism may entail the following:

cutting and pasting text from digital sources such as encyclopedias or digital journals, without using quotations marks and references;

cutting and pasting any text from the internet without using quotation marks and references; copying from printed material such as books, journals or encyclopedias without using quotations marks and references;

using a translation of the above texts in your own work, without using quotations marks and references; paraphrasing the above texts without using references. A paraphrase should never consist of merely replacing some words by synonyms;

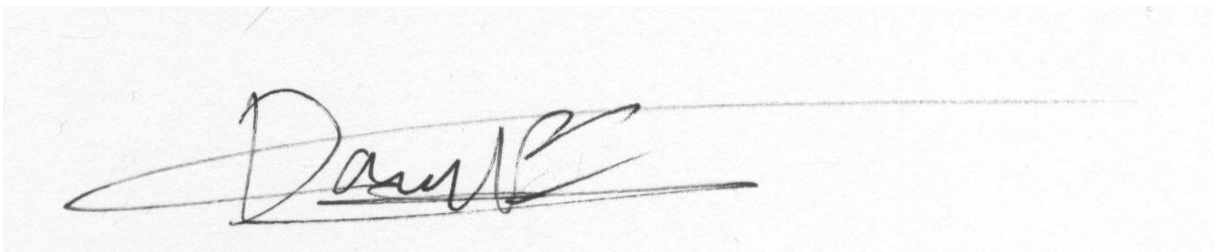
using pictures, sound recordings, or test materials produced by others without references, such that it appears that this is one's own work;

copying work by other students and passing this off as one's own work. If this is done with the other student's consent, the latter shall be an accomplice to the plagiarism;

even in cases where plagiarism is committed by one of the authors collaborating on a paper, the other authors shall be accomplices to plagiarism if they could or ought to have known that the first-mentioned author was committing plagiarism;

submitting papers acquired from a commercial source (such as an internet site offering summaries or complete essays) or written by someone else for payment."

I have read the above definition of plagiarism and certify with my signature below that I have not committed plagiarism in the appended thesis.



David Servatius Gregor Maria Franssen, 1 April 2015

Abstract

Mandarin Chinese is often regarded as a difficult language to master. In addition to being a tonal language it contains consonant and vowel sounds that speakers of Germanic and Romance languages may be unfamiliar with. Although it is well-established that pronunciation errors are associated with foreign accent and foreign accent can influence perceived personalities of speakers, little is known on how different types of pronunciation errors in L2 Chinese influence native listener's perception of attributes important to business communication in L2 Chinese compared to native listeners' perception of linguistic attributes. The present study was conducted to address exactly this question. Short sentences that contain various types of errors were composed and these sentences were then recorded by a Dutch learner of Chinese. Native speakers of Chinese were asked to rate these recorded sentences on several traits regarding personality and linguistic ability. The results indicated that erroneous speech had a significant impact on ratings for linguistic ability. The more errors in pronunciation, the lower his ability to speak the language was rated by native listeners. On the other hand, the effects of erroneous speech on ratings for business communication attributes were less clear. It was clear, however, that segmental errors had a stronger impact on the ratings for business communication attributes than tonal errors. Within the category of segmental errors, vowel errors led to significantly more negative scores than consonantal errors in two-thirds of the cases. A clear difference in the impact small and large tonal errors had on the ratings was only visible for linguistic attributes.

Keywords: Chinese; Erroneous; Speech; Pronunciation; Segmental; Tone

1 Introduction

The effects of foreign accented speech on the perceived personality of the speaker have been researched extensively. Kurowski et al. (1996) showed how erroneous speech was considered more foreign accented than non-erroneous speech. This claim was further supported by earlier research on what makes speech foreign accented (Blumstein et al., 1987; Gurd et al., 1988; and Ingram et al., 1992) However, these particular studies focused on Germanic or Romance languages. We are interested in the effects that erroneous production of Chinese, by L2 learners, has on a number of perceived business communication attributes. As a basis for the assumption that the relationship between erroneous speech and degree of foreign accented holds for Chinese as well, a vast body of literature will be summarised and discussed. Finally, literature on the relationship between degree of foreign accented and native listeners' perception of linguistic and business communication attributes will be discussed.

1.1 Literature review

Mandarin Chinese (also known as Putonghua, Standard Chinese, hereafter Chinese) is widely recognised as a hard language to master. It is not uncommon for L2 learners of Chinese to struggle with its pronunciation, in particular with the production of lexical tones. Previous research on degree of foreign accent and its effects mainly focused on three issues. First, what learner-specific factors influence the degree of foreign accent? Secondly, what constitutes degree of foreign accent from a linguistic point of view? Finally, how does the degree of foreign accent affect perceived personality traits?

To begin with, Piske et al. (2001) reviewed thirty years of literature on foreign accent and identified seven key learner-specific factors that influenced the degree of foreign accent. The first factor was the age of L2 learning (AOL). A correlating effect between the AOL and degree of foreign language was proven in numerous studies (Asher & García, 1969; Fathman,

1975; Seliger et al., 1975; Suter, 1976; Oyama, 1976; Tahta et al., 1981; Piper & Cansin, 1988; Flege, 1988; Patkowski, 1990; Thompson, 1991; Flege & Fletcher, 1992; Flege et al., 1995; Flege et al., 1999; Moyer, 1999). The data gathered on this topic so far suggested that a younger AOL led to a lower degree of foreign accent (p. 195-197).

The second factor was the length of residence (LOR), measured as the number of years spent in the country of the target language. The influence of LOR on degree of foreign accent has been heavily debated; some studies reported an influence of LOR on degree of foreign accent (Asher & García, 1969; Purcell & Suter, 1980; Flege & Fletcher, 1992; Flege et al., 1995; Flege et al., 1999) while other studies did not find an effect of LOR on degree of foreign accent (Oyama, 1976; Tahta et al., 1981; Flege, 1988; Piper and Cansin, 1988; Thompson, 1991; Elliott, 1995; Moyer, 1999). The general conclusion from the studies on LOR appeared to be that whenever data showed LOR functioned as a predictor for degree of foreign accent, LOR was a less accurate predictor than AOL (Piske et al., 2001, p. 197-199).

The third factor was gender. Literature review of studies on the effect of gender on degree of foreign accent showed conflicting results. Gender was found to have an effect on degree of foreign accent in several studies (Asher & García, 1969; Tahta et al., 1981; Thompson, 1991). However, a large number of studies on the topic contradicted the views posed by these studies and claimed that gender did not function as a predictor for degree of foreign accent (Olson & Samuels, 1973; Suter, 1976; Snow & Hoefnagel-Höhle, 1977; Purcell & Suter, 1980; Flege & Fletcher, 1992; Elliott, 1995). Although previous studies showed conflicting results, in most cases where gender did serve as a predictor for degree of foreign accent, females scored better than males (Piske et al., 2001, p. 199-200).

The fourth factor was formal instruction. Although research by Flege & Fletcher (1992) showed that the amount of time spent receiving formal instruction functioned as a predictor for degree of foreign accent, more often research did not find evidence for this claim

(Thompson, 1991; Elliott, 1995; Flege et al., 1995; Flege et al., 1999). The general conclusion from the studies on formal instruction appeared to be that there was little evidence this factor influenced degree of foreign accent (Piske et al., 2001, p. 201-202).

The fifth factor was motivation. The evidence found on the accuracy of motivation as a predictor for degree of foreign accent varied heavily. The most convincing claim was that speakers with a strong professional motivation were more likely to reduce their foreign accent than speakers without a strong professional motivation (Bongaerts et al., 1995; Moyer, 1999). On the whole, however, research on the influence of motivation on degree of foreign accent seemed to indicate that although motivation did show to have some influence on degree of foreign accent, the use of motivation as a predictor for foreign accent was not preferable to using AOL or LOR (Piske et al., 2001, p. 201-202).

The sixth factor was language learning aptitude. Research on the effect of a participant's musical ability did not find evidence that language learning aptitude functioned as a predictor for degree of foreign accent (Tahta et al., 1981; Thompson, 1991; Flege et al., 1995). On the other hand, research on the effect of a participant's ability to mimic sounds by Tahta et al. (1981) and Thompson (1991) did provide evidence to support the claim that the performance of the participants on mimicking sounds could serve as a predictor for degree of foreign accent.(Piske et al., 2001, p. 202-203).

The final factor was language use. The amount of time spent speaking the target language without living in a country where that language is predominant, was also researched as a possible predictor for degree of foreign accent. Both Suter (1976) and Purcell & Suter (1980) found that a combination of AOL and pattern of language use could be used as a significant predictor for degree of foreign accent. However, they did not identify language use as a significant predictor for degree of foreign accent on its own. Further research also failed to find any significant correlation between language pattern and degree of foreign accent

(Tahta et al., 1981; Thompson, 1991; Flege & Fletcher, 1992; Elliot, 1995). However, research by Flege et al. (1995) identified the number of times an L2 was used in a bilingual setting as a significant predictor for degree of L2 accent and follow-up research by Flege et al. (1997) provided more evidence for the influence of L2 use on degree of L2 accent. Flege et al. (1999) and Guion et al. (1999, 2000) provided more support for this idea (Piske et al., 2001, p. 203-204).

Research on the influence of learner-specific factors on degree of foreign accent showed a huge number of possibilities for variety in the degree of foreign accent of L2 learners of any language. Considering the literature discussed so far it seems safe to assume that no two L2 learners can be selected to score identically on all the factors described. If a research were to focus on determining how different types of pronunciation errors influence degree of foreign accent, it seems plausible that the influence of the learner-specific factors described so far should be limited to a minimum. To distinguish the linguistic elements that might influence degree of foreign accent, a closer look is needed at both phonemic and prosodic features of language. Will poor pronunciation of the target language's vowels or consonants affect degree of foreign accent? To what extent will failure to adapt to L2 rhythm, stress and intonation patterns influence the degree of foreign accent?

Mareüil & Vieru-Dimulescu (2006) hypothesised that "if both phonemic and prosodic characteristics are important, […] a foreign accent should reflect both segmental and suprasegmental dimensions" (p. 3). However, most research has limited itself to focusing on either the segmental or suprasegmental dimension. Flege (1991), Flege et al. (1995), Flege et al. (1997) and more recently Tsukada et al. (2004) and Tsukada (2005) focused on this segmental dimension, and found evidence that a deviation from native-like pronunciation of the L2 language's vowels and consonants correlated with degree of foreign accent. Flege & Eefting (1987) and Flege et al. (1997) showed how deviations from native-like vowel

pronunciation were frequent and prevalent, even to L2 speakers with near unlimited access to native L2 speaking conversational partners. Virtually no research has been done, however, on mixing up vowels entirely in L2 speech production. As far as consonants go, research by for example Kurowski et al. (1996) showed that degree of foreign accent could be influenced by over-aspiration in the case of patients with foreign accent syndrome (FAS). Interestingly, the deviations in voicing, place and manner of articulation led to an increased degree of foreign accent. Kurowski et al. (1996) found formant frequency and formation to be the key factors determining the effect vowel pronunciation had on the degree of foreign accent.

Research on the influence of prosody on degree of foreign accent repeatedly showed that prosody was directly linked to the degree of foreign accent (Grover et al., 1987; Munro, 1995; Tajima et al., 1997; Wayland, 1997; Magen, 1998; Missaglia, 1999; Pennington & Ellis, 2000; Munro et al., 1999). Rhythm, stress and intonation all had an audible effect on the degree of foreign accent. A large amount of in-depth research is available on this topic. The general consensus has been that any deviation from native-like speech, either segmental or suprasegmental, adds to the degree of foreign accent of a speaker. Although there is no evidence that this also applies to Chinese tones in particular, it will be entirely in line with virtually all research conducted on this topic if it proves to do so. However, the abundance of evidence for practically any deviation from native-like pronunciation to influence degree of foreign accent also implies that when researching the influence of segmental errors on degree of foreign accent suprasegmental errors should not be allowed to find their way into an experiment, and vice-versa.

The main focus of the present study is to collect data on how degree of foreign accent affects business communication attributes for learners of Chinese. Literature review by Gluszek and Dovidio (2010) described research on general encounters between native and non-native speakers of a language and summed up a vast amount of evidence for stereotypes

and prejudices towards non-native speakers of a language. More specifically, they noted that individuals who had non-native accents were viewed as less intelligent (Rubin et al., 1997; Lindemann, 2003), less competent (Bresnahan et al., 2002; Boyd, 2003), and as speaking the language poorly (Lindemann, 2003; Hosoda et al., 2007) (p. 2-6). These studies were conducted using different methods, speaker nationalities and participants in different circumstances, yet all found that a more recognisable foreign accent led to poorer scores on personality attributes.

Rubin et al. (1997) used the speech evaluation instrument (SEI) to measure the students' rating of attributes related to their teachers' knowledgeability and professional competence as physicians. The speakers in the Rubin et al. experiment were selected by the severity of their South-Asian accents, and analysis of the data showed that a more severe South-Asian accent was met by lower ratings from their students across the board, but most notably for intelligence and intelligibility.

Lindemann (2003) used Korean L2 learners of English, both male and female, and compared their scores to Mid-Western dialect speakers of US English. Out of a large pool of speakers a jury of two native English speakers selected the best two male and best two female speakers for each of the two languages. All the speakers were asked to read the same texts for recording, the very few oddities in their prosodic execution were later revisited to make sure they did not impact the outcome of the experiment inadvertently. The rated attributes for each of the speakers ranged from attributes to do with professionalism, such as: *ambition*, *incompetence* or *intelligence*, to linguistic attributes such as: *nice to listen to, speaks poorly* or *native speaker*. The data collected in this experiment showed a clear relationship between the scores on the linguistic attributes and that for intelligence.

Boyd (2003) featured L2 speakers of Swedish that worked as school teachers in Sweden, but were born and raised in other countries with different primary languages. In this

experiment their potential employers, in the form of school principals, were asked to rate both on a set of attributes explicitly focusing on linguistic attributes, such as: *fluency*, *word choice* and *pronunciation*, as well as their general *competence* as teachers. The data collected in this experiment showed a direct relationship between the scores on linguistic attributes and perceived competence as a teacher.

Bresnahan et al. (2002) conducted an experiment using two L2 speakers of American English and a single native speaker of American English. Before recording the stimuli, the three speakers were judged by a panel of 20 native listeners and the two L2 speakers were shown to score very differently. The first L2 speaker was rated equally intelligible as the native speaker; the second L2 speaker was rated much less intelligible than the others. The speakers were recorded and over 300 students were asked to give them scores using the SEI. In this experiment, however, the participants were also asked to rate them in different roles as either a potential friend or potential teaching assistant. The data collected in this experiment showed a relationship between the degree of foreign accent and perceived competence.

The study by Hosoda et al. (2007) resembled the studies by Lindemann (2003) and Rubin et al. (1997) to a large extent; speakers showed different degrees of Asian accent in their English speech were compared to native American English speakers. The data collected in this experiment also showed a relationship between degree of foreign accent and ratings for communicative ability. All in all, there was a clear connection between degree of foreign accent and lower scores on a great number of attributes, varying from prestige, friendliness and pleasantness (Gluszek & Dovidio 2010) to teaching competence (Rubin & Smith, 1990).

The literature reviewed in this section showed evidence for the claim that degree of foreign accent is influenced by many different factors and, in its turn, influences the perceived personality of a speaker as well as his professional competence. The degree of foreign accent is influenced by a number of learner specific attributes, ranging from AOL to the amount of

time spent speaking the target language. Meanwhile the speaker's performance in pronouncing the target language correctly will likely be influenced by these factors, while it will be impossible for the learner-specific factors to be influenced by the performance of the speaker. This means that when researching the influence of, for example, AOL on degree of foreign accent, erroneous speech should be a by-product of selecting speakers with different backgrounds. However, since the current study focuses on researching the different influences different kinds of errors in speech production have on degree of foreign accent, it seems reasonable to reduce the learner-specific factors to a minimum.

1.2 Research questions and hypotheses

The current study aims to investigate the effects of foreign accent on perceived language ability and business communication attributes. A comparison is made between the effects erroneous speech has on perceived language ability and the effects it has on business communication attributes. Furthermore, different types of segmental and tonal errors are compared to determine if there is a difference in the effect they have on these perceived attributes. In so doing, we address the following research questions:

1) *Do errors in speech production influence the ratings for language ability based attributes differently than they influence the business communication attributes of an L2 speaker of Chinese?*
2) *Do different types of errors influence the perceived business communication attributes of an L2 speaker of Chinese differently?*

In addition, the following hypotheses are put forward:

1) *Language ability based attributes of an L2 speaker of Chinese are more heavily influenced by speech production errors than perceived business communication attributes.*

2) *Different types of errors have different effects on perceived business communication attributes of an L2 speaker of Chinese.*

Hypothesis 1 (henceforth H1) was proposed because both research on the influence of segmental as well as that of suprasegmental errors on degree of foreign accent showed that whatever the deviation from a standardised pronunciation, scores for language ability were nearly always affected (Grover et al., 1987; Flege, 1991; Flege et al., 1995; Munro, 1995; Flege et al., 1997; Tajima et al., 1997; Wayland, 1997; Magen, 1998; Missaglia, 1999; Pennington & Ellis, 2000; Munro & Derwing, 2001; Tsukada et al., 2004; Tsukada, 2005). Research on the influence of degree of foreign accent on personality attributes that did not necessarily focus on the type of errors but rather used audibly foreign accented speakers to record their stimuli either showed significantly lower scores for language ability attributes (Hosoda et al., 2007; Lindemann, 2003), or made sure they scored low on language ability attributes to validate they were fit to function as speakers for the experiment, i.e. they could be considered speakers with a foreign accent (Bresnahan et al., 2002). Scores for other personality attributes, however, were not always affected in the same way. Research showed an influence of foreign accent on attributes varying from prestige, friendliness and pleasantness (Gluszek & Dovidio 2010) to teaching competence (Rubin & Smith, 1990). It appears viable to hypothesise that low scores on language ability attributes will be guaranteed while the effects on other attributes may differ.

Hypothesis 2 (henceforth H2) was proposed because, as stated earlier, research on the influence of segmental errors on degree of foreign accent has shown that incorrect pronunciation leads to lower scores for language ability. Earlier research showed that segmental errors influenced the degree of foreign accent (Flege, 1991; Flege et al., 1995; Flege et al., 1997; Tsukada et al., 2004; Tsukada, 2005). Research on deviations from native-like vowel pronunciation showed that these deviations happened on a regular basis (Flege &

Eefting, 1987; Flege, 1997), but did not investigate the specific effect these deviations had on personality attributes or even degree of foreign accent. Since different types of mistakes led to different degrees of perceived foreign accent in earlier research, these can be expected to also lead to different ratings for business communication attributes in the study at hand.

2 Method

Perception of someone's personality can be expected to involve a multitude of factors involving both parties' outward appearance as well as personal preferences. Simultaneously, setting and surroundings can play a role as well. To be able to test the hypotheses in a sound way, methodological decisions were made to minimise the influence of random factors (i.e. factors that were not systematically varied in the experiment).

2.1 Participants

We were interested in intercultural business communication. To this end, higher educated native speakers of Chinese were invited to participate in the experiment. This group could be expected to participate in intercultural business proceedings in the near future. The participants therefore were Chinese university students or recent graduates from different universities and with different geographic backgrounds, representing the next generation that will partake in intercultural business exchanges. Twenty native Chinese students and recent graduates (16 females, 4 males, M_{age} = 27.25, SD = 2.1) participated in the experiment.

2.2 Materials

2.2.1 Stimuli

The experiment focused on the different types of errors L2 speakers of Chinese are likely to make and the effect these errors have on perceived characteristics of speakers from L1 listeners' point of view. Four types of pronunciation errors were included, divided in the two categories segmental errors and tonal errors. Each of these categories consisted of two types of errors. The first type of segmental error was the vowel error. The vowels chosen to replace their correct counterparts were selected so that a vowel was never replaced by the same vowel twice. In addition, the new vowel had to create a clearly audible error. The selection was

made in careful deliberation with a phonetically trained native speaker of Chinese. For example, a vowel error would be: *"bi2 yong4 xie4,"* instead of the correct "*bu2 yong4 xie4."* Secondly, consonantal errors were included. A majority of Chinese syllables have a consonant+vowel (CV) structure. For this reason, consonantal errors were placed in syllable-initial position. The consonants chosen to replace their correct counterparts were selected through the use of a set of constraints.

1. A consonant was never replaced by the same consonant twice in the whole set of the stimuli
2. A voiced consonant was replaced by its unvoiced counterpart, and vice-versa
3. If a consonant did not have a voiced or unvoiced counterpart, it was replaced by another consonant with similar voicing and a place of articulation as close as possible to the original.

For example, a consonantal error would be: *"mu2 yong4 xie4,"* instead of the correct "*bu2 yong4 xie4."*

In addition to segmental errors, tonal errors were embedded in the stimuli. Chinese has four different lexical tones and so there are many ways in which to pronounce the tones erroneously. The difference between these tones is primarily pitch-based. The following image shows the waveform and pitch contour of the sound-sequence *ma* for all four tones of Chinese:

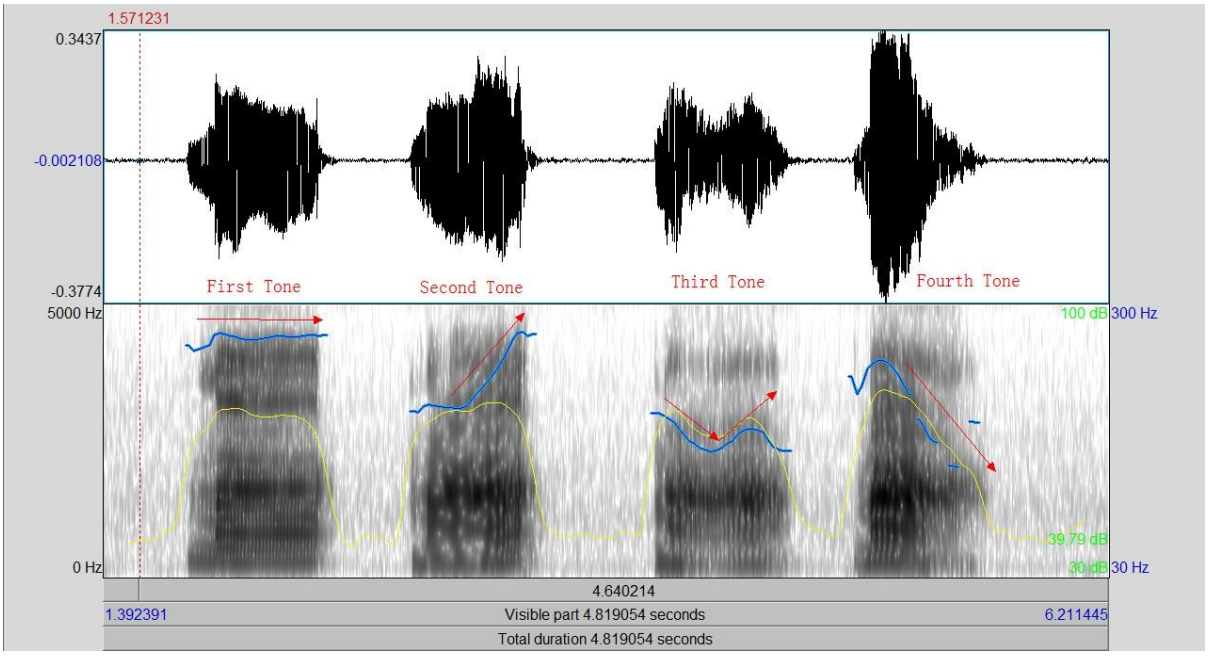


Figure 1. Pronunciation of "ma" in four different tones (taken from the Praat manual of Hong Kong University)

The red arrows in figure 1 show the changes in pitch over the course of syllable pronunciation. The first tone starts high pitched and stays high pitched. The second tone is a rising tone, whereas the fourth tone is a falling tone. Pronouncing the syllable using the third tone causes the pitch to fall and then rise again. Each of these tones modifies the vowel to an extent where the meaning of the word changes, creating a minimal pair. Chen (1999) explained that tone "is a production unit, just like segment, which can be selected from mental storage and inserted into a positional frame during phonological encoding" (p. 290). Since this experiment distinguished two types of segmental errors, the question was raised if there was a similar division possible for tonal errors. The variety of tones in Chinese allowed for two different types of tonal errors to be distinguished in the experiment. These were named *relatively small* and *relatively large tonal differences*. Although these could have been construed as two levels of the same error type, the difference in gravity of the error is so severe that the audible difference between the error types seemed not too far from the

difference between vowel and consonantal errors. Based on the direction of the change in pitch for each tone, these two types of tonal errors consisted of the following changes in tone:

Relatively Small Difference			Relatively Large Difference		
Tone 1	→	Tone 2	Tone 1	→	Tone 4
Tone 2	→	Tone 3	Tone 2	→	Tone 4
Tone 3	→	Tone 2	Tone 3	→	Tone 1
Tone 4	→	Tone 3	Tone 4	→	Tone 1

Figure 2. System of tonal swaps for small and large tonal error categories

The first category swapped tone 1 for tone 2, which ends around the same pitch height. Tone 2 was swapped for tone 3 as both show a rise in pitch. Tone 3 was swapped for tone 2 for the same reason and tone 4 was swapped for tone 3 as both initially show a fall in pitch height. For example, a small tonal error would be: *"bu3 yong4 xie4 ,"* instead of the correct "*bu2 yong4 xie4."* The second category swapped tone 1 for tone 4 as these tones show the largest difference in pitch height at the end of the syllable. Tone 2 was swapped for tone 4 as well as they move in opposite directions. Tone 3 was swapped for tone 1 as it starts with a dip in pitch height, whereas tone 1 is consistently high-pitched. Tone 4 was swapped for tone 1 for the reason stated earlier. For example, a small tonal error would be: *"bu4 yong4 xie4,"* instead of the correct "*bu2 yong4 xie4."*

To support the choice to distinguish these two categories we asked two native Chinese linguists to do an evaluation test. For each stimulus, each of the linguists received two sound-files. In one of the files the stimulus was manipulated to contain the small difference tonal error, in the other file the stimulus was manipulated to contain the large difference tonal error. They were given the sentence in simplified Chinese characters but were not given any additional information on the difference between the sound-files they received. These files were coded 1A, 1B, 2A, 2B and so forth. Files 1A, 2A, 3B and 4B contained a small

difference tonal error, while files 1B, 2B, 3A and 4A contained a large difference tonal error. They were then asked which of the two deviated more from what they would expect to hear based on the written form. The responses from both linguists matched the expected results based on our assumed degree of erroneousness. This justified our manipulation of the degree of erroneousness. In addition, two more categories were created for the sake of comparisons with the already specified categories, namely combinations of tonal and consonantal errors and combinations of tonal and vowel errors. For this category the relatively largely different tonal errors were combined with the segmental errors described earlier. For example, a tonal-vowel error would be: *"bi4 yong4 xie4,"* whereas a tonal-consonantal error would be: *"mu4 yong4 xie4,"* instead of the correct "*bu2 yong4 xie4."*

By combining two different types of errors the comprehensibility of the utterance in the stimulus should be reduced. This idea is supported by Friederici (2002) who used brain-scans to investigate the effect of erroneous or de-lexicalised speech on the process of auditory-sentence comprehension. She argued that "a bilateral temporo-frontal network subserves auditory-sentence comprehension" (p. 83). She described how this network is home to a series of processes and how "violations of pitch for lexical elements in a tonal language […] result in modulation of activation in the left frontal operculum adjacent to Broca's area" (p. 82). This increased activity would then result in an altered comprehensibility of the heard sentence (p. 81-83). In short; when words are being mispronounced it takes significantly longer for a listener to understand what is being said. Friederici's result thus suggested that sentences containing both tonal and segmental errors in the same words would be harder to comprehend than the sentences that have only one of these types of errors, and that this will subsequently reflect in the listener's judgement of the speaker's degree of foreign accent and fluency.

To establish a clear connection between the type of error and the perceived personality traits the error had to clearly stand out in an otherwise very correctly spoken sentence. Due to the nature of Chinese and the difficulties L2 learners have in pronouncing it correctly, the aim was to use the smallest possible strings of speech that still function as a complete sentence. This way, the error would be clearly audible within its surroundings, without its surroundings having too much of an impact on the judgement of the native listener. Tri-syllabic sentences are quite common in Chinese and were the best fit for this design. Examples are 不用谢 'you're *welcome',* and 晚上好 '*good evening'*. Eighteen short statements that can be used frequently due to their relatively general nature were selected for this experiment, as shown in Table 1.

#	**Chinese**	**Syllable 1**	**Syllable 2**	**Syllable 3**	**English**
1	别介意	bie2	jie4	yi4	Never mind
2	不用谢	bu2	yong4	xie4	You're welcome
3	不客气	bu2	ke4	qi0	You're welcome
4	大家好	da4	jia1	hao3	Hello everyone
5	对不起	dui4	bu4	qi3	I'm sorry
6	礼拜五	li3	bai4	wu3	Friday
7	恭喜你	gong1	xi3	ni3	Congratulations
8	回头见	hui2	tou2	jian4	See you later
9	我很好	wo3	hen3	hao3	I'm fine
10	没问题	mei2	wen4	ti2	No problem
11	没关系	mei2	guan1	xi0	It's okay
12	明天见	ming2	tian1	jian4	See you tomorrow

13	早上好	zao3	shang4	hao3	Good morning
14	您先请	nin2	xian1	qing3	After you
15	请签名	qing3	qian1	ming2	Please sign
16	请吃吧	qing3	chi1	ba0	Please eat
17	太好了	tai4	hao3	le0	Very good
18	晚上好	wan3	shang4	hao3	Good evening

Table 1: Experiment stimuli in Chinese, pinyin and English

However, sentence stress needed also be taken into account. Sentence stress can occur in different locations in Chinese, as in languages like English and Dutch. A stressed word is realised with a larger pitch excursion and a longer duration (Jin 1996). As such, stressed words are realised with prosodic prominence and consequently are perceptually salient. It was therefore important to avoid differences in perceptual salience of pronunciation errors by placing pronunciation errors in stressed words only.

To determine which syllable should get sentence stress in each of the 18 sentences, three phonetically trained native speakers of Chinese were asked to indicate for each sentence where they felt the sentence stress would appear most natural. Whenever there was no complete agreement but there was a 2 to 1 majority vote, this majority vote was selected for sentence stress. In fourteen of the original eighteen cases there was a majority vote for either the first or the final syllable, the only exceptions being stimuli 5, 9, 15 and 16. Stimulus 15 was removed before recording the stimuli because of its ambiguous nature, as 您贵姓 *'What is your name?*' is only used to ask for someone's name in a formal setting. It was replaced with an alternative that is more comparable to the other stimuli, i.e. 请签名 *'Please sign*'. Sentence stress for the new stimulus 15 was assigned to the initial syllable unanimously. Stimulus 16 showed stress on the pre-final syllable, which was unexpected. However, in this

stimulus the final syllable is tone-neutral, meaning this final syllable cannot be stressed. As a result, the tone shifted to the final syllable capable of carrying sentence stress. Stimuli 5 and 9 were left undecided and in these instances the experiment's speaker was asked to cast the decisive vote. A detailed representation of the results for this tally can be found in appendix A.

2.2.2 Stimulus recording

For this experiment a variation on the matched guise technique was used for recording the stimuli. Normally this setup would require a single bilingual speaker to produce all the stimulus conditions to rule out confounds based on voice-specific factors. While studies purely focusing on the degree of foreign accent and perceived personality usually have different speakers to compare, this study focused on different types of mistakes and therefore it was desirable to reduce the number of other factors that could influence the listeners' ratings to a minimum. However, for this experiment a bilingual speaker was not recruited. As the experiment was not designed to have multiple speakers with different degrees of foreign accent, the matched guise technique was not ideal. As the experiment was meant to shed light on the effects of pronunciation errors on business communication attributes in an intercultural setting the speaker would preferably be an intermediate learner of Chinese, the type that is likely to speak a little Chinese in an intercultural business setting and is able to pronounce the target sentences with a foreign accent but with an awareness of the intended pronunciation. The speaker selected for this experiment was a linguistically trained native Dutch-English bilingual, also fluent in German and French, who had achieved an intermediate level proficiency of Chinese as an L2 (the author). As a linguistically trained multilingual the speaker was expected to be more responsive to instructions during the recording given by a phonetically trained linguist and native speaker of Chinese.

Use of a second, female, speaker was considered. Research by Costa et al. (2001) showed that differences in perceived personality traits based on speech are linked to the local culture and gender role division. Although there is no evidence to date that male and female second language speakers of Chinese are consistently judged differently by Chinese native listeners, there is some literature on the influence of gender on perceived personality in general. Costa et al. also reported that "gender differences were most pronounced in European and American cultures in which traditional sex roles are minimized" (p. 322). Recent figures from He (2014) showed that the number of Chinese women that perform an important role in businesses is above average. The same was found for their role within governmental organisations (China Permanent Mission), while the workforce participation of Chinese women is traditionally higher than that of American or Japanese women (Chen et al. 1997). As such, it was speculated that in a business setting, gender differences in China should not serve as a predictor for the language capacity attributes in this experiment. Although this would have made a very interesting hypothesis, the addition of a second, female, speaker would have led to an increase in possible learner-specific factors influencing the perceived degree of L2. The main focus of this experiment was to distinguish between the different errors in phonetic execution of the stimuli. We therefore decided to use the material produced by one male speaker as experimental stimuli.

During the recording, the L2 Chinese speaker was closely monitored by the phonetically trained linguist. She paid close attention to sentence stress, phonetics and tone. Whenever the speaker failed to pronounce the stimulus as intended she immediately instructed him to retry, offering advice on the pronunciation or giving the correct pronunciation where necessary. For most of the 108 stimuli the number of attempts ranged from 1 to 3. All attempts were recorded and in case of poor quality of the recording, parts of the recordings were rerecorded. Three recording session were scheduled over a period of three

weeks in order to make sure all the recorded stimuli were as intended by the experimental design.

2.2.3 Attributes

Over the course of the last half-century speech evaluation research has been a popular topic of study. Many different ways of assessing an individual's response to speech were developed, and in 1985 Zahn & Hopper published a paper on the speech evaluation instrument, or SEI. The SEI divided speech evaluation up into three factors: superiority, attractiveness and dynamism (117-119). The SEI offered a long list of attributes that can be used for this type of research but these are best fitted for socio-psychological research. However, the attribute *fluency* was used from this scale, which represents the factor of superiority combined to some extent with dynamism. This attribute was described as a sort of language-capacity based indicator for socio-economic status (119). Tsalikis et al (1992) showed how participants that were asked to listen to a sales-pitch showed significantly better ratings in the speech evaluation of sales pitches presented them in Guatemalan Spanish. The sales pitches that were given in foreign accented Spanish received far worse ratings on the speech evaluation attributes credibility, friendliness, effectiveness and competence. These four attributes were used in the current study to rate the business communication qualities of the speaker. Furthermore, degree of foreign accent was used as an attribute for speech evaluation. By adding degree of foreign accent as an attribute a link could be established between the degree of degree of foreign accent and the other attributes. In short, the experiment included six different attributes, namely: credibility or 可信度, competence or 能力, degree of foreign accent or 外国口音, friendliness or 友好, effectiveness or 有效性 and fluency or 流利(程度). The first four of these attributes were included to measure the business communication qualities of the speaker, as perceived by the participants. The other two attributes, degree of foreign accent and fluency, are indicators of the speakers' competence at the target language. By studying

the interaction between the four personality-attributes and the two language-attributes possible evidence might be found to shed some light on the relationship between language competence and perceived personality.

2.2.4 Distribution

As mentioned in section 2.2.1, the stimuli consisted of eighteen pre-recorded tri-syllabic sentences, each recorded erroneously in six different ways. In this paragraph, *sentence* refers to one of the 18 sentences described in 2.2.2, whereas *stimuli* refers to one of the 108 (18 x 6) stimuli available in appendix B. The hundred-and-eight stimuli were divided into six blocks according to a full Latin Square design. Each block contained 3 stimuli from every one of the six different subcategories. The following table shows how these stimuli were distributed for every block:

	Vowel	Consonant	Tone small	Tone large	Tonal-vowel	Tonal-consonantal
Block 1	1,2,3	4,5,6	7,8,9	10,11,12	13,14,15	16,17,18
Block 2	4,5,6	7,8,9	10,11,12	13,14,15	16,17,18	1,2,3
Block 3	7,8,9	10,11,12	13,14,15	16,17,18	1,2,3	4,5,6
Block 4	10,11,12	13,14,15	16,17,18	1,2,3	4,5,6	7,8,9
Block 5	13,14,15	16,17,18	1,2,3	4,5,6	7,8,9	10,11,12
Block 6	16,17,18	1,2,3	4,5,6	7,8,9	10,11,12	13,14,15

Table 2. Distribution of stimuli across the experiment: the first block will contain sentence 1, 2 and 3 with a vowel error, sentence 4, 5 and 6 with a consonantal error, etc.

For every block, the eighteen stimuli were put in a random order with the sole criterion that no two stimuli from within the same category could directly follow one another.

Distributed amongst six different versions of the experiment, all these blocks could be rated for one of the six attributes described earlier. The following table shows how these blocks and attributes were distributed for every version:

	Version 1	Version 2	Version 3	Version 4	Version 5	Version 6
Block 1	可信度	能力	外国口音	友好	有效性	流利(程度)
Block 2	能力	外国口音	友好	有效性	流利(程度)	可信度
Block 3	外国口音	友好	有效性	流利(程度)	可信度	能力
Block 4	友好	有效性	流利(程度)	可信度	能力	外国口音
Block 5	有效性	流利(程度)	可信度	能力	外国口音	友好
Block 6	流利(程度)	可信度	能力	外国口音	友好	有效性

Table 3. Distribution of attributes across the different versions of the experiment

2.3 Procedure

Through personal networks, forty-eight native Chinese university students or recent graduates were approached with one of the six versions of the experiment and asked to download the audio-file. The instructions and the experiment itself were thus presented to the participants digitally. These instructions and an English translation of the instructions are available in Appendices C and D. The .wav audio-file containing all hundred-eight stimuli and one .wav example-file were made available through Dropbox.com, WeTransfer.com and SendSpace.com. The participants all listened to the same audio-file containing six blocks of eighteen stimuli.

The .wav file was created by pasting the stimuli together with silences in between using Praat. Each of the six blocks started with a beep, followed by a three-second pause and then the first stimulus. The eighteen stimuli in the block were played with a three second pause between each of the stimuli. At the end of each block another beep sounded, followed by a twenty second pause before the start of the next block. The .wav file could be played using any regular digital media player.

The participants were asked to rate a certain attribute of the speaker from one (lowest) to five (highest). For each block the participants rated one of the six attributes, based on their

version according to the distribution shown in table 3. Since the order of the stimuli across the blocks was always the same, participants only received different versions of the experiment. While listening to the audio-file, the participants could use the answer sheets they were sent to score the stimuli. The documents were prepped so as to make sure the participants only needed to click any of the boxes in order to tick it. The complete documents received by the participants thus consisted of seven pages. After the one-page Chinese introduction found in appendix C, each block occupied one page. These pages were created to only require the participants to click the most appropriate option for each stimulus. Those pages looked like this:

Attribute 1

友好

	1	2	3	4	5
1. 对不起	☐	☐	☐	☒	☐
2. 恭喜你	☐	☒	☐	☐	☐

Figure 3. Start of the second page of a returned document from the experiment

Participants were encouraged to select the most appropriate option intuitively. As stated in the instructions, participants were requested to return their documents by email upon completing the experiment.

3 Results

The data gathered from this experiment were analysed using SPSS Statistics 22. For all statistical tests an alpha level of .05 was used. Twenty of the forty-eight approached participants returned completed documents for the experiment, twenty-eight were non-responsive. The distribution of the participants over the six versions was as follows: versions 1, 2 and 4 received 2 responses, version 3 received 3 responses, version 6 received 5 responses and versions 5 received 6 responses.

Upon receiving the completed documents from the participants all the scores of 1 to 5 were manually copied to Windows Excel. A spreadsheet containing information on the participant's name, age, gender, dialect, profession, location and Chinese proficiency as well as the scoring of each stimuli and the attribute it was scored for was created. Because the attribute *foreign accent* was the only negative attribute out of the six, the scores were reversed so that an original score of 1 (no foreign accent) became 5, 2 became 4, and vice versa. This spreadsheet was then used to calculate the mean, standard deviation and standard error of mean of each participant's scores for every stimulus for every error type for every attribute. These were then used to calculate the mean, standard deviation and standard error of mean of each participants scores for all types of mistakes for each stimuli based on type of attribute (linguistic or business communication). These scores were then copied to SPSS Statistics for further investigation. The mean scores calculated were as follows:

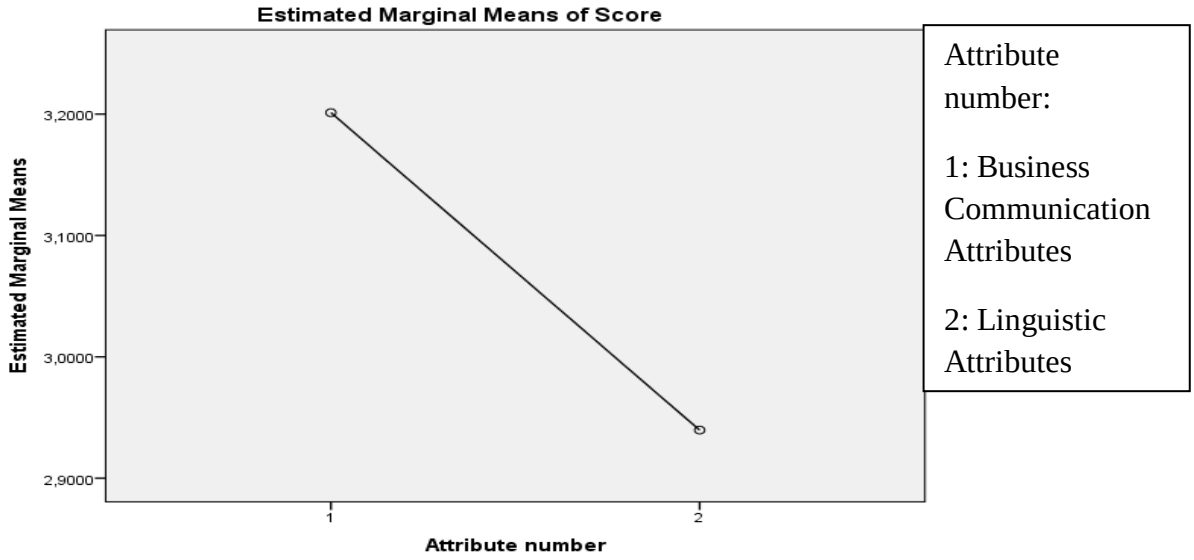
		Segmental		Tonal		Combination	
Attribute group	Attribute	Vowel	Consonant	Tone large	Tone small	Tone+ vowel	Tone+ consonant
Business communication attributes	Credibility	2.98	3.38	3.70	3.65	2.78	3.33
	Competence	2.84	2.98	3.60	3.83	2.65	2.74
	Friendliness	3.00	3.42	3.45	3.43	2.92	3.24

	Effectiveness	2.34	3.15	3.87	4.08	2.55	2.91
Linguistic attributes	Fluency	3.06	3.57	3.58	4.02	2.75	3.20
	Degree of Foreign accent (reversed)	2.33	2.63	2.58	3.25	1.99	2.32

Table 4. Mean scores for each error type for each of the attributes

Comparing the mean scores for each error type it was clear the segmental error type vowel consistently received worse ratings than the other segmental error type consonant. The same effect could be noticed, albeit a bit less extreme, in the tone+vowel error type compared to the tone+consonant error type. This seemed to suggest that vowel errors had a larger impact on attribute scores than consonantal errors. The differences in the mean scores for the error types tone large and tone small appeared very limited for the business communication attributes. For the linguistic attributes these differences between the means of scores were somewhat larger. This seemed to suggest that large tonal errors did not have a larger impact on business communication attribute scores than small tonal errors, but did have a larger impact on linguistic attribute scores. These mean scores showed a general trend, but further tests had to be conducted to determine if this trend was likely to have been coincidental.

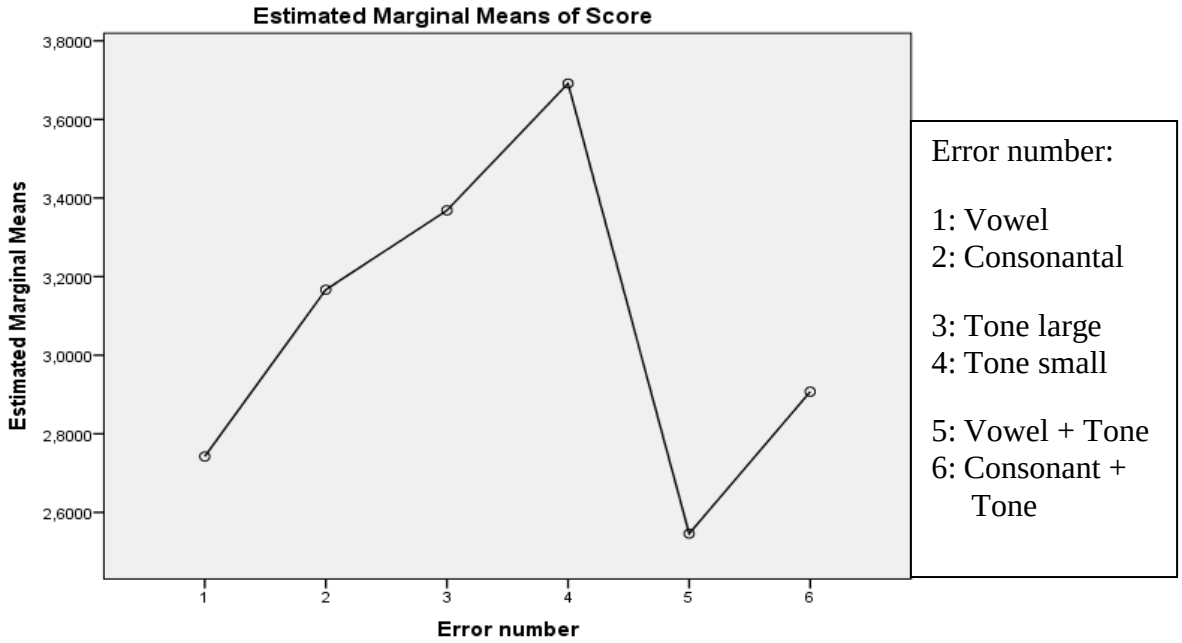
In order to check whether language based attributes were more heavily influenced by pronunciation errors than business communication attributes, a two-way analysis of variance (hereafter ANOVA) was conducted. In this analysis, the dependent variable was each participant's average score for every combination of error type and attribute type, either linguistic or business. The fixed factors were error number (one to six) and attribute number (one to two). For this analysis Levene's test was non-significant. Thus, the assumption of homogeneity of variance was met. Data from all the 20 participants were included into this analysis. The results showed the following:



Graph 1. Difference in means of scores for linguistic and business attributes

This graph shows the main effect for attribute type. The means of scores for business communication attributes were higher than those for linguistic attributes when looking at all error types combined. The analysis showed that this effect is significant ($F(1, 228) = 12.228$, $p = 0.001$).

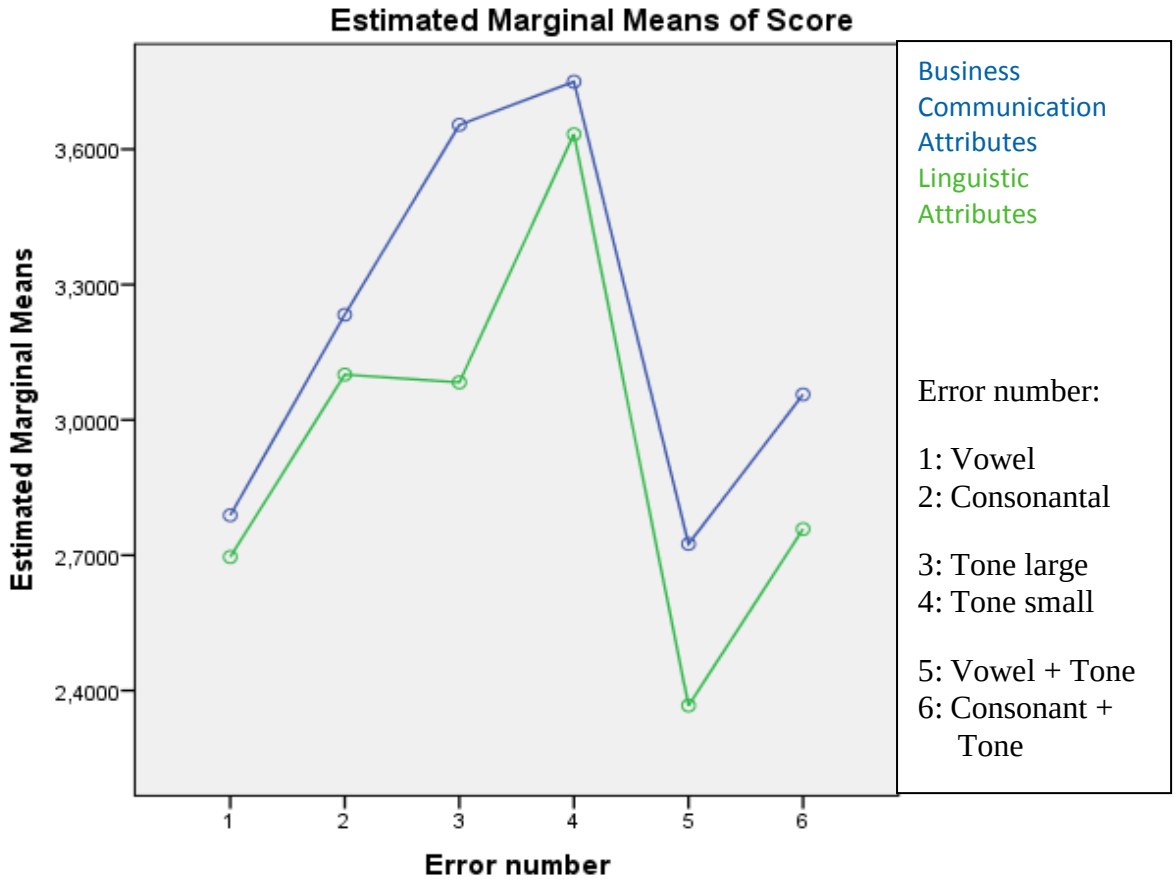
A similar plot was drawn for the effect of different types of error on means of scores:



Graph 2. Difference in means of scores for error types

This graph shows the main effect for error type. The means of scores for stimuli with tonal errors were higher than those for stimuli with segmental errors and stimuli with combination errors scored even worse. The analysis showed that this effect was significant ($F(1, 228) = 21.299$, $p = 0.000$).

Finally, the means of scores for each error type for the two types of attributes was plotted:



Graph 3. Difference in means of scores for error types based on attribute type

This graph shows the interaction of error type and attribute type. There was a clear pattern for linguistic attributes to receive lower ratings than business communication attributes. An interesting break in this pattern appeared to be the mean of scores for stimuli with a large tonal error (error number 3). The mean of scores for linguistic attributes for this error type was much lower than that for business communication attributes.

Although the analyses plotted in graphs 2 and 3 showed a significant effect for error type, they did not specify whether this effect was similar for all attributes individually. The category of combination errors showed a difference in the scores to vowel and consonantal errors similar to those in the category of segmental errors for both attribute categories. For further testing the category of combination errors was removed from the dataset. The reason for this deletion was the fact that the combination error types contained either a combination of vowel error and large tonal error or consonantal error and large tonal error. As such, comparing this category with the segmental and tonal categories would have meant comparing some errors with themselves. This was useful to establish a general trend in the mean scores, but to keep further analysis simple and clear only the means of scores for segmental and tonal errors were compared.

In order to judge the effects of different types of pronunciation errors that were distinguished in this research two different points of view were taken. First of all, the difference in the effects of segmental and tonal errors on each of the attributes was tested. Secondly, the difference in the effects of the two error types in each of these categories on each of the attributes was tested. For the first round of testing a one-way repeated measure ANOVA was conducted on SPSS for each attribute to establish whether the segmental and tonal error categories had significantly different influences on the scores for each attribute. In each analysis the dependent variable was the average score of the attribute at issue and the independent variable was type of pronunciation error. Data from all the 20 participants were included into the analyses. For each of the attributes Mauchly's test was non-significant, thus indicating that the assumption of sphericity was met. Furthermore, for the entire dataset of responses Levene's test was non-significant for all dependent variables. Thus, the assumptions of homogeneity and of homogeneity of covariance matrices were met. In total,

six one-way repeated measures ANOVA were conducted. The results are summarised in Table 5.

Attribute group	Attribute	Significant difference between categories	P-score	Mean score in the 'segmental error' condition	Mean score in the 'tonal error' condition
Business communication attributes	Credibility	YES	.012	3.18	3.68
	Competence	YES	.000	2.91	3.72
	Friendliness	NO	.125	3.21	3.44
	Effectiveness	YES	.000	2.75	3.97
Linguistic attributes	Fluency	YES	.000	3.32	3.80
	Degree of Foreign accent (reversed)	YES	.026	2.48	2.92

Table 5. Results for one-way repeated measures ANOVA for segmental and tonal errors

Table 5 shows that segmental errors were met with lower scores for all attributes than tonal errors for five out of the six attributes. The effect was significant for all these occurrences except when asked about the friendliness of the speaker. Combined with the mean scores from table 4 these results showed that segmental errors did have a consistently and significantly more negative impact on all the attributes than tonal errors.

Finally, tests were conducted to investigate whether the different error types within the same error category had different effects on the scores for each of the attributes. The results of which can be found in this table:

Attribute group	Attribute	Category	Significant difference between subcategories	p-score
Business attributes	Credibility	Segmental	NO	.089
	Credibility	Tonal	NO	.780
	Competence	Segmental	NO	.382
	Competence	Tonal	NO	.265

	Effectiveness	Segmental	YES	.003
	Effectiveness	Tonal	NO	.090
	Friendliness	Segmental	YES	.041
	Friendliness	Tonal	NO	.942
Linguistic attributes	Degree of Foreign accent	Segmental	YES	.047
	Degree of Foreign accent	Tonal	YES	.003
	Fluency	Segmental	YES	.009
	Fluency	Tonal	YES	.002

Table 6. Summary of results from one-way repeated measures ANOVA for subcategories

As shown in table 6 there was a visible difference between the language related attributes and the business communication related attributes. The ANOVA tests only showed a significant main effect for segmental errors on effectiveness and friendliness. The ANOVA tests ran for degree of foreign accent and fluency showed a significant main effect in all four cases. These results showed that vowel errors in 4 out of 6 cases had a consistently and significantly more negative impact on the attribute in question than consonantal errors. The difference between small tonal errors and large tonal errors appeared to only be significant for linguistic attributes. These observations were completely in line with the trend that was visible across the means of scores shown in table 4.

4 Discussion and conclusions

The research questions posed at the end of chapter 1 can now be discussed in the light of the results reported in section 4. The following research questions were raised:

RQ_1: Do errors in speech production influence the ratings for language ability based attributes differently than they do the business communication attributes of an L2 speaker of Chinese?

RQ_2: Do different types of errors influence the perceived business communication attributes of an L2 speaker of Chinese differently?

Concerning RQ1, the mean scores shown in table 3 were significantly lower for linguistic attributes than for business communication attributes for every error type. This effect was especially noticeable for the attribute degree of foreign accent. This trend was visualised in graph 3. A two-way repeated measures ANOVA showed that the difference in the effect that attribute category had on the means of scores was significant.

Concerning RQ2, table 5 showed that tonal errors were repeatedly proven less harmful to the perceived business communication attributes of the speaker than segmental errors. Research by Xi et al. (2010) showed that tonal errors elicited mismatch negativity triggered brain activity, showing categorical perception of lexical tones by native speakers of Chinese (p. 223). It seems unlikely the difference in impact between tonal and segmental errors was caused by any inability to recognise the tonal errors on the side of the participants. A study by Li et al. (2000) showed how speakers of Chinese with a different dialect background varied in their pronunciation of both segmental and tonal speech components (p. 486-488). Since both segmental and tonal errors seemed to trigger a reaction from the participants in the study at hand, it seems the participants for some reason simply favoured tonal errors over segmental errors. It is possible the participants were biased to more easily forgive tonal mistakes,

perhaps because they were aware of the differences between tonal and non-tonal languages. The reason behind this behaviour is unclear, however, and future research might provide more information on this phenomenon.

Table 5 also showed how the ratings for friendliness were not significantly different for tonal and segmental errors. The lack of a significant effects might suggest that perceived friendliness was more dependent on vocal characteristics than non-erroneous speech production.

The difference between the effects of small and large vowel mistakes on the mean scores for business communication attributes was not significant, while the difference between the effects of vowel and consonantal errors was significant in two out of four cases. This suggests that the error types can be arranged in order of negative influence on business communication attributes, vowel errors being the most influential, followed by consonantal errors, followed by large tonal errors and small tonal errors having the smallest influence of the four. In this case it should be noted that the difference between large and small tonal errors was never significant, although the mean scores did show that small tonal errors received better overall ratings in this experiment. This could suggest that learning to correctly pronounce vowels and consonants should be prioritised over learning to correctly pronounce tones. However, the stimuli used in this experiment were common expressions in everyday communication. It is possible the effects of erroneous pronunciation of tone will prove significant for stimuli that do not belong to this category of everyday expressions. Further research could provide more evidence for or against this idea.

There is enough evidence to evaluate the research questions and the hypotheses as posed at the end of chapter one. We hypothesised that:

H_1: Language ability based attributes of an L2 speaker of Chinese are more heavily influenced by speech production errors than perceived business communication attributes.

H_2: Different types of errors have different effects on perceived business communication attributes of an L2 speaker of Chinese.

Evidence was found to support the first hypothesis. The consistent significance of the effect of the different types of errors on the attributes of degree of foreign accent and fluency supports the idea that erroneous speech production leads to a less positive or even more negative view of the speaker's ability to use the language.

Evidence was found to support the second hypothesis. Data indicated that different types of errors showed different effects on different attributes. Tonal errors were significantly less impactful than segmental errors. Vowel errors proved to be significantly more impactful than consonantal errors in half the cases, while small and large tonal errors did not appear to influence the means of scores differently from one another.

To be able to test the hypotheses, the number of factors influencing the scores had to be limited to a minimum. As such, further research could be conducted on the effect of gender and voice characteristics of the speaker or profession of the participants. Other interesting variables could be familiarity of the listeners with Chinese spoken by foreigners or regional differences in the perception of intercultural business related attributes. Furthermore, the data showed an interesting difference in scores for linguistic attributes for small and large tonal errors, which might also be addressed in future research.

The participants in this study gave significantly different ratings for stimuli with tonal errors and stimuli with segmental errors on five out of six attributes. The only attribute not rated significantly different for these two categories was friendliness. This difference might be explained by the fact that friendliness is more judged by first impression whereas competence, effectiveness and credibility might be more influenced by evidence on the speaker's functioning. However, this research did not make this distinction between the business communication attributes. Further research might address this issue.

Despite the many limitations of this research, it does provide interesting implications for second language acquisition and business-focused Chinese language courses in particular. The difference in the effects caused by the different error types implies it might be wise to focus first and foremost on the correct pronunciation of vowels and consonants. The ratings between small and large tonal errors were not significantly different for any of the attributes, which makes the effort of learning how to perfectly use tones in Chinese pronunciation appear overly time-consuming. Further research might focus on determining to what extent the ability to effectively recognise and/or produce tones in L2 Chinese is necessary for communication purposes at different levels of L2 competence.

5 Acknowledgments

I would like to thank Anqi Yang for her advice, expertise and help in finding participants. Also, I would like to thank Zenghui Liu for her advice, expertise, help in translating the instructions and constant moral support throughout this process. My thanks go out to Tom Lentz as well, for helping me find my way around SPSS and statistics in general. Above all, I would like to thank my supervisor Aoju Chen for her time, expertise and seemingly never-ending patience, which have guided me through every step of this learning process. Finally, my thanks go out to all the participants in the Netherlands, Hong Kong and People's Republic of China for their contribution to this research project.

References

Asher, J. J., & Garcia, R. (1969). The Optimal Age to Learn a Foreign Language. *The Modern Language Journal*, 53(5), 334-41.

Bongaerts, T., Planken, B. & Schils, E. (1995). Can Late Learners Attain a Native Accent in a Foreign Language? A Test of the Critical Period Hypothesis. In Singleton, D.M. & Lengyel, Z. (Eds.), *The Age Factor in Second Language Acquisition*, 30-50. Clevedon: Multilingual Matters.

Boyd, S. (2003). Foreign-born Teachers in the Multilingual Classroom in Sweden: The Role of Attitudes to Foreign Accent. *International Journal of Bilingual Education and Bilingualism*, 6(3-4), 283-95.

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y. & Shearman, S. M. (2002). Attitudinal and Affective Response toward Accented English. *Language & Communication*, 22(2), 171-85.

Blumstein, S. E., Alexander, M. P., Ryalls, J. H., Katz, W. & Dworetzky, B. (1987). On the Nature of the Foreign Accent Syndrome: A Case Study. *Brain and Language*, 31(2), 215-44.

Chen, C. C., Yu, K. C. & Miner, J. B. (1997). Motivation to Manage: A Study of Women in Chinese State-Owned Enterprises. *The Journal of Applied Behavioral Science*, 33(2), 160-73.

Chen, J-Y. (1999). The Representation and Processing of Tone in Mandarin Chinese: Evidence from Slips of the Tongue. *Applied Psycholinguistics*, 20, 289-103.

China. Permanent Mission of the People's Republic of China to the United Nations Office at Geneva and Other International Organizations in Switzerland. *Gender Equality and Women's Development in China*. N.p., n.d.

Costa, P. Jr., Terracciano, A. & McCrae, R.R. (2001). Gender Differences in Personality Traits across Cultures: Robust and Surprising Findings. *Journal of Personality and Social Psychology*, 81(2), 322-31.

Elliot, R. A. (1995). Field Independence/dependence, Hemispheric Specialization, and Attitude in Relation to Pronunciation Accuracy in Spanish as a Foreign Language. *The Modern Language Journal*, 79, 356-71.

Fathman, A. (1975) The Relationship Between Age And Second Language Productive Ability. *Language Learning*, 25(2), 245-53.

Flege, J. E. (1988). Factors Affecting Degree of Perceived Foreign Accent in English Sentences. *Journal of the Acoustical Society of America*, 84(1), 70-79.

Flege, J. E. & Eefting, W. (1987). Cross-language Switching in Stop Consonant Perception and Production by Dutch Speakers of English. *Speech Communication*, 6(3), 185-202.

Flege, J. E. & Fletcher, K. L. (1992). Talker and Listener Effects on Degree of Perceived Foreign Accent. *The Journal of the Acoustical Society of America*, 91(1), 370.

Flege, J.E., Munro, M. J. & MacKay, I. R.A. (1995). Factors Affecting Strength of Perceived Foreign Accent in a Second Language. *The Journal of the Acoustical Society of America*, 97(5), 3125-134.

Flege, J. E., Frieda, E. M. & Nozawa, T. (1997). Amount of Native-language (L1) Use Affects the Pronunciation of an L2. *Journal of Phonetics*, 25, 169-186.

Flege, J.E., Yeni-Komshian, G. H. & Liu, S. (1999). Age Constraints on Second-Language Acquisition. *Journal of Memory and Language*, 41(1), 78-104.

Friederici, A. D. (2002). Towards a Neural Basis of Auditory Sentence Processing. *Trends in Cognitive Sciences*, 6(2), 78-84.

Gluszek, A., & J. F. Dovidio. (2010). The Way They Speak: A Social Psychological Perspective on the Stigma of Nonnative Accents in Communication. *Personality and Social Psychology Review*, 14(2), 214-37.

Grover, S., Jamieson, D. G. & Dobrolovsky, M. B. (1987). Intonation in English, French and German: Perception and Production. *Language and Speech*, 30, 277-96.

Guion, S. G., Flege, J. E. & Loftin, J.D. (1999). The Effect of L1 Use on Foreign Accent Ratings in Quichua-Spanish Bilinguals. *Proceedings of the 14th International Congress of Phonetic Sciences*, 2, 1471-474.

Guion, S. G., Flege, J. E. & Loftin, J.D. (2000). The Effect of L1 Use on Pronunciation in Quichua–Spanish Bilinguals. *Journal of Phonetics*, 28(1), 27-42.

Gurd, J.M., Bessell, N.J., Bladon, R.A.W. & Bamford, J.M. (1988). A Case of Foreign Accent Syndrome, with Follow-up Clinical, Neuropsychological and Phonetic Descriptions. *Neuropsychologia*, 26(2), 237-51.

Ingram, J. C., McCormack, P. F. & Kennedy, M. (1992). Phonetic Analysis of a Case of Foreign Accent Syndrome. *Journal of Phonetics*, 20(4), 457-74.

He, D. (2014). Career Women in China Take Charge. *China Daily USA*, n.p., 24 March 2014.

HKU Manual for Praat. *Chapter 1 Introduction*. Hong Kong University, n.p., n.d., 14 July 2014.

Hosoda, M., Stone-Romero, E. F. & Walter, J. N. (2007). Listeners' Cognitive And Affective Reactions To English Speakers With Standard American English And Asian Accents. *Perceptual and Motor Skills*, 104(1), 307-26.

Kurowski, K.M., Blumstein, S. E. & Alexander, M. (1996). The Foreign Accent Syndrome: A Reconsideration. *Brain and Language*, 54(1), 1-25.

Li, A., Zheng, F., Byrne, W., Fung, P., Kamm, T., Liu, Y., ... & Chen, X. (2000). CASS: a Phonetically Transcribed Corpus of Mandarin Spontaneous Speech. *Interspeech*, 485-488.

Lindemann, S. (2003). Koreans, Chinese or Indians? Attitudes and Ideologies about Non-native English Speakers in the United States. *Journal of Sociolinguistics*, 7(3), 348-64.

Magen, H. S. (1998). The Perception of Foreign-accented Speech. *Journal of Phonetics*, 26(4), 381-400.

Mareüil, P. B. de & Vieru-Dimulescu, B. (2006). The Contribution of Prosody to the Perception of Foreign Accent. *Phonetica*, 63(4), 247-67.

Missaglia, F. (1999). Contrastive Prosody in SLA: An Empirical Study with Adult Italian Learners of German. *Proceedings of the 14th International Congress of Phonetic Sciences*, 1, 551-54.

Moyer, A. (1999). Ultimate Attainment In L2 Phonology. *Studies in Second Language Acquisition*, 21(1), 81-108.

Munro, M. J. (1995). Nonsegmental Factors in Foreign Accent. *Studies in Second Language Acquisition*, 17(1), 17-34.

Munro, M. J., Derwing, T. M. & Flege, J. E. (1999). Canadians in Alabama: A Perceptual Study of Dialect Acquisition in Adults. *Journal of Phonetics*, 27(4), 385-403.

Olson, L. L. & Samuels, J. S. (1973). The Relationship between Age and Accuracy of Foreign Language Pronunciation. *Journal of Educational Research*, 66, 263-68.

Oyama, S. (1976). A Sensitive Period for the Acquisition of a Nonnative Phonological System. *Journal of Psycholinguistic Research*, 5(3), 261-83.

Patkowski, M. S. (1990). Age and Accent in a Second Language: A Reply to James Emil Flege. *Applied Linguistics*, 11(1), 73-89.

Pennington, M. C. & Ellis, N. C. (2000). Cantonese Speakers' Memory for English Sentences with Prosodic Cues. *The Modern Language Journal*, 84(3), 372-89.

Piper, T. & Cansin, D. (1988). Factors Influencing The Foreign Accent. *Canadian Modern Language Review*, 44(2), 334-42.

Piske, T., Mackay, I. R. A. & Flege, J. E. (2001). Factors Affecting Degree of Foreign Accent in an L2: A Review. *Journal of Phonetics*, 29(2), 191-215.

Purcell, E. T. & Suter, R. W. (1980). Predictors Of Pronunciation Accuracy: A Reexamination. *Language Learning*, 30(2), 271-87.

Rubin, D. L. & Smith, K. A. (1990). Effects of Accent, Ethnicity, and Lecture Topic on Undergraduates' Perceptions of Nonnative English-speaking Teaching Assistants. *International Journal of Intercultural Relations*, 14(3), 337-53.

Rubin, D. L., Healy, P., Gardiner, T. C., Zath, R. C. & Moore, C. P. (1997). Nonnative Physicians as Message Sources: Effects of Accent and Ethnicity on Patients' Responses to AIDS Prevention Counseling. *Health Communication*, 9(4), 351-68.

Seliger, H. W., Krashen, S. D. & Ladefoged, P. (1975). Maturational Constraints in the Acquisition of Second Language Accent. *Language Sciences*, 36, 20-22.

Snow, C. E. & Hoefnagel-Hohle, M. (1977). Age Differences in the Pronunciation of Foreign Sounds. *Language and Speech*, 20, 357-65.

Suter, R. W. (1976) Predictors Of Pronunciation Accuracy In Second Language Learning. *Language Learning*, 26(2), 233-53.

Tahta, S., Wood, M. & Loewenthal, K. (1981). Foreign Accents: Factors Relating to Transfer of Accent from the First Language to a Second Language. *Language and Speech*, 24(3), 265-72.

Tajima, K., Port, R. & Dalby, J. (1997). Effects of Temporal Correction on Intelligibility of Foreign-accented English. *Journal of Phonetics*, 25(1), 1-24.

Thompson, I. Foreign Accents Revisited: The English Pronunciation of Russian Immigrants. *Language Learning*, 41(2), 177-204.

Tsalikis, J., Ortiz-Buonafina, M., & Latour, M. S. (1992). The Role of Accent on the Credibility and Effectiveness Of the International Business Person: The Case of Guatemala. *International Marketing Review*, 9(4), 57-72.

Tsukada, K. (2005). Cross-language Speech Perception of Final Stops by Australian-English, Japanese and Thai Listeners. *Proceedings of the ISCA Workshop on Plasticity in Speech Perception '05*, 244– 247.

Tsukada, K., Birdsong, D., Mack, M., Sung, H., Bialystok, E. & Flege, J. E. (2004). Release Bursts in English Word-final Voiceless Stops Produced by Native English and Korean Adults and Children. *Phonetica*, 61, 67–83.

Wayland, R. (1997). Non-native Production of Thai: Acoustic Measurements and Accentedness Rating. *Applied Linguistics*, 18, 345-73.

Xi, J., Zhang, L., Shu, H., Zhang, Y. and LI, P. (2010). Categorical Perception of Lexical Tones in Chinese Revealed by Mismatch Negativity. *Neuroscience*, 170, 223-231.

Zahn, C. J. & Hopper, R. (1985). Measuring Language Attitudes: The Speech Evaluation Instrument. *Journal of Language and Social Psychology*, 4(2), 113-23.

Appendix A – Sentence stress of stimuli

	Speaker 1	Speaker 2	Speaker 3
1	别介意	别介意	别介意
2	不用谢	不用谢	不用谢
3	不客气	不客气	不客气
4	大家好	大家好	大家好
5	对不起	对不起	对不起
6	礼拜五	礼拜五	礼拜五
7	恭喜你	恭喜你	恭喜你
8	回头见	回头见	回头见
9	我很好	我很好	我很好
10	没问题	没问题	没问题
11	没关系	没关系	没关系
12	明天见	明天见	明天见
13	早上好	早上好	早上好
14	您先请	您先请	您先请
15	您贵姓	您贵姓	您贵姓
16	请吃吧	请吃吧	请吃吧
17	太好了	太好了	太好了
18	晚上好	晚上好	晚上好
Add.	请签名	请签名	请签名

Appendix B – Complete list of stimuli per category

Segmental

Vowel errors

1	别介意	biu	2	jie	4	yi	4
2	不用谢	bi	2	yong	4	xie	4
3	不客气	ba	2	ke	4	qi	0
4	大家好	da	4	jia	1	hou	3
5	对不起	du	4	bu	4	qi	3
6	礼拜五	li	3	bai	4	wo	3
7	恭喜你	ging	1	xi	3	ni	3
8	回头见	hui	2	tou	2	jun	4
9	我很好	wo	3	hen	2	hu	3
10	没问题	mai	2	wen	4	ti	2
11	没关系	mu	2	guan	1	xi	0
12	明天见	ming	2	tian	1	jiu	4
13	早上好	zao	3	shang	4	huo	3
14	您先请	nan	2	xian	1	qing	3
15	请签名	qiu	3	qian	1	ming	2
16	请吃吧	qing	3	chu	1	ba	0
17	太好了	tou	4	hao	3	le	0
18	晚上好	wan	3	shang	4	ho	3

Consonantal errors

1	别介意	pie	2	jie	4	yi	4
2	不用谢	mu	2	yong	4	xie	4
3	不客气	fu	2	ke	4	qi	0
4	大家好	da	4	jia	1	gao	3
5	对不起	tui	4	bu	4	qi	3
6	礼拜五	li	3	bai	4	lu	3
7	恭喜你	kong	1	xi	3	ni	3
8	回头见	hui	2	tou	2	qian	4
9	我很好	wo	3	hen	2	kao	3
10	没问题	nei	2	wen	4	ti	2
11	没关系	lei	2	guan	1	xi	0
12	明天见	ming	2	tian	1	xian	4
13	早上好	zao	3	shang	4	fao	3
14	您先请	min	2	xian	1	qing	3
15	请签名	jing	3	qian	1	ming	2
16	请吃吧	qing	3	zhi	1	ba	0
17	太好了	dai	4	hao	3	le	0
18	晚上好	wan	3	shang	4	lao	3

Tonal errors

Relatively small tonal errors

1	别介意	bie	3	jie	4	yi	4
2	不用谢	bu	3	yong	4	xie	4
3	不客气	bu	3	ke	4	qi	0
4	大家好	da	4	jia	1	hao	2
5	对不起	dui	3	bu	4	qi	3
6	礼拜五	li	3	bai	4	wu	2
7	恭喜你	gong	2	xi	3	ni	3
8	回头见	hui	2	tou	2	jian	3
9	我很好	wo	3	hen	2	hao	2
10	没问题	mei	3	wen	4	ti	2
11	没关系	mei	3	guan	1	xi	0
12	明天见	ming	2	tian	1	jian	3
13	早上好	zao	3	shang	4	hao	4
14	您先请	nin	3	xian	1	qing	3
15	请签名	qing	2	qian	1	ming	2
16	请吃吧	qing	3	chi	2	ba	0
17	太好了	tai	3	hao	3	le	0
18	晚上好	wan	3	shang	4	hao	2

Relatively large tonal errors

1	别介意	bie	4	jie	4	yi	4
2	不用谢	bu	4	yong	4	xie	4
3	不客气	bu	1	ke	4	qi	0
4	大家好	da	4	jia	1	hao	1
5	对不起	dui	2	bu	4	qi	3
6	礼拜五	li	3	bai	4	wu	1
7	恭喜你	gong	4	xi	3	ni	3
8	回头见	hui	2	tou	2	jian	1
9	我很好	wo	3	hen	2	hao	4
10	没问题	mei	4	wen	4	ti	2
11	没关系	mei	1	guan	1	xi	0
12	明天见	ming	2	tian	1	jian	1
13	早上好	zao	3	shang	4	hao	1
14	您先请	nin	4	xian	1	qing	3
15	请签名	qing	4	qian	1	ming	2
16	请吃吧	qing	3	chi	4	ba	0
17	太好了	tai	1	hao	3	le	0
18	晚上好	wan	3	shang	4	hao	1

Combination errors

Tonal-vowel errors

1	别介意	biu	4	jie	4	yi	4
2	不用谢	bi	4	yong	4	xie	4
3	不客气	ba	1	ke	4	qi	0
4	大家好	da	4	jia	1	hou	1
5	对不起	du	2	bu	4	qi	3
6	礼拜五	li	3	bai	4	wo	1
7	恭喜你	ging	4	xi	3	ni	3
8	回头见	hui	2	tou	2	jun	1
9	我很好	wo	3	hen	2	hu	4
10	没问题	mai	4	wen	4	ti	2
11	没关系	mu	1	guan	1	xi	0
12	明天见	ming	2	tian	1	jiu	1
13	早上好	zao	3	shang	4	huo	1
14	您先请	nan	4	xian	1	qing	3
15	请签名	qiu	4	qian	1	ming	2
16	请吃吧	qing	3	chu	4	ba	0
17	太好了	tou	1	hao	3	le	0
18	晚上好	wan	3	shang	4	ho	1

Tonal-consonantal errors

1	别介意	pie	4	jie	4	yi	4
2	不用谢	mu	4	yong	4	xie	4
3	不客气	fu	1	ke	4	qi	0
4	大家好	da	4	jia	1	gao	1
5	对不起	tui	2	bu	4	qi	3
6	礼拜五	li	3	bai	4	lu	1
7	恭喜你	kong	4	xi	3	ni	3
8	回头见	hui	2	tou	2	qian	1
9	我很好	wo	3	hen	2	kao	4
10	没问题	nei	4	wen	4	ti	2
11	没关系	lei	1	guan	1	xi	0
12	明天见	ming	2	tian	1	xian	1
13	早上好	zao	3	shang	4	fao	1
14	您先请	min	4	xian	1	qing	3
15	请签名	jing	4	qian	1	ming	2
16	请吃吧	qing	3	zhi	4	ba	0
17	太好了	dai	1	hao	3	le	0
18	晚上好	wan	3	shang	4	lao	1

Appendix C – Instructions

您好！

感谢您参与本语言学实验。在接下来的10分钟里，你会听到荷兰语为母语的人说几句汉语。这几位说话人在中国从事商贸。但因为他们是外国人，他们的汉语不完美。

实验包括6个页面，每个页面上，您会被要求评估一个“特征”。每一页以哔声开始，开始后一共是18句话供评估。

在页面的左侧，您会看到他们想说的话。每次您需要给他们从1（最低）到5（最高级别）评分。评定准则是根据在列表中显示的“特征”。您不需要仔细思考，只是根据您的第一印象，直接选择相应的选项。

这个实验将持续10分钟。

当您准备开始时，请播放example.wav, 用耳机听。

例如：**害羞**

1(不害羞)	2	3	4	5(很害羞)
☐	☐	☐	☐	☐

他们说每一话，请只选中其中一个选项。

实验结束后,请您将此完成的文件发送至 chinese.experiment@gmail.com。将文件保存为:lastname_list5.docx. 比如说:fang_list5.docx。

当您准备好开始的时候，请来播放experiment.wav。

非常感谢您的参与！

年龄：　　性别：　　职业:

您的方言:

普通话水平：高(　)　比较高(　)　一般(　)　比较低(　)　很低(　)

Appendix D – Instructions (English Translation)

Hello !

Thank you for participating in this linguistic experiment. Over the next 10 minutes, you will hear Dutch-speaking people say a few words in Chinese. These speakers are in China for business. But, because they are foreigners, their Chinese is not perfect.

The experiment consists of six pages. On each page you will be asked to evaluate an "attribute." Each page starts with a beep, after which you will hear a total of 18 sentences to evaluate.

On the left side of the page, you will see they mean to say. For every sentence you need to rate them from 1 (lowest score) to 5 (highest score). The evaluation criteria are the "attributes" displayed on top of the list. You do not need to think too carefully about your answer. Just select the score based on your first impression.

This experiment will last 10 minutes.

When you are ready to start, play example.wav, listening with headphones.

Example: Shy

1(not shy)	2	3	4	5(very shy)
☐	☐	☐	☐	☐

For every sentence, please select only one option.

After finishing the experiment, please send this completed document to chinese.experiment@gmail.com. Save the file as: lastname_list5.docx. Example: Fang_list5.docx.

When you are ready to start, please play experiment.wav.

Thank you very much for your participation!

Age: Gender: Occupation: Your dialect:

Mandarin level: High () Relatively high () General ()

Relatively low () Very low()