

THE IMPACT OF AUDIO-VISUAL MEDIA ON ENGLISH PHONOLOGICAL ACQUISITION AND DEFOSSILIZATION BY LEARNERS IN CHINA

CHEN ZHANG

(University of Sheffield)

Abstract

This paper investigates the effectiveness of ‘audio-visual media’ as language input for English phonological acquisition and defossilization among Chinese-speaking learners of English in the Chinese context. Supported by the exemplar theory (Johnson 1996, Pierrehumbert 2001), the necessity of native-like phonological input is greatly emphasized. However, in the current Chinese context, the phonological input is generally non native-like. Also, given the fact that adult learners in China have a fossilized English phonology, it is important to find an alternative form of input. Audio-visual media have the potential to become the alternative, on account of being a provider of native-like English phonological characteristics and a provider of visual cues. The impact of audio-visual media was tested in a natural self-learning setting by implementing a pre-test and a post-test on vowel quality. Eighteen Chinese English learners participated. Over a half-month period, nine subjects received input from an English TV series, while the other nine did not receive any input. The results suggest that audio-visual media can have a positive impact on L2 (i.e. second language) phonological acquisition. Audio-visual media can also help Chinese L2 English learners break the deadlock in fossilized L2 phonology, and thus push forward the development of their interlanguage phonology.

1. Introduction

The present preliminary study aims to test the impact of audio-visual media as language input for English phonological defossilization by L2 English learners in China. This project is motivated by the need for native-like phonological input in second language research and the current situation of English phonological input in China. Native-like input is a necessary component in L2 phonological acquisition (see Section 2.1). However, considering the nature of L2 English education in China, the English input that Chinese-speaking learners of English receive is non native-like. The L2 learner’s interlanguage has obvious characteristics from Chinese transfer and has a tendency to be fossilized (Xiang 2009, Chen 2011). The non native-like phonological input can aggravate learners’ phonological fossilization and impede the process of defossilization. Considering this problem, it is essential that we find alternative L2 English input to better facilitate phonological acquisition and defossilization processes among Chinese-speaking learners of English (Zhang Y 2013). Due to the fact that audio-visual media simultaneously provide both unmodified target language input in the sound tracks and moving visual images, it would appear to be the best alternative form of input to traditional classroom teaching in the Chinese context, and this is tested in this study.

The impact of audio-visual media has gained increasing attention in the past few years. However, it is still a new research topic which needs to be comprehensively investigated. Few studies focus on testing the impact of audio-visual media on L2 phonological acquisition. Therefore, to fill this research gap, an intervention experiment, which consists of a pre-test and an immediate post-test, was conducted to test the effectiveness of audio-visual media. The main research question addressed was ‘Can audio-visual media have a positive impact on L2 acquisition of phonology?’

This paper has six sections. In Section 2, the research background, including the necessity of native-like language input in L2 phonological acquisition, is presented. The problem of English input in China and phonological fossilization is also considered alongside the alternative L2 input of audio-visual media. Section 3 introduces the research methodology of the experiment. Next, in Section 4, the data analyses and results are outlined. The discussion is presented in Section 5 and this paper is concluded in Section 6.

2. Research background

2.1. Native-like input in L2 phonological acquisition

The past few decades have witnessed a great development of language input research in the field of second language acquisition (Gass 1997, Ellis, R 1985, Ellis, R 1994, Krashen 1982, Krashen 1985, VanPatten 2004, Carroll 2007, Schwartz 1993, Rast 2008, Ellis N 2009, Piske & Young-Scholten 2009). The necessity of language input has been commonly demonstrated and acknowledged (Gass 1997). Learners cannot acquire the target language in a vacuum. Specifically, phonological input, as an integral part of language input, is undoubtedly essential in phonological acquisition. Flege (2009: 175) defines phonological input as ‘all L2 vocal utterances the learner has heard and comprehended’. Both native-like and non native-like L2 utterances are included in this definition. Comparing these two kinds of utterances, Jenkins (2000: 207) claimed that native-like utterances are unnecessary, with regard to intelligibility, when using English to communicate internationally. Thus the demand for native-like input in L2 phonological acquisition is reduced. However, as far as I am concerned, native-like input is far more important than non native-like input in L2 phonological acquisition. This viewpoint is supported by the exemplar theory (Johnson 1996, Pierrehumbert 2001).

The notion of exemplar was first introduced from psychological models for the study of speech sounds by Johnson (1996). This exemplar model works for both phonological perception and production. In this model, every phonological category is represented by ‘a large cloud of remembered tokens of that category’ in the mind (Pierrehumbert 2001: 140). The categories are systematically organized. The sounds which are perceived as similar phonological units are stored closely, while the dissimilar units are stored far away from each other. If new phonological tokens are perceived, they will be categorized into the ‘clouds’ according to their similarity to the existing tokens. If the ‘cloud’ has more numerous and activated exemplars, it will have advantages over the competition (Pierrehumbert 2001: 142). Different ‘clouds’ of sound units which are perceived from the linguistic input establish the recognition of the phonology of the target language. Phonological input is stored as exemplars and production originates from them. In exemplar theory the authenticity of phonological input is of great significance. Only if this is native-like can L2 learners locate the ‘clouds’ in the right place and thus facilitate their acquisition process in both perception and production.

From the above explanation, it is obvious that native-like input is of great necessity in L2 phonological acquisition. Therefore, phonological input should involve as much native-like input as possible.

2.2. The main problem with L2 English input in China and phonological fossilization

There is seemingly a strong desire among the Chinese public to acquire a foreign language proficiently for the purpose of worldwide political, economic and cultural communication. Governmental policies also encourage the public to learn a foreign language (The Ministry of Education of P. R. China 2011). In China, English has an unshakable status, and is considered as a very important worldwide communication tool. There has been an upsurge of enthusiasm towards English language learning. Though it appears that the whole of Chinese society realizes

the importance of English education and seems to be active in popularizing English among the public, there are admittedly still many problems in current English language teaching and learning. Among the problems, the lack of native-like phonological input is the most prominent one.

In the Chinese context, the phonological input generally comes from ‘teacher-talk’ in classrooms. It is the main source of English input that L1 (i.e. first language) Chinese learners can receive. According to Li & Wang (2006), more than eighty per cent of English input that the learners receive comes from the English language classroom. Liu (2010) obtained a similar result on the percentage of classroom input contribution (seventy-eight per cent) in the overall L2 English input in China. ‘Teacher-talk’ plays an important role in L2 target language learning. However, as is mentioned in Young-Scholten’s work (1995: 111), ‘teacher-talk’ is ‘typically L1 accented’ in foreign language classrooms. A similar problem is also found in Chinese L2 English classrooms, and it appears to be the main notable hindrance in phonological acquisition. The overwhelming majority of the language teachers in China are native speakers of Chinese but not native speakers of English. The speech that they deliver in classroom teaching is not native-like English, especially in terms of phonology (Jia 2011, Li R 2009). The English teachers in China are also English learners themselves. Most of them have not had the opportunity to go to English speaking countries (Li X 2007, Deng 2007, Xu & Guo 2009, Zhang J 2009, Chai & Lu 2009, Tu 2012) and they rarely interact with native speakers of English. Therefore, their English phonology is far from native-like, and their speech still carries very obvious interlanguage characteristics which come from Chinese phonology and individual dialectal accents (Yang & Ren 2009). Another problem which exists in ‘teacher-talk’ in the Chinese classroom is that the speech that teachers deliver is exaggeratedly modified. Adjustments frequently occur in pronunciation in language classrooms (Henzl 1979). In the Chinese context, teachers usually slow down the speech speed, use long pauses when waiting for a reaction from students, and accentuate certain words in the speech for emphasis (Li H 2011). These phonological modifications can potentially make the input much more comprehensible for L2 English learners (Loschky 1994, Krashen 2003) and might be beneficial for learners when understanding the teacher’s talk. However, they do not benefit the learners in phonological acquisition. In a word, ‘teacher-talk’ in China fails to present native-like English phonological characteristics. Thus, it is inefficient in facilitating English phonological acquisition. Given the fact that non-native ‘teacher-talk’ is the main source of phonological input in China, the lack of native-like phonological input is widespread.

In China, many English learners have a fossilized English phonology due to the lack of native-like phonological input. This phenomenon can be explained from the two measures of L2 phonological input proposed by Moyer (2009), that is the ‘age of onset’ and the ‘length of residence’ respectively. With respect to ‘age of onset’, it refers to the age when a L2 learner starts learning the target language. In Moyer’s claim (2009), early ‘age of onset’ is beneficial for L2 phonological acquisition. This viewpoint is supported by the ‘critical period’ hypothesis in L2 phonological acquisition (Scovel 1998, 2000), whereby puberty is the dividing point for speech learning ability. Before puberty, language learners have strong speech learning abilities. In contrast, after puberty, the speech learning ability declines due to neural maturation. If learners receive exposure to native-like L2 phonological input before puberty, they can possibly acquire the L2 phonology successfully. In China, English learners start learning English in the classroom from primary school at the age of six (The Ministry of Education of P. R. China 2011). In other words, they do have an early ‘age of onset’. However, this advantageous effect is not manifested in their English phonological acquisition, because the phonological input they receive before puberty is generally non native-like. Their knowledge of English phonology obtained from ‘teacher-talk’ has traits of Chinese phonology and dialectal accent, and the learner’s interlanguage is obviously influenced by L1 transfer. It is argued that L1 transfer is a

prominent affecting factor in L2 phonological acquisition and production (Hansen 2006, cf. Altenberg & Vago 1987, Benson 1988). Thus, the L1 Chinese learner's English phonology does not develop much before puberty. Considering the fact that exposure to non-native input is continuously given after the critical period, when their speech learning ability has declined, the L2 learner's English phonology is highly likely to be fossilized. In terms of 'length of residence', most L2 English learners in China do not have the experience of living in English speaking countries to receive native-like phonological input. This is restricted for various social and economic reasons. By 2010 in China, there were 400 million L2 English learners. It is expected that there will be a billion English learners by 2020 (Cao 2012). Given the huge number of Chinese speaking learners of English, it is impossible that many will have the opportunity to go abroad and become immersed in English speaking contexts. Therefore, generally speaking, L2 English learners in China have no 'length of residence'. Without 'length of residence' in English speaking countries, learners cannot be exposed to native-like English phonology, and thus cannot benefit from it to reorganize the exemplar 'clouds' and defossilize their English phonology.

Considering the lack of native-like phonological input and learner fossilization in L2 English phonology, it is desirable to find an alternative form of input which contains unmodified native-like English to facilitate L2 acquisition and defossilization. With the development of the internet and the opening-up of the media policy in China, audio-visual media in English are now widely accessible to English learners in China. Audio-visual media contain a large amount of unmodified native-like phonological input, and so it is an ideal alternative to 'teacher-talk'.

2.3. Audio-visual media as alternative phonological input

As language input, audio-visual media has several characteristics. Firstly, it mostly contains rich phonological characteristics spoken by native speakers of English. The quantity of audio-visual media is abundant and is available to the majority of English learners in China. The native-like input from audio-visual media has the potential to defossilize the L2 learner's English phonology. Secondly, audio-visual media simultaneously combines moving visual images with a sound-track. This combination makes the language input much easier for the learners to comprehend and process, because non-linguistic visual cues make the input more noticeable (Plass & Jones 2005: 472). Visual cues from audio-visual media can also help L2 learners recognise the exemplars and process the speech provided in the input. The learners can receive the input, including, for example, information about the lip contour, mouth location, mouth scale and orientations from the actor and the actress. Fanelli et al. (2009) affirm the role of audio-visual fusion in speech processing. They argue that video information provided together with the audio can significantly promote the accuracy of recognition. This viewpoint has also been approved by Macaluso et al. (2004) and Niyogi et al. (1999). Based on these characteristics, it is promising that audio-visual media can positively influence L2 English phonological acquisition by L1 adult Chinese learners. Besides these characteristics, audio-visual media can also convey cultural information, which is interesting to explore. However, since it is outside the research scope of the current study, cultural information is not discussed in this paper.

Past studies on the impact of audio-visual media mainly focus on other aspects of L2 acquisition, such as incidental vocabulary learning (Neuman & Koskinen 1992, Yuksel & Tanriverdi 2009, Karakas & Saricoban 2012). Just a few studies have been carried out on L2 phonological acquisition. From the current literature, Davis and Kim (2001) examine the impact of visual speech on the acquisition of L2 Korean by ten native-speakers of English. They compared the effectiveness of audio with visual speech (i.e. video) information (e.g. mouth

movements, lip contours) and the effectiveness of audio without visual speech information. The results revealed that the fusion of audio-visual cues is beneficial when English speakers learn Korean. Similar results were found in Ortega-Llebaria et al. (2001). In their study, thirty-six Spanish learners of English significantly improved in English phonology with the help of audio-visual (or video) stimuli. However, evidence demonstrating the impact of audio-visual media in the Chinese context is rare. Therefore, the current study contributes to the literature in this field and provides pedagogical implications.

Based on the characteristics and implications of previous studies, it is hypothesized that audio-visual media can have a positive impact on L2 English phonological acquisition by L1 Chinese learners.

3. Research methodology

3.1. Experimental focus and phonological language background

In the experiment, the focus was on aspects of vowel quality including monophthongs and diphthongs. The corresponding phonological language background information is presented as follows.

There are five vowels in Chinese, namely /i/, /y/, /u/, /ə/, /a/. When high vowels occur before another vowel, they behave as glides /j, ɥ, w/. /i/ and /u/ can also follow a non-high vowel to form a diphthong. For instance, /ai/ and /ou/ are diphthongs which are commonly found. However, compared to English, the Chinese diphthongs are less diphthong-like, based on the degree of Formant 1 (F1) and Formant 2 (F2) changes (Zheng & Zheng 2011).

It is also worth noting that there are language-specific phonological constraints in different languages (Klima & Bellugi 1979). There are several combinations of consonants and vowels which exist in English. However, these combinations are impossible in Chinese. These English-specific phonological constraints may be difficult for L1 Chinese-speaking learners. For instance, /hi:/ is a commonly seen consonant-vowel combination in English. However, it is impossible in Chinese. Chinese speaking learners might have difficulties in pronouncing these combinations without exposure to native-like English input. Thus, they might possibly refer to their L1 knowledge of phonology and choose the closest pronunciation (e.g. pronounce /hi:/ as /hei/). Because of this, there might be a chance that the degree of F1 and F2 changes of these vowels in the learner's interlanguage is higher than it should be in English.

3.2. Research hypotheses and predictions

Given the above discussion of audio-visual media as L2 input, the experimental focus and the nature of learners' L1 phonology, the research hypotheses of the current study are as follows:

- 1) Audio-visual media can have a positive impact on English phonological acquisition and defossilization of monophthongs.
- 2) Audio-visual media can have a positive impact on English phonological acquisition and defossilization of diphthongs.

Given the nature of audio-visual media, it is predicted that it may positively influence the L1 Chinese learners' pronunciation and defossilization of monophthongs and diphthongs.

3.3. Research subjects

Twenty-one Chinese L2 English learners were selected to participate in the experiment. After the pre-test, three participants were screened out, on the grounds that they watched films and

TV series for relatively greater amounts of time than the other subjects. There remained eighteen participants who had similar linguistic and educational backgrounds. At the time of testing, the average age of the participants was twenty-two. They had their English education in an EFL context from the age of six and they had been exposed to 'teacher-talk' in L2 English classrooms for about 16 years. All the participants were in China at the time of testing, none of the subjects had been to an English-speaking country and their L2 English input was mainly from classroom instruction.

3.4. Research materials

In this study, Received Pronunciation (RP) was used as the criteria of measurement. By definition, RP refers to the standard accent spoken in the United Kingdom (Wells 1982). When choosing the input stimuli and the baseline material, the main concern was whether it contained RP. In terms of the input stimuli, a comic TV series made by BBC Two in 2009 called *Miranda* was chosen. The main actress, whose name is Miranda Hart, speaks RP (Another Tongue 2015). The first two seasons were adopted in the experiment. Each season included six episodes and each episode lasts for about thirty minutes. The total input time of treatment was six hours. The chosen TV series provided sufficient consistent phonological input in stimuli during the given period and kept the participants interested in the experiment. Also, the TV series provided contained visual components, such as lip contours, facial expressions and body gestures from the actor and the actress. Considering the negative influence that orthographic input might have in L2 phonological acquisition (Young-Scholten 1995, Bassetti 2009), English subtitles were not included in the given TV series. With regard to the baseline material, since the Queen's English shows strong characteristics of RP (Cruttenden 2014: 79), the *Queen's address to Parliament during the Diamond Jubilee* was chosen.

3.5. Research procedures

This study was carried out in a natural self-learning setting. This intervention study consisted of a pre-test and an immediate post-test. A delayed post-test was considered when designing the experiment. However due to the limitations of the research time scale, it was not conducted. In the pre-test stage, all eighteen participants were required to record their reading of the Queen's address. After the recordings were collected, the participants were randomly allocated to either the experimental group (N=9) or the control group (N=9). During the input period of two weeks, the experimental group was asked to watch the given TV series, whereas the control group was not. The experimental group watched the TV series (two seasons) just once via the internet on home computers individually. They were told to watch the given TV series at their own pace and report to the researcher. The control group had no exposure to English input during the input period. Meanwhile, none of the participants in either the experimental group or the control group received any classroom instruction in English or other types of audio-visual media input. Immediately after the input period, all the participants were required to record their reading of the Queen's address again. They were not given time for pronunciation rehearsal. The recordings collected in the pre-test and the immediate post-test are the main data from the experiment. Table 1 below presents the general research procedures.

Table 1. Research procedure

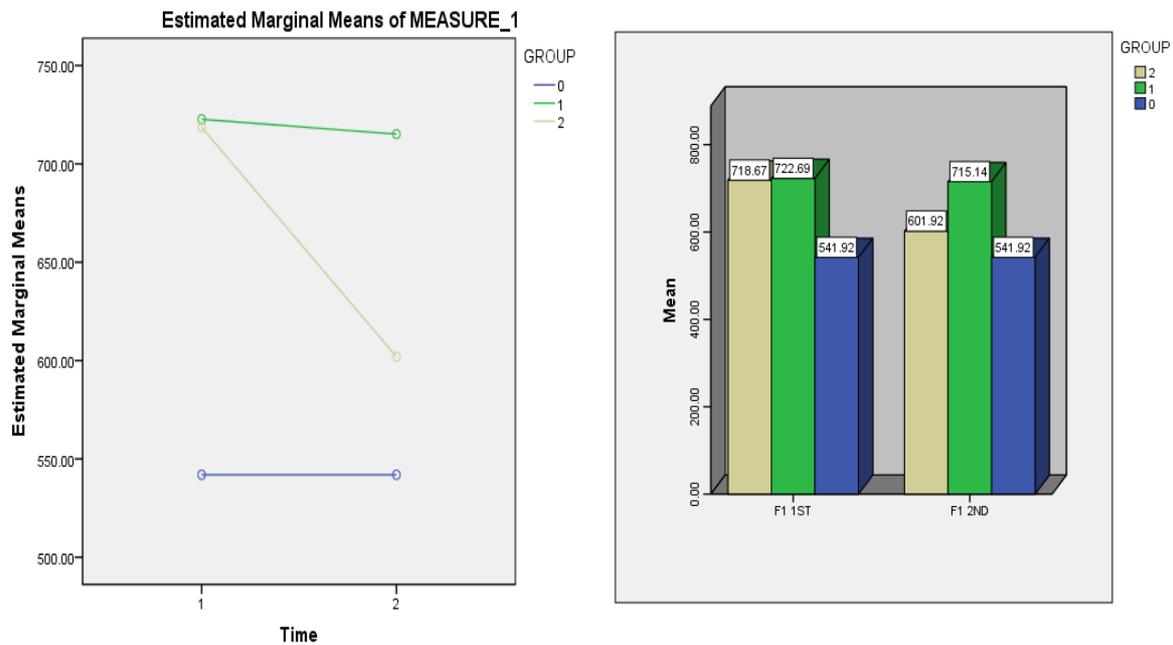
	Pre-test	Stimuli input period (two weeks)	Immediate post-test
Control group (N=9)	√	×	√
Experimental group (N=9)	√	√	√

4. Data analyses and results

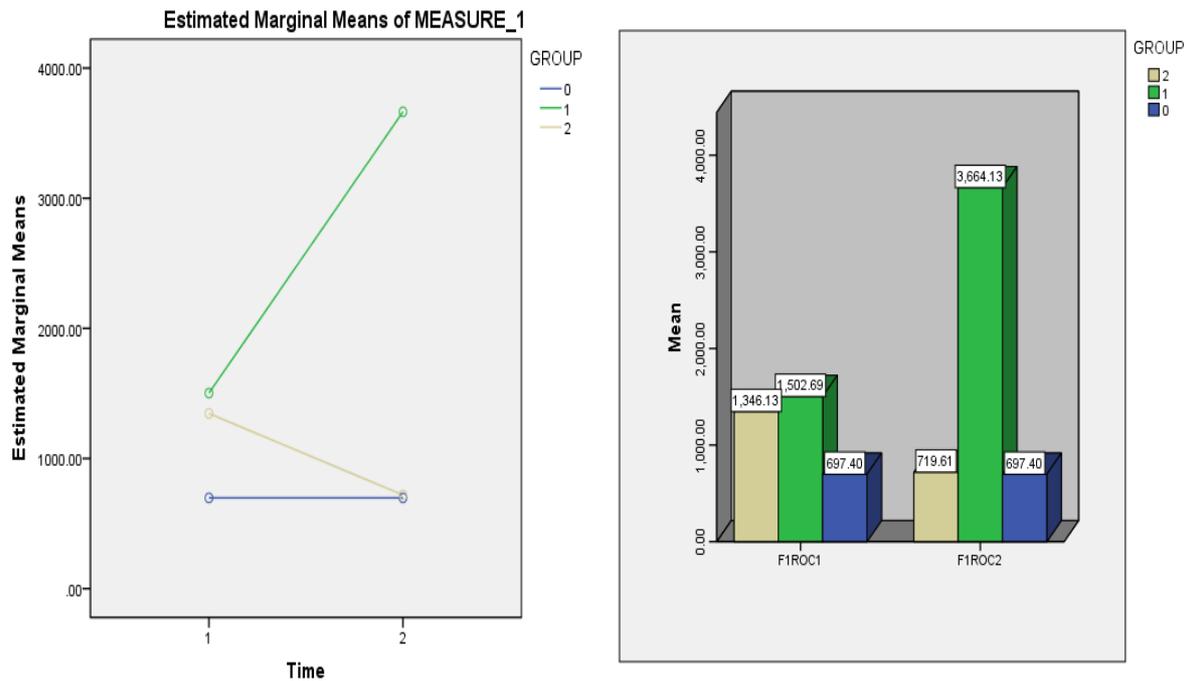
All the data were analyzed in the phonological analysis software ‘Praat’. Statistical analyses were conducted using SPSS. The next two subsections present the data analyses and results from the perspective of vowel quality, including monophthongs and diphthongs.

4.1. Vowel quality - monophthongs

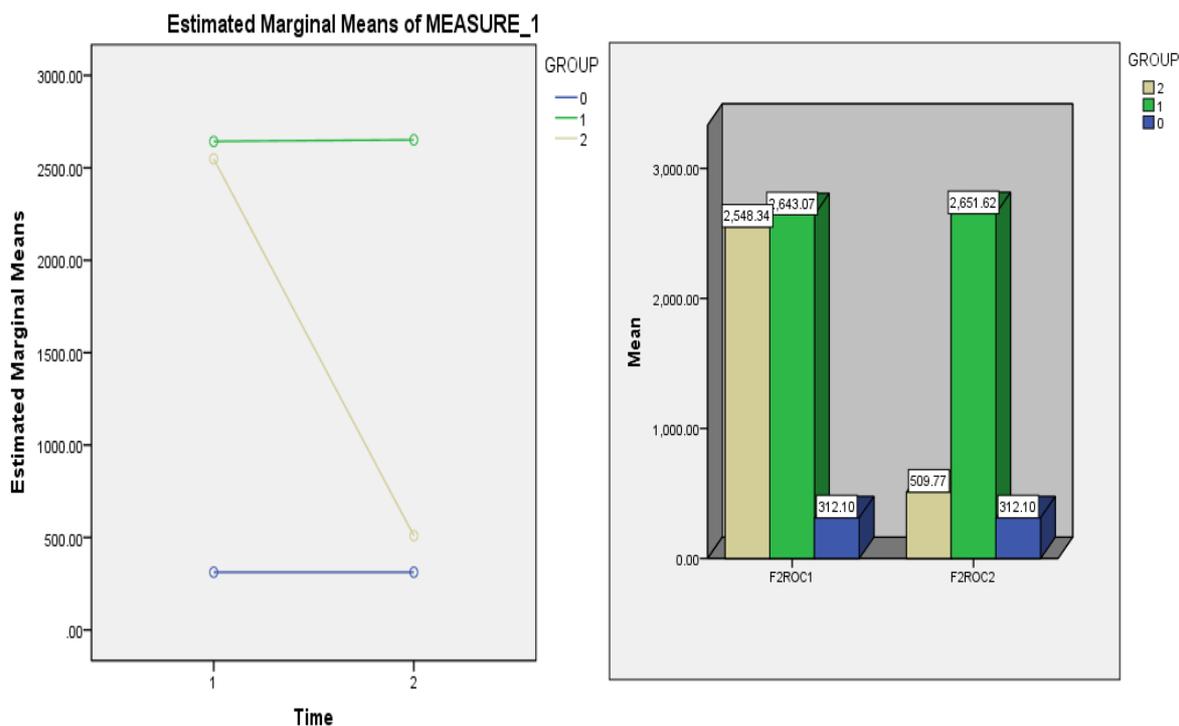
In the examination of monophthongs, it was found that /ɔ:/ had a tendency to be more open in participants’ interlanguage in the pre-test. This was observed based on the F1 values from the word ‘taught’ in the Queen’s address. Generally speaking, the F1 value represents the vowel openness. Compared to close vowels, open vowels have higher F1 values. From Figure 1 shown below, it is obvious that the F1 values of both the experimental group (Category 2 in yellow) and the control group (Category 1 in green) are much higher than the baseline (Category 0 in blue). A statistical comparison made in the F1 values revealed that there was no significant between-group difference in the pre-test ($p=0.957 > 0.05$). Therefore, the performances of the two groups are valid for comparison from the starting point. After the input period, the experimental group’s mean F1 value was much lower than the control group’s F1 value and much closer to the baseline in the immediate post-test. Though the changes that the two groups made from pre-test to immediate post-test were not statistically different ($p=0.117 > 0.05$), I argue that audio-visual media had a slight positive impact according to the level of mean F1 values.

Figure 1. Between-group comparison: mean F1 values of /ɔ:/

Apart from the finding of /ɔ:/, it was found that the monophthong /i/ in the word ‘office’ in the Queen’s address had a tendency to be diphthongal in the pre-test. This finding was based on the great rate of change in F1 (representing the vowel openness) and F2 (representing the vowel frontness) values in the pre-test. In the analysis, the diphthongal extent was measured by using the rate of change (ROC), calculated using the equation ‘(End F value- Start F value)/Duration’ (Deterding 2000). Big absolute ROC values indicate great changes. Figure 2 and Figure 3 below present the between-group comparisons of mean F1 ROC values and mean F2 ROC values of /i/ respectively.

Figure 2. Between-group comparison: mean F1 ROC values of /i/

From Figure 2, it is obvious that in the pre-test, both the experimental group (Category 2 in yellow) and the control group (Category 1 in green) had a larger ROC value in F1, compared to the baseline (Category 0 in blue). Statistical analysis demonstrated the group homogeneity in the pre-test ($p=1.000 > 0.05$). After the audio-visual input period, the experimental group's F1 ROC value dropped and was very much closer to the baseline. However, the control group's F1 ROC value continued to be larger in the immediate post-test. Though the statistical results did not show a significant difference between the experimental group and the control group in the immediate post-test ($p=0.318 > 0.05$), the F1 rate of change of the experimental group did become smaller.

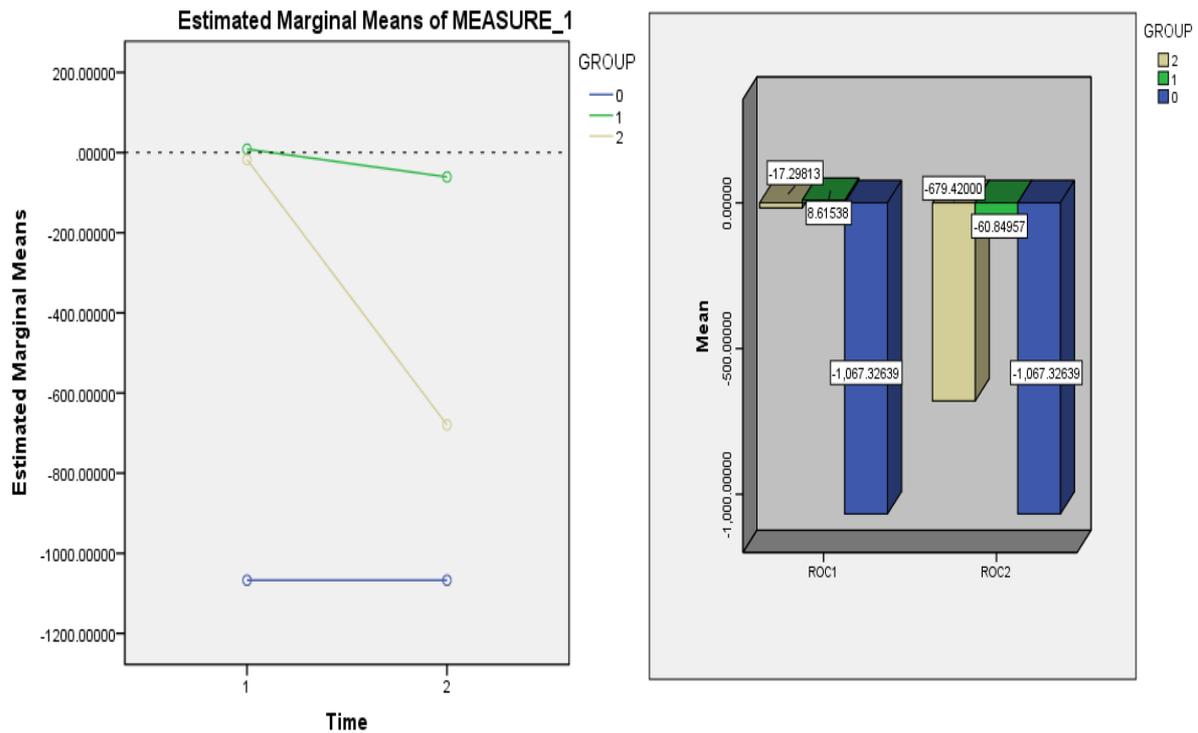
Figure 3. Between-group comparison: mean F2 ROC values of /i/

As is shown in Figure 3, in the pre-test stage the F2 ROC values of the experimental group and the control group were very similar. Statistical results show that the two groups were homogeneous in the pre-test ($p=1.000 > 0.05$). After the input period, in the immediate post-test, the experimental group's F2 ROC value significantly dropped down to the baseline level ($p=0.018 < 0.05$). Also, there was a significant difference between the experimental group and the control group in the immediate post-test ($p=0.009 < 0.05$).

When summarizing the results of F1 ROC values and F2 ROC values, it is obvious from Figure 2 and Figure 3 that /i/ pronounced by the experimental group is less diphthongal in the immediate post-test, compared to the control group. Thus, the audio-visual media input positively influenced the pronunciation of /i/ in the word 'office'.

4.2. Vowel quality - diphthongs

When examining the diphthongs, it was found that /əu/ was less diphthongal in the participants' interlanguage in the pre-test. In English, there is a clear glide from /ə/ to /u/ when pronouncing the diphthong /əu/. The F1 value, which represents the vowel openness, and the F2 value, which represents the vowel frontness, are both expected to change greatly. However, the interlanguage characteristics in the pre-test showed that the participants generally glided from /ə/ to /ɔ/, and then ceased to glide to /u/. Their F2 values changed as is the way in English, but their F1 values did not change much. Figure 4 below presents the between-group comparison of F1 ROC values for the diphthong /əu/.

Figure 4. Between-group comparison: mean F1 ROC values of /əʊ/

From Figure 4 we can see that the between-group difference in the pre-test is very subtle. This is also manifested in the statistics, where no significant difference is found between the experimental group (Category 2 in yellow) and the control group (Category 1 in green) at the starting point ($p=0.929 > 0.05$). Figure 4 also shows that the mean absolute F1 ROC value of the control group does not vary much from the pre-test to the immediate post-test. In comparison, the mean absolute F1 ROC value of the experimental group increases drastically from the pre-test to the immediate post-test and reaches towards the baseline value (Category 0 in blue). Statistical analyses reveal that the experimental group significantly improved in pronouncing /əʊ/ ($p=0.002 < 0.05$) from the pre-test to the immediate post-test. Apart from this, another significant difference was found in the between-group comparison in the immediate post-test ($p=0.015 < 0.05$). It is apparent that the audio-visual media appear to have had a positive effect on the English pronunciation of diphthong /əʊ/.

5. Discussion

This study investigated the effectiveness of audio-visual media on English phonological defossilization by Chinese L2 adult learners. The results shown in the previous section provided evidence for the hypothesis that audio-visual media can influence L2 English phonological acquisition positively, with regard to vowel quality (monophthongs and diphthongs). The effectiveness can be accounted for by the reduction of negative first language transfer, the visual speech aid and the input of native-like phonological exemplars.

As can be seen in the investigation of diphthongs, obvious transfer from L1 Chinese was detected in the participants' interlanguage. Their pronunciation of diphthongs was less diphthongal according to the F1 ROC values in the pre-test. As explained in Section 3.1, Chinese diphthongs are less diphthongal than English ones, and the characteristics of Chinese diphthongs are retained when participants speak English. Another interlanguage characteristic

from Chinese transfer was manifested in the diphthongal /i/. As stated in Section 3.1, there are several combinations of vowels and consonants which exist in English that cannot be realized in Chinese. Among them, /fi/ is an impossible combination in Chinese. Chinese learners would look for a similar combination in their knowledge of Chinese phonology and thus pronounce /fi/ as /fei/. Audio-visual media, which provide high quality native-like English input in sufficient amounts, weaken the negative transfer from the first language. As is shown in the results, /əu/ became more diphthongal and /i/ became less diphthongal in the experimental group's performance in the immediate post-test, and their pronunciations became much closer to the standard baseline. Audio-visual media can potentially break the fossilized deadlock in the interlanguage caused by negative first language transfer, and thus push forward the continuum of phonological acquisition.

In the data analysis of the diphthongs, it was found that in the pre-test, the participants glided from /ə/ to /ɔ/ and then ceased to glide to /u/ at the end. In other words, the participants moved the vowel from the front to the back correctly. However, they failed to control the vowel openness, as is the way in English. After the input stimuli of audio-visual media was given, the experimental group significantly improved in the control of vowel openness. I argue that this might be attributed to the visual speech aid provided in the audio-visual media. With the provided stimuli, it is hypothesized that participants in the experimental group could see the mouth movement and contours from the actors and the actress, which offered cues to vowel openness. By referring to the lip contours, the participants who received the stimuli glided from the front position to the back and at the same time closed the mouth in order to pronounce /əu/ correctly. This further provides demonstrative evidence for Davis and Kim's study (2001).

The results also show that /ɔ:/ has a tendency to be more open in participants' interlanguage in the pre-test. After the input period, the experimental group performed better than the control group, according to the mean F1 values, though not significantly so. This indicates that the experimental group might have conquered the problem of vowel openness with the help of audio-visual media. In my opinion, this might be due to the fact that audio-visual media can provide learners with effective and native-like phonological exemplars. The phonological inventory in the interlanguage might be positively influenced, and so be more likely to be reorganized, thus promoting the process of L2 English phonological acquisition.

This study also offers practical implications. To compensate for the lack of native-like phonological input in China, audio-visual media can be widely applied in classroom teaching, as well as individual learning. Chinese learners of L2 English can benefit from the language input in audio-visual media without being immersed in English-speaking countries.

Due to the scale of the project, this study has some limitations. Admittedly, the input period (2 weeks) is considered to be a little short. If the period could be extended, the effects of the input from films and TV series would be much more salient and more data about, for example, consonants or perhaps sentence intonation, could be elicited. Also, if time allowed, delayed post-tests could also be conducted, in order to see whether the audio-visual media can have a long-term effect.

6. Conclusion

The present study investigated the effectiveness of audio-visual media as language input for L2 English phonological defossilization by adult learners in China. An intervention study which consisted of a pre-test and an immediate post-test was conducted. The results revealed that audio-visual media might potentially have a positive influence on the acquisition and defossilization of vowel quality, with regard to monophthongs and diphthongs. The positive impact was attributed to the nature of audio-visual media input, which can weaken negative first language transfer, provide a visual speech aid and offer native-like phonological

exemplars. This suggests that audio-visual media can tackle the current problem of L2 input in China and can be widely applied in classroom teaching and individual learning.

References

- Altenberg, E. P. & Vago, R. M. (1987). Theoretical implications of an error analysis of second language phonology production. In Ioup, G. & Weinberger, S. H. (eds.), *Interlanguage phonology: the acquisition of a second language sound system*, 148-164. Cambridge, MA: Newbury House Publishers.
- Another Tongue. (2015). *Miranda Hart*. <http://www.anothertongue.com/artist/miranda-hart> [Retrieved on 15 April, 2015].
- Bassetti, B. (2009). Orthographic input and second language phonology. In Piske, T. & Young-Scholten, M. (eds.), *Input matters in SLA*, 191-206. Clevedon: Multilingual Matters.
- Benson, B. (1988). Universal preference for the open syllables as an independent process in interlanguage phonology. *Language Learning* 38, 221-235.
- Cao, D. 曹迪 (2012). 语言教育政策与文化软实力 [Language education policy and cultural soft power]. *语文学刊 Yuyan xuekan* 4, 7-8.
- Carroll, S. (2007). Autonomous induction theory. In VanPatten, B. & Williams, J. (eds.), *Theories in second language acquisition: an introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Chai, R. 柴瑞琴 & Lu, Y. 卢玉芳 (2009). 二语习得课堂中的教师话语分析 [Analysis of teacher talk in the L2 classroom]. *长春理工大学学报(社会科学版) Changchun ligong daxue xuebao (shehui kexue ban)* 22(4), 648-649.
- Chen, J. 陈君君 (2011). 二语学习者英语语音的石化现象研究 [Research on the English phonological fossilization of L2 learners]. *黑龙江教育学院学报 Heilongjiang jiaoyu xueyuan xuebao* 30(7), 147-148.
- Cruttenden, A. (2014). *Gimson's pronunciation of English*. Oxon: Routledge.
- Davis, C. & Kim, J. (2001). Repeating and remembering foreign language words: implications for language teaching systems. *Artificial Intelligence Review* 16, 37-47.
- Deng, L. 邓利蓉 (2007). 教师话语在二语习得课堂上的作用 [The use of teacher talk in the L2 classroom]. *西昌学院学报(社会科学版) Xichang xueyuan xuebao (shehui kexue ban)* 19(2), 142-145.
- Deterding, D. (2000). Measurements of the /eɪ/ and /əʊ/ vowels of young English speakers in Singapore. In Brown, A, Deterding, D. & Low, E. L. (eds.), *The English language in Singapore: research on pronunciation*, 93-99. Singapore: Singapore Association for Applied Linguistics.
- Ellis, N. (2009). Input and second language acquisition: the roles of frequency, form, and function introduction to the special issue. *The Modern Language Journal* 93(3), 329-335.
- Ellis, R. (1985). *Understanding second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Fanelli, G., Gall, J. & Gool, L. J. V. (2009). Hough transform-based mouth localization for audio-visual speech recognition, in 'British Machine Vision Conference'. British Machine Vision Association.
- Flege, J. (2009). Give input a chance! In Piske, T. & Young-Scholten, M. (eds.), *Input matters in SLA*, 175-190. Clevedon: Multilingual Matters.

- Gass, S. (1997). *Input, interaction and the second language learner*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Hansen, J. G. (2006). *Acquiring a non-native phonology: linguistic constraints and social barriers*. London: Continuum.
- Henzl, V. (1979). Foreigner talk in the classroom. *International Review of Applied Linguistics* 17, 159-165.
- Jenkins, J. (2000). *The phonology of English as an international language*. Oxford: Oxford University Press.
- Jia, H. 贾红安 (2011). 浅谈当代大学英语教师的综合素质 [Brief discussion of essential qualities of contemporary college English teachers]. *价值工程 Jiazhi gongcheng* 28, 162.
- Johnson, K. (1997). Speech perception without speaker normalization. In Johnson, K. & Mullennix, J. (eds.), *Talker variability in speech processing*, 145-165. San Diego: Academic Press.
- Karakas, A. & Saricoban, A. (2012). The impact of watching subtitled animated cartoons on incidental vocabulary learning of ELT students. *Teaching English with Technology* 4, 3-15.
- Klima, E. & Bellugi, U. (1979). *The signs of language*. Cambridge, MA: Harvard University Press.
- Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.
- Krashen, S. (1985). *The input hypothesis: issues and implications*. New York: Longman.
- Krashen, S. (2003). *Explorations in language acquisition and use: the Taipei lectures*. Portsmouth: Heinemann.
- Li, H. 李卉 (2011). 高校英语课堂教师话语: 问题与对策 [Teacher talk in higher education English classes: problems and solutions]. *内蒙古农业大学学报 (社会科学版) Neimenggu nongye daxue xuebao (shehui kexue ban)* 13(1), 144-145.
- Li, L. 李力 & Wang, H. 王海波 (2006). 大学英语口语教学问题的探讨与对策 [A discussion of oral English teaching problems and counter-measures in colleges]. *浙江万里学院学报 Zhejiang Wanli xueyuan xuebao* 19(4), 149-151.
- Li, R. 李瑞 (2009). 浅析英语课堂上教师的语言输入 [Brief discussion on teacher's language input in English classrooms]. *内江科技 Neijiang keji* 4, 60-61.
- Li, X. 李霞 (2007). 教师话语及其对英语课堂教学的影响 [The influence of teacher talk in the EFL classroom]. *沈阳农业大学学报 (社会科学版) Shenyang nongye daxue xuebao (shehui kexue ban)* 9(4), 580-583.
- Liu, X. 刘晓明 (2010). 非英语专业大学生语言输入现状调查与研究 [An investigation of the current situation concerning language input among non-English major students in Chinese universities]. *科技信息 Keji xinxi* 29, 240-241.
- Loschky, L. (1994). Comprehensible input and second language acquisition. *Studies in second language acquisition* 16, 303-323.
- Macaluso, E., George, N., Dolan, R., Spence, C. & Driver, J. (2004). Spatial and temporal factors during processing of audiovisual speech: a PET study. *Neuroimage* 21(2), 725-732.
- Moyer, A. (2009). Input as a critical means to an end: quantity and quality of experience in L2 phonological attainment. In Piske, T. & Young-Scholten, M. (eds.), *Input matters in SLA*, 159-174. Clevedon: Multilingual Matters.
- Neuman, S. & Koskinen, P. (1992). Captioned television as comprehensible input: effects of incidental word learning from context for language minority students. *Reading Research Quarterly* 27(1), 94-106.

- Niyogi, P., Petajan, E. & Zhong, J. (1999). Feature based representation for audio-visual speech recognition. In *proceeding of audio-visual speech processing*, Santa Cruz, USA. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.399.3465&rep=rep1&type=pdf> [Retrieved on 15 April, 2015].
- Ortega-Llebaria, M., Faulkner, A. & Hazan, V. (2001). Auditory-visual L2 speech perception: effects of visual cues and acoustic-phonetic context for Spanish learners of English. *Speech, Hearing and Language: work in progress* 13, 39-51.
- Pierrehumbert, J. B. (2001). Exemplar dynamics: word frequency, lenition and contrast. In Bybee, J. & Hopper, P. (eds.), *Frequency and the emergence of linguistic structure*, 137-158. Amsterdam: John Benjamins.
- Piske, T. & Young-Scholten, M. (2009). *Input matters in SLA*. Clevedon: Multilingual Matters.
- Plass, J. & Jones, L. (2005). Multimedia learning in second language acquisition. In Mayer, R. (ed.), *The Cambridge Handbook of Multimedia learning*, 467-488. Cambridge: Cambridge University Press.
- Rast, R. (2008). *Foreign language input: initial processing*. Clevedon: Multilingual Matters.
- Schwartz, B. (1993). On explicit and negative data effecting and affecting competence and linguistic behaviour. *Studies in Second Language Acquisition* 15, 147-163.
- Scovel, T. (1998). *A time to speak: a psycholinguistic inquiry into the critical period for human speech*. Cambridge, M.A.: Newbury House.
- Scovel, T. (2000). A critical review of critical period research. *Annual Review of Applied Linguistics* 20, 213-223.
- The Ministry of Education, P. R. China. (2011). *English course standard of compulsory education*. Beijing: Beijing Normal University Press.
- Tu, J. 屠静 (2012). 教师话语对中学二语习得课堂教学的影响 [The influence of teacher talk on second language teaching in middle schools]. *中国科教创新导刊 Zhongguo kejiào chuàngxīn dǎo kān* 8, 154.
- VanPatten, B. (2004). Input processing in second language acquisition. In VanPatten, B. (ed.), *Processing instruction: theory, research and commentary*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wells, J. C. (1982). *Accents of English*. Cambridge: Cambridge University Press.
- Xiang, T. 项湍舟 (2009). 英语语音学习过程中的“石化现象” [Fossilization in English pronunciation learning]. *宜春学院学报 Yichun xueyuan xuebao* 31(1), 149-151.
- Xu, L. 徐丽仙 & Guo, H. 郭弘 (2009). 教师话语研究及其对英语课堂教学的启示 [Research on teacher talk and its implications for English classroom teaching]. *内蒙古师范大学学报(教育科学版) Neimenggu shifan daxue xuebao (jiaoyu kexue ban)* 22(3), 112-115.
- Yang, Y. 杨艳慧 & Ren, T. 任婷 (2009). 大学英语精读课教师话语研究 [Research on teacher talk in college English intensive reading]. *吉林工程技术师范学院学报 Jilin gongcheng jishu shifan xueyuan xuebao* 25(2), 6-8.
- Young-Scholten, M. (1995). The negative effects of ‘positive’ evidence on L2 phonology. In Eubank, L., Selinker, L. & Sharwood Smith, M. (eds.), *The current state of interlanguage*, 107-122. Amsterdam: Benjamins.
- Yuksel, D. & Tanriverdi, B. (2009). Effects of watching captioned movie clip on vocabulary development of EFL learners. *The Turkish Online Journal of Educational Technology* 8, 48-54.
- Zhang, J. 张洁君 (2009). 大学英语课堂教师话语存在的问题及解决策略 [The problems of teacher talk in college English classroom teaching and the solutions]. *考试周刊 Kaoshi zhoukan* 3(3), 126-127.

- Zhang, Y. 张尧 (2013). 英语电影对大学生学习英语的重要性 [The importance of English movies to English learning among college students]. *创新教育 Chuangxin jiaoyu* 6, 60.
- Zheng, X. R. 郑鲜日 & Zheng, X. J. 郑晓杰 (2011). RP 英语和 PTH 汉语双元音声学特征对比研究 [Research on the comparison of the acoustic characteristics of diphthongs between RP English and PTH Chinese]. *东疆学刊 Dongjiang Journal* 28(3), 99-105.

Chen ZHANG
School of English
University of Sheffield
4.80b Jessop West
1 Upper Hanover Street
SHEFFIELD
S3 7RA
UK

czhang17@sheffield.ac.uk