

Improving English Pronunciation Teaching and Learning via Speech Corpora of Learners with Dialectal Backgrounds

<http://dx.doi.org/10.3991/ijet.v11i04.5461>

Ruilan CAO

Anhui University of Technology, Ma'anshan, P. R. China

Abstract—Speech corpora play an important role in phonetic research and have been applied in daily life. Their application can be found in pronunciation teaching and learning. This article takes Jianghuai mandarin, a transitional dialect between northern and southern dialects in China, as an example to illustrate the necessity and procedures about the construction of an English speech corpus of learners with dialectal backgrounds. It shows that such corpus can be applied to help improve English pronunciation teaching and learning with its technical power in saving large audio data and its annotation with voice editing and analyzing software. It assists, when working together with Praat, to visually demonstrate spectrogram of similarities and differences between English and native language. It also functions to raise pronunciation awareness and help improve learners' pronunciation. It is expected to be useful in stimulating teachers and learners to study pronunciation based on data analysis of physical features of recorded audio sound.

Index Terms—corpus, dialect, English pronunciation, learner, speech

I. INTRODUCTION

A corpus is a collection of data. It is constructed with a certain technology and method for research or application. There are corpora collecting spoken and/or written data. A speech corpus contains speech data, usually in audio form, with the help of recorders, computers, recording software, and voice editing and/or analyzing software. Speech corpora and relevant technologies make it possible for researchers to analyze features of speech data, and to seek and reveal representative speech phenomena and patterns through quantitative and qualitative analysis.

In linguistic field, much attention has been paid to the construction and application of speech corpora because of the important role that speech corpora play in linguistic research, especially in phonetic research. Phonetic synthesis, phonetic recognition and modern phonetics, to name just a few, are areas that are fruitful in the application of speech corpora. The development of voice reading and dialing, based on phonetic synthesis and large-scale speech corpora, is a welcoming achievement of the application. Speech corpora, with their research stepping forward steadily, are found to be helpful in language teaching and learning as well, including pronunciation teaching and learning.

II. DEVELOPMENT OF SPEECH CORPORA

Generally speaking, there are three stages in the development of speech corpora. In the first stage, with the invention of recorders, researchers were able to collect speech data and save the data for future study. But the scale of database could not reach what it could be in the second stage. Furthermore, transcription of the collected acoustic data has to be done with ears by writing notes on pieces of paper, which is obviously time-consuming and hard for data analysis. In the second stage, speech corpora were built with computers and software. Much more data could be collected and saved in disks with higher quality. Transcription could be achieved with software, making it more efficient and easier for data analysis. The last two decades fall in the third stage which witnesses the boost of speech corpus construction, its research and application.

Concerning acoustic features of learners' English, many English speech corpora have been constructed. Some of them are of global influence. Speech Accent Archive in America is an example of many well-constructed speech corpora. It contains 2061 samples of accents with participants coming from 246 native language backgrounds [1]. In China, available well-constructed large-scale speech corpora of learners include College Learners' Spoken English Corpus (CLSEC), Spoken and Written English Corpus of Chinese Learners (SWECCL), and English Speech Corpus of Chinese Learners (ESCCL).

III. CONSTRUCTION OF ENGLISH SPEECH CORPORA OF LEARNERS WITH DIALECTAL BACKGROUNDS

A. Necessity of Constructing English Speech Corpora of Learners with Dialectal Backgrounds

Non-native English speakers are often noticed to exhibit pronunciation characteristics. Debate in linguistic studies is heated about why learners demonstrate varieties of their second language. Lado (1957) claimed that differences between a learner's native language and the learner's second language would bring problems and difficulties in second language acquisition [2]. Concerning pronunciation, it is believed with this idea that phonetic and phonological differences between native language and second language would bring more problems and difficulties in second language acquisition. However, Flege (1995) proposed that similarities would lead to more problems in language acquisition than differences. That is to say, in second language pronunciation acquisition, similar phones would be more problematic than new ones in the learner's pronunciation acquisition of a second

language [3]. With the fast development of computer technology, corpus-based linguistics becomes important in understanding the nature of language [4] and language acquisition [5]. English speech corpora offer audio data as systemic and robust evidence in the study of language acquisition to investigate pronunciation development of non-native speakers [6].

This paper focus on English learners in China to illustrate the construction of English speech corpora of learners with dialectal backgrounds due to the fact that they contribute to a large percentage of English learners around the world. They have diversified native language backgrounds when their dialectal mother tongues are concerned after the use of Mandarin Chinese being promoted only several decades ago in mainland China.

Though research has claimed that learners' native languages should be taken into account when designing a corpus [7], most English speech corpora in China collected speech data without special concern of various dialects spoken by their participants. ESCCL is one of the few corpora which have considered dialectal impact on English speech output. Participants of this large-scale corpus are from most of the dialectal areas in China. However, the speech material was not designed with special concern of Chinese dialectal differences. Therefore, some dialectal impacts on English speech of Chinese learners may not be found in its collected data.

According to the dialect map issued by Chinese Academy of Social Sciences, there are ten major dialects in China [8]. Another practical classification states that there are southern dialects and northern ones, at the junction of which Jianghuai mandarin is spoken as a transitional dialect. Jianghuai mandarin indicates the developmental history of Chinese dialects with its unique phonetic and phonological features. Therefore, the construction of English speech corpus of learners with dialectal backgrounds introduced in this paper will be exemplified by that of learners who speak Jianghuai mandarin.

In most research about Jianghuai mandarin, phonetic features and the fourth tone of this dialect are concerned, with mandarin Chinese as a reference. Few research is found on Jianghuai mandarin with English phonetics as its reference. Only several studies have been conducted on merely segmental output of English speech produced by Jianghuai mandarin learners and learners are noticed with many problems in English pronunciation learning and production [9-12]. One of the obstacles against the development of the research and learners' pronunciation learning lies in the lack of a large-scale speech corpus designed for English learners with Jianghuai mandarin as their mother tongue.

By offering large-scale and systematic speech data, the construction of an English speech corpus of Jianghuai mandarin learners would promote the research on English interlanguage speech of Jianghuai mandarin learners in segmental and suprasegmental aspects. The revealing of features and patterns of learners' English would, consequently, lead researchers and teachers to reveal features and patterns of learners' English as an interlanguage and help raise learners' awareness on pronunciation problems and difficulties. With reasonable application of the speech corpus, it is possible to improve English pronunciation teaching and learning.

Similarly, the construction of English speech corpora of learners with other dialectal backgrounds would help with the research on these dialects and their speakers' English production. Their application would help with English speech teaching and learning.

B. Procedures in Constructing English Speech Corpora of Learners with Dialectal Backgrounds

Procedures to construct English speech corpora that are exemplified with the construction of English speech corpus of Jianghuai mandarin learners involve four main steps, namely, the selection of participants, the design of speech material, the collection and classification of linguistic data, and the annotation of collected phonetic data.

The first step is the selection of participants. Referring to ESCCL, 160 English learners who are native speakers of Jianghuai mandarin were selected from four Jianghuai mandarin dialectal regions where citizens speak with unique phonetic features of Jianghuai mandarin. In each region, 40 learners were selected from junior high schools, senior high schools, and first and fourth grades in colleges and universities, with 10 from each. The male and the female participants are of an equal proportion.

The second is the design of speech materials. Speech materials involve check points of segmental and suprasegmental elements designed in two types of tasks. One is reading-aloud task of English sentences and dialogues. Another is about questions and answers. They have been designed with the consideration of learners' English levels. Reading materials are simple and easy with mostly high-frequency words. Each questions should be answered in a complete sentence with several words.

These two tasks have their own advantages and disadvantages, and exert complementary effect for the construction of this phonetic database. The reading-aloud task satisfies requirements for those check points in learners' output. However, the output may be influenced by participants' phonetic awareness. If participants notice the purpose of the task and pay more attention to pronunciation to this task, their production will deviate from their daily performance. That generally brings better output and results in fewer problems and difficulties in the database. The task of questions and answers is useful in focusing their attention on contents and grammar of answers and diminish learners' awareness of pronunciation. This task helps collect phonetic and phonological awareness free audio data.

The third step is the collection and classification of linguistic data. Information of the participants was registered and verified, including name, gender, age, and whether having been trained of Chinese and/or English pronunciations. Instruments used in speech collection include a Creative Sound Blaster and a Creative HS-300 headset. Voice editing software, Cool Edit Pro 2.1, was adopted to collect speech audio with sampling rate at 44100 Hz. The classification of collected audio data was conducted by researchers concerning task types, segmental and suprasegmental features of English speech, participants' gender and educational level.

Finally, it is the annotation of collected phonetic data. It is achieved with the voice editing and analyzing software of Praat. The speech data annotation has been designed for seven levels, involving segment, rhythm, stress and tone, with reference to the annotation design of ESCCL.

The first level of the annotation is about the description of actual speeches presented with words; the second and third levels are for standardized speeches. The second level is presented in form of syllables, and the third phonemes. These are annotations on segmental levels.

Levels four to seven are annotations of suprasegmental features, including rhymes, stresses and intonations. The fourth level involves the annotation of intonation phrase boundary, middle phrase boundary, prosodic words' boundary and bound morpheme. The fifth level annotates stressed syllables in sentences. The sixth and seventh levels are for intonation in accordance with the British intonation pattern and the American intonation pattern to describe continuum of tone groups in the British intonation pattern, and pitch events and boundary tone in the American intonation pattern.

IV. APPLICATION OF THE CORPORA IN ENGLISH PRONUNCIATION EDUCATION

English pronunciation teaching is booming following the prosperous English speech study. This is especially true in mainland China in recent years. Teachers and learners care more about English pronunciation than before. Modern technologies have been gradually applied in English phonetic research and education. As achievements of modern technologies, English speech corpora contribute to English pronunciation education and research in various ways. The followings demonstrate the application of English speech corpora of learners with dialectal backgrounds with an English speech corpus of learners speaking Jianghuai mandarin as an example.

A. Explaining Contrastive Differences between English and Jianghuai mandarin with Spectrographic Analyses

Differences have been found in segmental and suprasegmental areas. In segmental sphere, some phonemes, incomplete explosions, liaisons in English are not found in Jianghuai mandarin. In suprasegmental sphere, Jianghuai mandarin, a dialectal Chinese, is a syllable-timed language, while English is a stress-timed language.

Without the application of data from speech corpora of learners, it would be hard for teachers to explain segmental and suprasegmental differences between a native-like sentence and a typically dialect-impacted one from the acoustic perspective. For instance, when explaining stress difference between Jianghuai mandarin and English, teachers often explain that English learners whose native language is Jianghuai mandarin tend to stress all words with no obvious rhythm. Though it is true, it is weak to clarify differences between the two versions. After the explanation, even if learners may have a vague idea about the problem they have in word and sentence stresses, it would be hard for the learners to improve. Fig. 1 and Fig. 2 show how the corpus that is annotated with voice analyzing software helps explain pronunciation differences and problems.

Because modern voice analyzing software can visually demonstrate waveform and analysis of audio data, teachers can help learners compare spectrographs of a learner's pronunciation and a reference pronunciation (usually Received Pronunciation in pronunciation class). It becomes easier for teachers and learners to notice the difference between the two, if there is any, by considering parameters such as duration, pitch, and intensity.

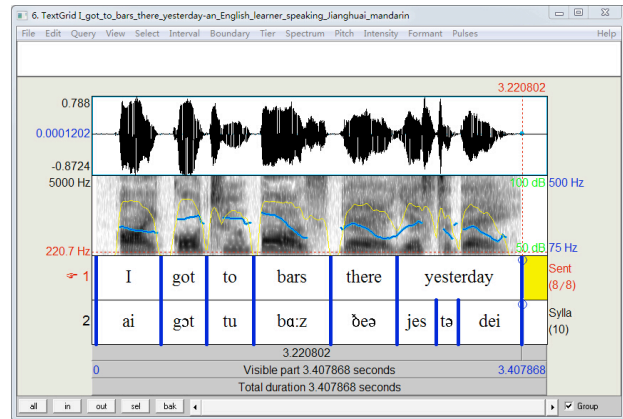


Figure 1. Annotation of the English sentence "I got to bars there yesterday." uttered by an English learner speaking Jianghuai mandarin

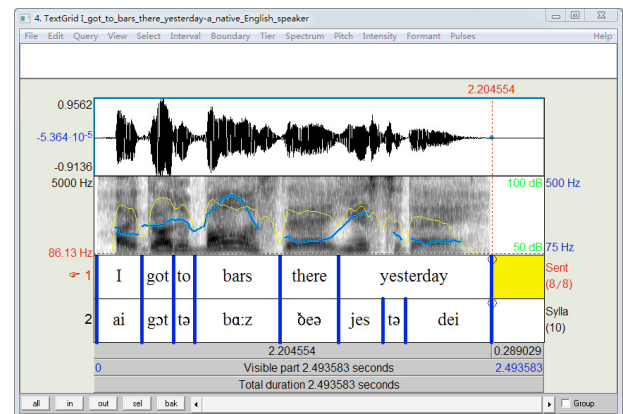


Figure 2. Annotation of the English sentence "I got to bars there yesterday." uttered by a native English speaker

The two annotations for data in the corpus demonstrate acoustic features of the English sentence "I got to bars there yesterday." uttered by two speakers. One is an English learner speaking Jianghuai mandarin, and the other is a native English speaker. Fig. 1 is for the sentence uttered by the English learner speaking native Jianghuai mandarin, and Fig. 2 for that orally produced by the native English speaker. In both figures, the yellow lines represent intensities of the utterances, which are indicators of the stress of utterances. In these figures, the blue lines show pitch contours of the utterances. They usually develop in the directions the same with those of the syllable tones.

The first annotation (Fig. 1) shows that the total duration of the recording lasts 3.407868 seconds which is the same with the duration of visible part, and the utterance lasts 3.220802 seconds.

The second annotation (Fig. 2) shows that the total duration of the recording lasts 2.493583 seconds which is the same with the duration of visible part, and the utterance lasts 2.204554 seconds.

According to the above mentioned claim that differences between learners' native language and second language may lead to second language acquisition problems and difficulties, English learners speaking different Chinese dialects would have different pronunciation problems brought about by their dialects. Based on this assumption, learners speaking a certain dialect may refer to an English speech corpus specialized on data reflecting impact of this dialect so that their English pronunciation learning could be more specific and efficient. Furthermore, they could

learn from an English speech corpus specialized on data reflecting impact of some other dialects because of their own interest in various accents.

Nevertheless, starting from the idea that, rather than new phones, similar phones between learners' native language and second language cause more difficulties and problems in learners' second language acquisition, teachers and learners should be aware which phonetic elements tend to be more problematic than others though they are found essentially similar in both languages.

Therefore, no matter whether similarity or difference leads to acquisition difficulties and problems, English speech corpora like these would help learners avoid negative native language transfer in English pronunciation learning, and help them improve pronunciation learning with recorded sampling data that are authentic and convincing.

B. Preparing for Research of English Pronunciation

Without English speech corpora, learners English pronunciation learning probably remains experiential. With English speech corpora of learners with dialectal backgrounds, teachers and learners are able to study learners' pronunciation theoretically by analyzing data in English speech corpora and generalizing patterns of English pronunciation displayed. It would help teachers better understand teaching focus and provide them empirical support for pronunciation teaching. Furthermore, learners of special interest in pronunciation may develop their linguistic study from their own pronunciation learning. Such study would enhance pronunciation teaching and learning.

V. CONCLUSION

The improvement of English pronunciation is an issue for many learners of English as a second language. The construction of English speech corpora of learners with dialectal backgrounds is required in the circumstance of phonetic research and application boom. With the construction involving participant selection, material designing, data collection, classification and annotation, the corpora would function in several aspects. It explains contrasts between English and learners' native language, improves learners' awareness of these contrasts, reminds teachers and learners of possible negative native language transfers, and helps improve pronunciation learning with authentic sampling data. Furthermore, data in the corpora offer source for future research on pronunciation education based on data analysis of physical features of recorded audio sound.

ACKNOWLEDGMENT

I would like to thank Professor Haixiao WANG and Professor Hua CHEN for their encouragement and enlightening ideas in pronunciation education and research. I hereby thank administrative officers in Anhui University

of Technology for their support on my pronunciation teaching and research. I also want to thank my students in their active participation in my pronunciation class and research.

REFERENCES

- [1] <http://accent.gmu.edu/> last visited July 2015.
- [2] R. Lado, *Linguistics across Cultures: Applied Linguistics for Language Teachers*. Ann Arbor, Michigan: University of Michigan Press, 1957.
- [3] J. E. Flege, "Second language speech learning: Theory, findings, and problems," in *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues in Crosslanguage Speech Perception*, W. Strange, Eds. Timonium, MD: York Press, 1995, pp. 233-277.
- [4] W. Chafe, "The importance of corpus linguistics to understanding the nature of language," in *Directions in Corpus Linguistics*, J. Svartvik, Eds. Berlin / New York: Mouton de Gruyter, 1992, pp. 79-104.
- [5] G. R. Yepes and R. Krishnamurthy, "Corpus linguistics and second language acquisition — The use of ACORN in the teaching of Spanish grammar," *Lebende Sprachen*, vol. 55, pp. 108-122, November 2010. <http://dx.doi.org/10.1515/les.2010.008>
- [6] J. Schertz and M. Ernestus, "Variability in the pronunciation of non-native English the: Effects of frequency and disfluencies," *Corpus Linguistics and Linguistic Theory*, vol. 10, pp. 329-345, August 2014. <http://dx.doi.org/10.1515/clit-2014-0024>
- [7] B. Wu, "A phonetic and phonological study of Jianghuai mandarin," Doctoral Dissertation, Fudan University, 2007.
- [8] Z. X. Zhang, *Language Atlas of China*, 2nd ed.. Beijing: Commercial Press, 2012.
- [9] R. L. Cao, "OT analysis of acquisition of sC/sCC onset consonant clusters in English," *Journal of Anhui Normal University (Hum. & Soc. Sci.)*, vol. 38, pp. 458-463, July 2010.
- [10] R. L. Cao, "A study of factors affecting Chinese EFL learners' production of English onset consonant clusters," Doctoral Dissertation, Nanjing University, 2010.
- [11] R. L. Cao, "Predictability of markedness differential hypothesis on the acquisition of onset consonant clusters," *Journal of Anhui University of Technology (Social Sciences)*, vol. 31, pp. 53-54, 58, July 2014.
- [12] R. L. Cao, "A study of Chinese EFL learners' production of English onset consonant clusters: From the perspective of markedness differential hypothesis," *Foreign Languages and Their Teaching*, vol. 282, pp. 8-12, June 2015.

AUTHOR

Ruilan CAO is with Anhui University of Technology, School of Foreign Languages, Ma'anshan, Anhui Province, 243002, P. R. China (e-mail: crl2010@ahut.edu.cn).

This work was supported in part by National Social Science Foundation of China (15BYY024), by Key Teaching Research Program Foundation of Anhui Province, China (2012jyxm183), by Excellent Resource Sharing Course of Quality Project Foundation of Anhui Province, China (2012gxxk040) and by Excellent University Youth Talents Supporting Program of Anhui Province, China (Anhui Education Personnel [2014] Reference No.: 181).

Submitted 15 August 2015. Published as resubmitted by the author 23 September 2015.

APPENDIX

Task I Reading-aloud

Directions: Please read the following sentences and dialogues aloud. Pause about two seconds after each sentence or dialogue is read.

1. There are 20 sheep on the ship.
2. The master said he must finish the work within 2 days.
3. We're all very well, thank you.
4. I got to bars there yesterday.
5. For God's sake, show me something new.
6. They went to the zoo last Sunday.
7. This jeep is not cheap.
8. If I had wing, I would win.
9. Yesterday, she had fun with her mother at the party.
10. Seeing is believing, but feeling's the truth.
11. First come, first served.
12. Cigar butts littered the ground.
13. Eat at a pleasure, drink with measure.
14. This big tin is filled with milk.
15. That man has taken my black jacket.
16. Put it on!
17. There is a bottle of ink on the desk.
18. Take care not to eat too much at the party.
19. It would be a good idea to ask some friends in to celebrate the New Year.
20. A: What did you have this morning?
B: I had an egg, two pieces of bread, and a glass of milk.
21. A: "How many days are there in a week?" Bob asked.
B: "Seven. They are Sunday, Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday." Ben answered carefully.
22. A: It's a fine day, isn't it?
B: Mmm. I don't think so. It's a bit hot.
23. A: I don't think Mary speaks Chinese.
B: I'm sorry. Mary does speak Chinese.
24. A: Would you like tea or coffee?
B: I'm a Chinese. Of course, I prefer tea to coffee.
25. A: Do you want to go to the cinema tonight?
B: Yes, I'd love to, but only if there's something good on.
A: I heard that that film from the United States is really good. I think it's called Home Alone.
B: Oh, yes. Somebody told me that it's a popular one. Let's go!

Task II Questions and Answers

Directions: Please answer each of the following questions with a complete sentence. Pause about two seconds after each question is answered.

1. Where do you come from?
2. Which subject do you like most, English, Chinese or Latin?
3. Do you listen to English radio programs everyday?
4. Would you like something to drink?
5. Can you play football?
6. What time is it now?
7. What is the weather like?
8. What are you going to do this evening?
9. Can you speak Japanese?
10. What do you like doing in your free time?