A Study on the English Monophthong Pronunciation of Tibetan Lhasa Natives Based on Phonetic Parameters

Lei Guo, Guiyun Guan, Qian Zhou

Key Laboratory of China's Ethnic Languages and Information Technology of Ministry of Education, Northwest Minzu University, Lanzhou, Gansu 730000, China

Keywords- Vowels, Tibetan, Lhasa native speakers, formant, duration

Abstract: Based on previous studies and in accordance with relevant specifications, this paper divides, annotates, and extracts the collected English vowel speech signals of Tibetan Lhasa native speakers, and establishes an “English vowel speech acoustic parameter database”. Using the parameter database, a comparative experimental study was made on the pronunciation of English vowels in college students of native Lhasa dialects. Extract the formant and duration parameters corresponding to the 12 phonemes in the English language of college students in Lhasa dialect area, combine the standard pronunciation parameters of English foreign teachers to analyze the acoustics features of the vowel pronunciation in the Tibetan Lhasa dialect area, summarize the parameter features of tongue position and oral cavity changes when vowels are pronounced.

1. Introduction

Tibetan belongs to the Tibetan branch of the Tibetan-Burmese family of the Sino-Tibetan language family. It is mainly distributed in the five provinces and autonomous regions of Tibet, Sichuan, Yunnan, Qinghai, and Gansu in my country. It has ancient Pinyin characters and vast literature. There are about 6.8 million Tibetans in China. Tibetan students have unique language, culture, religious beliefs and customs. Most Tibetan students have difficulty learning English. With the continuous advancement of globalization, English learning has become an indispensable communicative tool in international communication. Therefore, from teachers to students, Tibetan areas are in urgent need of English talents. The study of second language acquisition of different ethnic minorities, especially like the study of second language acquisition of Tibetan compatriots who use Tibetan dialects, that has a great practical significance to further promote national integration and common progress. At present, most of the relevant research abroad is concentrated on the main European languages. In contrast, although domestic research on second language acquisition has developed rapidly, however, it is limited to research techniques and methods. There are not many research results on both foreign language acquisition of ethnic minorities and the acoustic and perceptual speech digitization.[1]

2. The Research Method

The Tibetan language is a phonetic alphabet with 30 consonants and 4 vowel symbols. The Tibetan syllable (character) has a special morphological structure. It is composed of 30 letters and 4 vowel letters vertically and horizontally superimposed.The vowel system of ancient Tibetan is simpler only five vowels, while Lhasa now has eight short vowels and eight long vowels corresponding to it, there are only 4 true complex vowels. English is also a kind of pinyin script. The pronunciation has 20 vowel phonemes, including 12 monophthongs and 8 diphthongs. The acoustic characteristics of vowels are mainly determined by the harmonic frequency structure. The formant and duration parameters corresponding to the 12 monophthongs in the English of Lhasa dialect students are extracted, and the standard English pronunciation of English foreign teachers is combined to analyze the second language learning.[2] The acoustic characteristics of vowel sounds are summarized as the characteristics of the tongue position and oral cavity during vowel sounds. We analyze the acoustic characteristics of the pronunciation of English vowels in the Lhasa dialect

  DOI: 10.25236/icembe.2020.047

of Tibetan, and summarize the parameter characteristics of tongue position and oral cavity changes when vowels are pronounced.

2.1 Brief Introduction to the speaker

This study was conducted using the methods of phonetic acoustics and phonetic perception. We selected 30 college students whose native language is Tibetan dialect in Lhasa dialect, aged 18-25 years, based on the questionnaires, phonetic tests and the English graded test. Finally 15 low-level and high-level subjects were selected, and the speaker's listening and speaking ability was normal.

2.2. Recording signal

Record acoustic signal so of speech sound: collect digital voice signals in a professional recording studio through a computer, microphone, and external sound card, use 16-bit quantization and 22kHz sampling frequency, and transfer the recording content to the computer to be marked. The recording format is a PCM code for the wav file.[3]

2.3 Corpus segmentation and parameter determination

The corpus segmentation is mainly divided into two: first, using Audition software to artificially segment the collected speech samples for the first time. The segmentation is based on hearing recognition, cutting off the noise segment and the invalid segment, and retaining valid speech segments. Secondly, according to the content of the syllable study, the speech fragments of the previous stage are divided into individual speech fragments, and they are saved and named in the corresponding folders according to the order of division.According to the research content, the acoustic parameters to be extracted in this study are mainly the duration of the 12 monophthongs in English and the first and second formant parameters.

3. The Study on the Influence of Tibetan Lhasa Dialect on English Vowel Pronunciation

This article takes Tibetan Lhasa native-speaking as the research object, studies the specific errors caused by the mother tongue when they pronounce English vowels, and discusses the causes of the errors. Specifically, according to the theory of bias analysis, through the research methods of experimental phonetics and simple statistical methods, find the pronunciation errors of English vowels by the speaker, and finally find the problem through comparison under the guidance of comparative analysis theory.

3.1 Statistics of Pronunciation duration of English Vowels in Lassa Dialect

Statistics for the pronunciation duration of all the unit pronunciations of native Tibetan Lhasa English speakers is in milliseconds. For comparison and discussion, the following table will count the average duration of each vowel in the High Score Group (HSG) and Low Score Group (LSG). It is summarized with standard values (SV), and each set of data is averaged. The standard mean of native speakers comes from the article by James M Hillenbrand (1995).

Table 1 High and Low Grouping and Standard Values (Unit of Time: Milliseconds)

Monophthong	[iː]	[ɪ]	[e]	[æ]	[ɑː]	[ɔ]	[ɔː]	[u]	[uː]	[ʌ]	[ɜː]	[ə]
HSG	139	126	141	138	166	175	151	129	164	122	175	106
LSG	105	93	105	102	125	104	136	125	133	104	148	96
SV	275	215	294	305	295	296	318	221	270	207	292	222

In order to more intuitively show the difference between the length of English vowels of Lhasa native speakers and English native speakers, [4]the above table is now converted into the following histogram (in milliseconds):

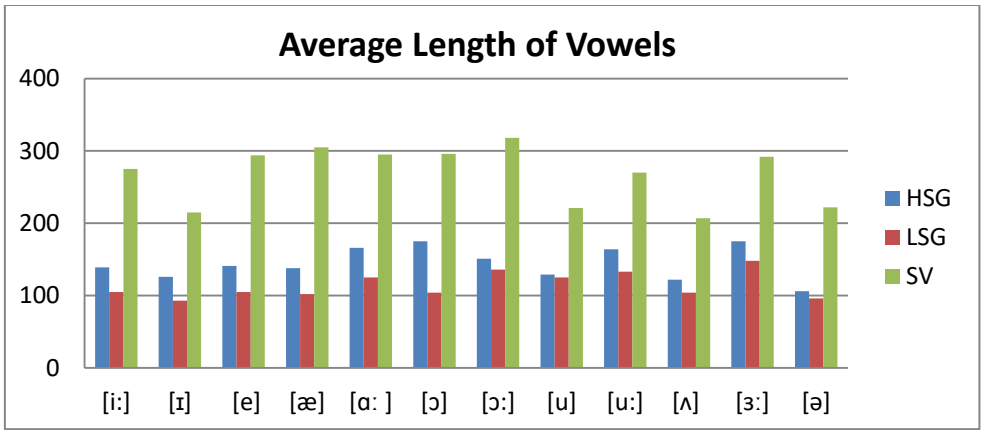

Fig.1 Average Length of Vowels

The above picture clearly shows that the average duration of each vowel of the speaker is much shorter than that of the native speaker, whether it is a high or low group, or even less than half of the standard value. In addition, from the standard value, there are 8 sounds in English vowels that are significantly longer than the other four sounds. However, looking at the length of English vowels of Lhasa native speakers, it is found that Lhasa native speakers (whether high or low) do not reflect this difference in their pronunciation. It can also be clearly seen from the above figure that the average duration of all vowels in the low group is less than the average duration of the speaker in the high group, that is, “the performance of the high group in the vowel duration is better than the low group”, which One thing is as expected. [5]

This study was recorded in a standard voice recording studio. There were 30 speakers and 21 effective recordings were obtained. In this article, the recordings of the top 15 speakers are selected as the research object. Analyze it through Praat, and use the data analysis function of Exce, we put the date F1 and F2 of each vowel into an acoustic vowel diagram with the horizontal axis as F1 and the vertical axis as F2, this method is based on the scatter plot from James M Hillenbrand's paper. Our purpose is to obtains the standard categories of each vowel (each ellipse in the figure below corresponds to the category of a vowel of the native speaker), which helps to judge the pronunciation of the speaker. In the statistical process, points that fall into the category of the native language of the corresponding vowel are correct pronunciations, and those that do not fall into are the wrong pronunciation. Scattered points at the boundary of the category are counted as a correct sound only if they enter more than half of the category; On the contrary, the date will be counted as a wrong sound if they are less than half. If it is exactly at 1/2, it is counted as the correct sound.

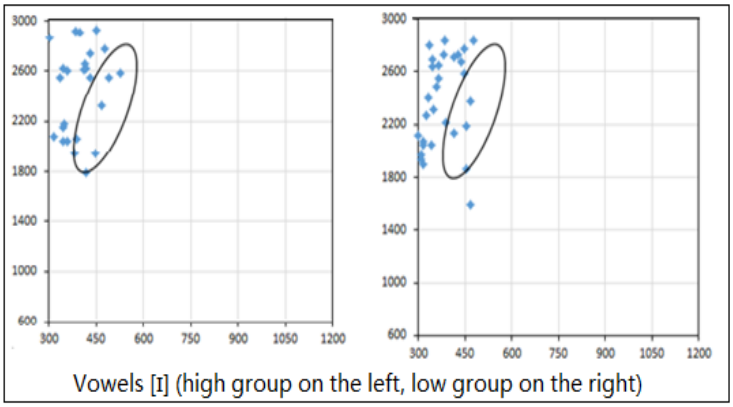

Fig.3 L [ɪ] Pronunciation Category Diagramy

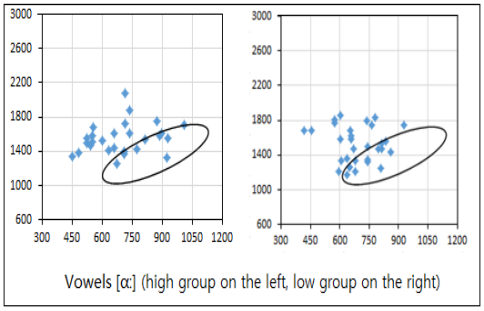

Figure 3. Vowel [ɑ:] Pronunciation Category Diagramy

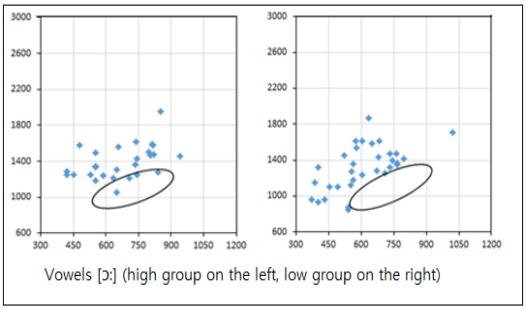

Fig.5 Ls [Ɔ:] Pronunciation Category Diagramy

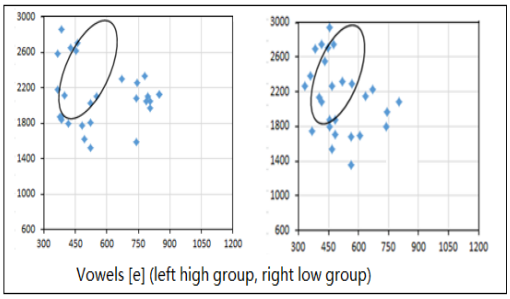

Figure 5. Vowels [e] Pronunciation Category Diagramy

The above shows four groups of vowels with large pronunciation problems. After analysis, the pronunciation accuracy of 12 groups of monophthongs is counted. Among them, (HSG) means high group and (LSG) means low group. Percent Correct (PC) indicates the correct rate. The correct rate of the English vowels of the Tibetan Lhasa native speakers is as follows:[6]

Table 2 Grouping Statistics Of Vowel Correctness of Native Speakers of Tibetan Lhasa Dialect

Vowel	HSG		LSG	
[i:]	PC(%)	70	PC(%)	56
[ɪ]	PC(%)	18	PC(%)	12
[e]	PC(%)	17	PC(%)	22
[æ]	PC(%)	56	PC(%)	49
[ɑ:]	PC(%)	30	PC(%)	37
[ɔ]	PC(%)	21	PC(%)	24
[ɔ:]	PC(%)	13	PC(%)	0
[u]	PC(%)	18	PC(%)	31
[u:]	PC(%)	81	PC(%)	39
[ʌ]	PC(%)	23	PC(%)	29
[ɜ:]	PC(%)	24	PC(%)	35
[ə]	PC(%)	70	PC(%)	63

The data shows that the native speakers of Lhasa have a good grasp of the English vowels: [i:], [u:] and [ə], both the high and low groupings show a good grasp of these three sounds. Secondly, for the sound of [ɑ:], the pronunciation accuracy of the low group exceeds of the high group. The

worst pronunciation of vowel is [ɔ:], both the high group and the low group has the same results of the lowest correct rate, and the low group has the correct rate of 0. From the analysis of this set of data, there is no tendency to show whether the pre-vowels, central vowels, or the post-vowels have better pronunciation, and each group has a phoneme with poor pronunciation and better pronunciation. In addition, as can be seen from Table 2, the vowel accuracy does not appears as higher groups are better than the low groups.

4. Conclusion

According to the “phonetic fossilization theory” (interlanguage: if the new language is not accessible before the age of 7, it will be difficult for the learner to correctly acquire the pronunciation, which is precisely due to the pronunciation habits of the mother tongue Caused by “solidification” in the brain and oral cavity.) enlightenment: As native speakers of Lhasa who started to learn English over the age of 8, and had insufficient contact with English before puberty, which caused their certain kind of oral muscle articulation patterns and habits solidify, it is difficult to acquire correct English pronunciation. Judging from the length of vowel pronunciation, there are long and short vowel phonemes in English monophthongs, while Tibetan Lhasa native speakers have little difference in length when they pronounce long and short vowels, making it difficult to distinguish between long and short vowels. From the perspective of the front and back dimensions of the tongue position, there are only the three vowels [i:], [u:] and [ə] are relatively well pronounced, also the front low vowel and the back low vowel pronunciation accuracy are very low.

Acknowledgment

The successful of this paper was subsidized by Central Universities Fundamental Research Funds for the project: A study of English phonemes of Lhasa Tibetan college students based on perception (31920190110), and Ethnic Research Project of the State People's Committee: A Comparative Study on English Learning among College Students from Minority Nationalities (2019-GMD-066).

References

[1] Best, C. T. A Direct Realist View of Cross-Language Speech Perception. Strange, no.1, pp. 171-204, 1995.

[2] Teaching English Intonation to EFL/ESL Students. [online] Available: http://iteslj.org/Techniques/Celik-Intonation.html.

[3] Debaene, M. Perceiving and Producing Native and Non-native Vowels- An Experimental Study on the Effects of First Language Regional Variation, University of Ghent, 2013.

[4] Flege, J.E. The Relation Between L2 Porduction and Perception. In Proceedings of ICPhS99 (San Francisco), no.12, pp.1273-1276, 1999.

[5] Debaene, M. Perceiving and Producing Native and Non-native Vowels- An Experimental Study on the Effects of First Language Regional Variation, University of Ghent, 2013.

[6] Lin, C.Y. Perception and Production of Five English Front Vowels by College Students. English Language Teaching, vol. 7, no. 9, pp. 14-20, 2014.