# The Study of the Salient Vowel Mispronunciations of Tibetan Adult English Learners by Means of Experimental Phonetics

Qian Zhou[+]

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

**Abstract.** Assisted by the computer software "Praat", this paper studied the salient vowel mispronunciations of Tibetan adult English learners by checking whether the F1(First Format)and F2 (Second Formant)of the pronounced vowel of the tested speaker can fit into the corresponding vowel cluster and by comparing the duration of the vowels with the standard mean time length. This innovative method is a great improvement in finding the vowel mistakes, as the difference among vowels are very subtle and the traditional judgment methods （the judgement of right and wrong is based on human listening are very subjective.

**Keywords:** Vowel Mispronunciations, Vowel Cluster, Duration

## 1. Introduction

Generally speaking, Language has three sorts of substructures: phonology, morphology and syntax[1]. Among them, phonology is the means of expression employed by language in daily verbal communication and is the basics of language.

In second language learning the most important and difficult thing is to learn the pronunciation of that language[2], and also pronunciation teaching is the most important. Especially, as English is an alphabetic writing system, great relations between pronunciation and writing are formed in English. A good grasp of pronunciation can certainly promote the learning of vocabulary and grammar, and to facilitate speaking and listening skills, and by doing so, to fulfill the communication goal in language teaching. Above all, as far as we can see, pronunciation does have a significance in foreign language teaching of English [3].

Tibetan people is one of the 56 nationalities in China and it has 7 million people in China alone. They have their own language and writing systems which is called Tibetan. Tibetan language belongs to the Tibetan branch, Tibetan-Burmese group, Chinese-Tibetan family. There are three dialect areas, U-Tsang, Kham and Amdo[4] and they are greatly different. For this research, the subjects are the Amdo speakers --- which means the mother language of the selected speaker for this study is Amdo dialect.

As for Amdo and English, English has 20 vowels, among which 5 long vowels, 7 short vowels and 8 diphthongs. But Amdo has only 6 vowels: [i], [e], [a], [ə], [o] and [u] and no long vowels[5], this paper is going to identify the salient mispronunciations by means of experimental phonetics. The past researches or discussions about vowel mispronunciations are usually auditory judgements or just based on researchers' teaching experience, which is very subjective and has lack of accuracy.

## 2. The Design of the Experiment

This paper is analyzing the salient vowel mispronuncia tions of Tibetan adult English learns by means of experimental phonetics. The pronunciation of vowels of the selected subjects will be analyzed by "praat"

---

[+]   Corresponding author. Tel.: + 13609392562.
     *E-mail address*: smiling424@126.com.

(praat is a linguistic tool for studying sounds and is an often used software by language pronunciation experimenters.) and the data of "first formant" (F1) and "second formant" (F2) will be taken out from Praat's sound spectrogram for comparison and analysis.

As for vowels, the first formant and second formant can determine the quality of the vowel[6], and are very important experimental data. The data taken out from this experiment will be compared with James M Hillenbrand [7]' data from "Acoustic characteristics of American English vowels", which is published on The Journal of the Acoustical Society of America in June 1995. This study was a replicate and extension of Peterson and Barney (PB)'s classical experiment [8], but with more talkers, better equipment and more sophisticated experiment plans. In their study, the F1 and F2 of 10 English vowels (i, ɪ, ɛ, æ, a, ɔ, U, u, ʌ, ɜ ) were measured and were presented on the Formant Frequencies chart along with ellipses fit to each vowel category. Like the following:
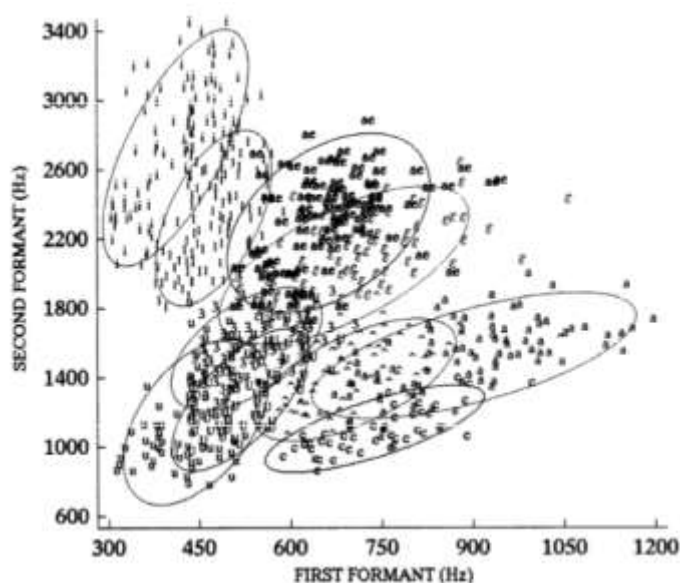


Fig.1: Vowel (i, ɪ, ɛ, æ, a, ɔ, U, u, ʌ, ɜ ) clusters for men, women, and child native talkers

In this study these 10 vowels (i, ɪ, ɛ, æ, a, ɔ, U, u, ʌ, ɜ ) are also the testing targets. If the F1 and F2 of the tested vowel of the speakers fall into the ellipse of the vowel cluster in Fig. 1, the pronunciation of this vowel of the talker would be considered as standard and vice versa.

And also, as Amdo language doesn't have long vowels, so we predict that there might be a duration confusion between long vowel category and short vowel category. In this study, we will measure the vowel duration of Amdo speakers and make comparison with the native speaker.

## 3. The Preparation of the Experiment.

### 3.1. Subjects

Four talkers, all university students, at the same age (22 years old), with the same dialect of Amdo and at the same level of English have been selected through a screening test to take this pronunciation experiment. The screening test is an oral test (read a passage of 121 words) judged by Pronunciation expert. And one more important thing is that they all have received 5 or 6 years of formal English education and their oral English tests are all graded level B.

### 3.2. Reading list

The two famous studies which measured the quality of English vowels all used the /h-d/ phonetic context to form words and use them as the reading list. But this study did not take that. Because half of the words of their word list are not common words or are made-up words, they are difficult for non-native speakers to pronounce. In order to avoid the mispronunciation of the vowels were caused by the wrong-guessing of the pronunciation of the words. The vowels checked should be put in common and simple words but not in /h-d/ context. And also, in order to make the influence of the adjacent consonants less, the syllable of the words

should be as little as possible. And the word should better be a CVC (Consonant-Vowel-Consonant) phonetic structure word.

# 4. About the Experiment

## 4.1. Recordings

The word list would be given to the students 20 minutes before the recordings for them to practice and get familiar with, but without any guidance or tutoring. And each word on the reading list will be read twice, when make the recording.

The recordings were made with a digital audio recorder and a microphone at standard voice recording lab in Northwest Minzu University in China. The recording was sampled at 16 kHz with 12 bits of amplitude resolution on a standard computer and saved as wav. type document.

## 4.2. Data collecting

### 4.2.1. Data collection of F1 and F2

The F1, F2 and duration of each vowel of each speaker was measured by hand and with the facilitation of "Praat". The detailed steps of operation are as following:

When one word is read into "Praat", press the "show Formant" button, we can clearly see five formants marked respectively by five red lines on the spectrogram. The red line at the bottom is F1 and the one above it is F2. The data of F1 and F2 are taken from the middle and the most steady state of "the red line". These two figures are the data which should be collected in this study.

### 4.2.2. Data collection of vowel duration

Vowel duration: the starting and ending times of each vowel were measured by hand from high-resolution gray-scale digital spectrograms and by using standard measurement criteria by Peterson and Lehiste,1960[9].

## 4.3. Data

The following table are the data collected according to the above mentioned experiment method.

Table 1: The formant data and duration data of four selected speakers. Dur=Duration

| | Speaker1 | | | Speaker2 | | | Speaker3 | | | Speaker4 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | F1 | F2 | Dur | F1 | F2 | Dur | F1 | F2 | Dur | F1 | F2 | Dur |
| i | 479 | 2498 | 0.290 | 413 | 2465 | 0.129 | 313 | 2366 | 0.344 | 413 | 2796 | 0.154 |
| ɪ | 380 | 2697 | 0.292 | 313 | 2564 | 0.191 | 313 | 2267 | 0.332 | 479 | 2763 | 0.190 |
| ɛ | 876 | 2068 | 0.328 | 810 | 2068 | 0.254 | 479 | 2763 | 0.197 | 644 | 1969 | 0.225 |
| æ | 777 | 2167 | 0.164 | 777 | 1969 | 0.192 | 677 | 1902 | 0.260 | 942 | 2035 | 0.153 |
| a | 909 | 1406 | 0.281 | 810 | 1207 | 0.461 | 711 | 1108 | 0.380 | 1042 | 1572 | 0.190 |
| ɔ | 790 | 1307 | 0.209 | 711 | 1141 | 0.203 | 876 | 1493 | 0.172 | 810 | 1207 | 0.152 |
| U | 611 | 1406 | 0.284 | 479 | 1241 | 0.149 | 512 | 1671 | 0.145 | 512 | 1108 | 0.111 |
| u | 545 | 1075 | 0.444 | 380 | 810 | 0.328 | 290 | 990 | 0.275 | 512 | 843 | 0.228 |
| ʌ | 1009 | 1770 | 0.187 | 909 | 1505 | 0.170 | 843 | 1340 | 0.164 | 1075 | 1836 | 0.133 |
| ɜ | 744 | 1307 | 0.167 | 677 | 1538 | 0.147 | 578 | 1439 | 0.192 | 545 | 1075 | 0.254 |

# 5. Data Analysis

## 5.1. The analysis of accuracy of vowel quality.

The data in the above table is further processed by excel to fit in to the chart of James M Hillenbrand (see fig.1 of this passage), to see whether the tested vowel can fit into the ellipses (cluster) of the corresponding standard vowel and the results are established in the following diagram.
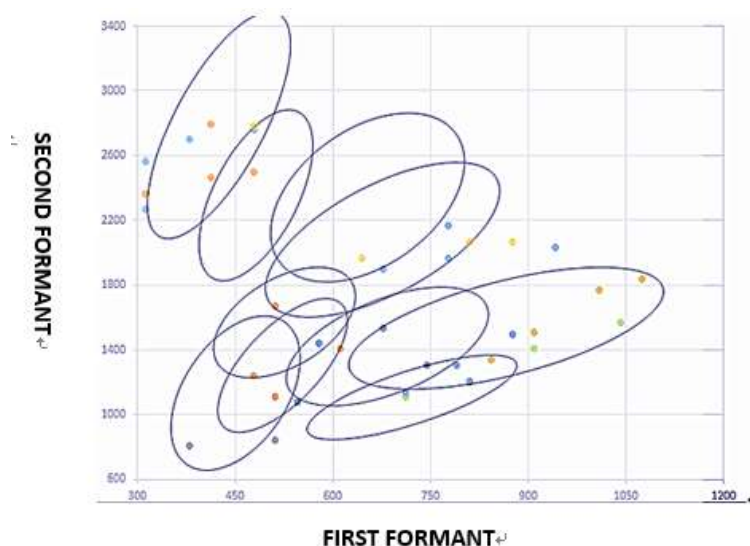
Fig. 2: The data of F1 and F2 of the tested speaker are put into the chart of Fig.1.

If the F1 and F2 of the tested vowel of the speakers fall into the ellipse of the vowel cluster in Fig. 1, the pronunciation of this vowel of the talker would be considered as right and standard and vice versa. From this chart, we can see that the accuracy rate of the ten vowels checked in this passage is like the following:

i:75%；ɪ:25%；ɛ: 50%；æ:0%；a:75%；ɔ:75%；U:75%；u:25%；ʌ:0%；ɜ:25%

## 5.2.    The analysis of vowel duration.

The previous study [7] has shown that the mean duration of vowels between male and female are quite different. Generally speaking the mean duration of female are longer than male for each vowel. So, in this study, the comparation of vowel duration is also separated by male and female.

Table 2: S1=Speaker1; S2=Speaker2; S3=Speaker3; S4=Speaker4. F in the brackets means female speaker; M in the brackets means male speaker; The yellow highlighted figures are the figures of the long vowels (When pronounce, these vowels should be longer than the rest.); The figure of the standard duration comes from James M Hillenbrand (1995).

|   | S1(F) | S2(F) | S3(F) | Mean Dur | Standard(F) | S4(M) | Standard(M) |
|---|-------|-------|-------|----------|-------------|-------|-------------|
| I | 0.290 | 0.129 | 0.344 | 0.254 | **0.306** | 0.154 | **0.243** |
| ɪ | 0.292 | 0.191 | 0.332 | 0.272 | **0.237** | 0.190 | **0.192** |
| ɛ | 0.328 | 0.254 | 0.197 | 0.197 | **0.254** | 0.225 | **0.189** |
| æ | 0.164 | 0.192 | 0.260 | 0.241 | **0.332** | 0.153 | **0.278** |
| A | 0.281 | 0.461 | 0.380 | 0.374 | **0.323** | 0.190 | **0.267** |
| ɔ | 0.209 | 0.203 | 0.172 | 0.195 | **0.353** | 0.152 | **0.283** |
| U | 0.284 | 0.149 | 0.145 | 0.193 | **0.249** | 0.111 | **0.192** |
| U | 0.444 | 0.328 | 0.275 | 0.349 | **0.303** | 0.228 | **0.237** |
| ʌ | 0.187 | 0.170 | 0.164 | 0.174 | **0.226** | 0.133 | **0.188** |
| ɜ | 0.167 | 0.147 | 0.192 | 0.169 | **0.321** | 0.254 | **0.263** |

From the above table we can see that like the native speakers the duration of vowels of male speaker is much shorter than the female speakers.

For the vowels, like: i, æ, a, ɔ, u, ɜ which suppose to be longer than the rest, both the female and male speaker grasped the duration of a and u, but for the rest, all appears to be apparently shorter than normal and standard mean duration.

For the vowels like: ɪ, ɛ, U, ʌ, which all suppose to be shorter, all the speakers of the experiment seems to do them right.

Generally speaking, the Tibetans kind of mixed between the long vowels and short vowels and they tend to treat most of the long vowels short. This is probably caused by the influence of their mother language ------ Amdo, as in Amdo language, there is no distinction between long vowel and short vowel.

## 6. Conclusion

From this analysis, we can see that the grasp of pronunciation of native Amdo speakers need a lot of

improvements. The teacher and the learners should pay special attention to the pronunciation of ɪ, ɛ, æ, u, ʌ, ɜ, as the accuracy rate of these vowels are all under 50% percent (ɪ:25%, ɛ: 50%, æ:0%, u:25%, ʌ:0%, ɜ:25%). And After this the pronunciation to i:75%, a:75%, ɔ:75%, U:75%, should also be guided and corrected. Also, great attention also should be payed to the duration of the vowels, especially the duration of long vowels. By doing so, their pronunciation will be improved quickly and greatly.

## 7. Ackonledgements

## 8. Reference

[1]  Lindao, Wanglijia. 2013. Phonetics Beijing: Beijing University Press,2013, p5

[2]  Yuanren Zhao, Language Study. Beijing: Commercial Press, 1980

[3]  Ninghong Peng, The Study of Learning Policies between English Majors and Non-English Majors[J]. Foreign Language Teaching Study, 2014(2)

[4]  https://baike.baidu.com/item/%E8%97%8F%E6%97%8F/130324?fr=aladdin

[5]  Deng guangyu, Hua kan, 1995. The discussion and analysis of English pronunciation of Tebitans from Gansu and Qinhai. [J]. The Academic Journal of Northwest Minzu University 1995(02):74-78.

[6]  Dao Lin, Jiali Wang. Phonetics. 2edition. Beijing: Peiking University Press, 2013.08

[7]  Hillenbrand J1, Getty LA, Clark MJ, Wheeler K, 1995.Acoustic characteristics of American English vowels, The Journal of the Acoustical Society of America, June 1995

[8]  Peterson, G. E. and Barney, H. I. (1952). "Control methods used in the study of vowels", Journal of the Acoustical Society of America, 24, 75-184.

[9]  Peterson,G.E. and Lehiste, I.,(1952). "Control methods used in a study of the vowels," J.  Acoust. Soc. Am. 24, 175-184.