Application of Experimental Phonetic Technology in Acoustic Study of Putonghua Retroflex Vowels Acquisition

Hongli Deng[1], Xinzhong Liu[2], Xianming Bei[3], Jingjing Wang[4]

{361389651@qq.com[1], liuxinzhongjn@163.com[2], xianmingbei@163.com[3], wangjj1204@126.com[4]}

[1]School of Liberal Arts, Jinan University, Guangzhou, Guangdong, China, College of Culture and Communication, Guangxi Science and Technology Normal University, Laibin, Guangxi, China

[2]School of Liberal Arts, Jinan University, Guangzhou, Guangdong, China, Corresponding author

[3]School of Chinese Language and Culture, Guangdong University of Foreign Studies, Guangzhou, Guangdong, China, Corresponding author

[4]College of Vocational and Technical Education, Guangxi Science and Technology Normal University, Laibin, Guangxi, China, Corresponding author

Abstract. The article conducts acoustic experiments on the pronunciation of the “er” which means “two” in Chinese by learners from different language background by referring Praat speech analysis software and scripts, and analyses acoustic parameters of the slope, F1, F2, F3, difference by combined with the five-level scoring method and one-way ANOVA method.The study shows that as for the “er” pronunciation of learners who speak in S&P, their F2 of retroflex vowels rises, and F3 falls, and the difference value between F3 endpoint and F2 endpoint is small, which means that F3 and F2 are closer to each other. The key characteristics to the learner's “er” pronunciation acquisition depend on the slope of F3 and the difference value of F3. The greater the slope of F3 falls, the smaller the difference value, which indicates the “er” pronunciation of learners from different language background will draw near to S&P “er”.

Keywords. experimental phonetic; retroflex vowels; slope of F3; acoustic characteristics

1 Introduction

1.1 Existing researches on retroflex vowels in S&P

Retroflex vowel of Putonghua has the property of compound vowels(Zhou Dianfu & Wu Zongji 1963). [1] Retroflex vowel of Putonghua is R-sound.they have in common is that they are acoustic structures rather than tonal modes - the third formant usually drops (Peter Lagefoged and Maddison 2015, [2]Chen ZhongMin 2022[3]). Sun Guohua (1994) believed that there are two pronunciation patterns for the retroflex vowels “er” in Putonghua.[4] 1. Pronunciation pattern One - The vocal cords vibrate first, then followed by the roll-up movement. The acoustic pattern - F1 rises, F2 falls, and F3 falls; 2. Pronunciation pattern Two - The tongue rolls before the vocal cords vibrate. The acoustic pattern - F1 falls, F2 rises, F3 goes flat. Although in the latter one we can not see the situation that F3 decreases in the time

ITEI 2022, December 23-25, Harbin, People's Republic of China

DOI 10.4108/eai.23-12-2022.2329195

domain, the starting target value of F3 is lower than the latter's starting target value of 300Hz compared with the “er” sound which has a downward trend. Bei Xianming (2021) believed that the acoustic characteristic of retroflex vowels is that F3 decreases rapidly, and moved towards F2 rapidly, which made the value of F3-F2 smaller.[5] Liu Xinzhong, etc (2022) believed their direct observation of the pronunciation of the “er” syllable is that first the tip of the tongue is raised upwards and then moves backwards; Then the soft palate sounds close the nasal channel, while the channels of the oral cavity and pharynx remain open. This is the typical “er” syllable (as shown in Figure 1). The three corresponding acoustic cues for the “er” syllable are that F3 decreases, F2 rises slightly, and the resulting formants F2 and F3 move closer to each other.(as shown in Figure 2) [6]. It can be seen that scholars basically agree that retroflex vowel has the property of compound vowels, and their acoustic characteristic is F3 falling and F2 rising, and the two move closer to each other.

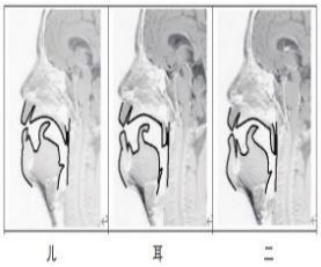

Fig. 1. Three-syllable interface MRI vector diagram of the o move closer[6]

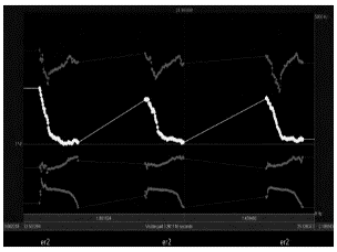

Fig. 2. The first 4 formants of the three syllables of the “er” syllable[6]

Yu Jue,etc (2008) believed that the main vowel of “二” is [ɐ], and F3 falls.[7] There are few studies have been conducted on the acquisition of retroflex vowels in Putonghua, Although Yu Jue, etc (2008) [7], Yu Jue, etc (2020) [8]studied the acquisition of retroflex vowels in Putonghua, but they failed to conduct cross-language comparisons to analyze the key acoustic performance of the acquisition of Putonghua retroflex vowels in multiple native-language contexts. so this paper will use phonetic experimental technology to analyze the acoustic characteristics of Putonghua "er" in six native language backgrounds.

1.2 Theoretical basis

Based on phonetic experimental techniques, this paper analyzes the acoustic performance of retroflex vowels in learners from six native language backgrounds. The theoretical basis is experimental phonetics. Experimental phonetics is a discipline that uses various experimental

instruments to study and analyze phonetic, and is an important branch of linguistics. The study of experimental phonetics includes physiological phonetics, acoustic phonetics, and perceptual phonetics. This paper focuses on acoustic phonetics, which mainly analyzes acoustic waveforms that record the pressure fluctuations of retroflex vowels over time. The formant is the basic acoustic parameter of a vowel. F1 is proportional to the high and low positions of the vowel tongue, and F2 is related to the anteroposterior position of the vowel tongue, and F3 is related to the retroflex position.

2 Experiment introduction

2.1Experiment method

2.1.1 Phonetic experiments

By using Praat speech analysis software and scripts written by Xiong Ziyu (2016), the author first annotates the retroflex vowels in the recording samples, and extracts the starting point values, midpoint values and endpoint value of the first three formants F1, F2, and F3. Formula (1) can calculate the slope value of the first half and the second half of F3. Where （Xa, Ya) and (Xb, Yb) are the coordinates of any two points in the plane, and K is the slope of this straight line.

$$k = \frac{y_{\mathrm{a}} - y_b}{x_a - x_b} \tag{1}$$

y2 in the first half of F3 is the midpoint value of F3, y1 is the starting point value of F3, x2 is the time point of the midpoint value, x1 is the time point of the starting point value, and y2 in the second half is the ending point value of F3, y1 is the midpoint value of F3, x2 is the time point of the ending point value, and x1 is the time point of the midpoint value. We set the slope of the first half as k1 and the slope of the second half as k2, so the k1 of F3 = (F3 intermediate value - F3 start value)/(the time point of the midpoint value - the time point of the start value), k2=(F3 end value - F3 midpoint value)/(the time point of the end value - the time point of the midpoint value). It is the same as the calculation to F2. In order to reduce interpersonal discrepancies, it is convenient to make a comparison among different people with different native languages.

Formula (2) (Schroeder, Ataland Hall 1979)convert the resonant peak to the Bark value, which can be used as the scale of the acoustic vowel diagram, and it is an auditory equivalent value. Among them, f represents the rate of the resonance peak. Then use formulas (3) and (4) to convert to V values

$$Bark = 7\ln(\frac{f}{650} + \sqrt{(\frac{f}{650})^2 + 1}) \tag{2}$$

This formula shows the coordinate value of each vowel in the standard graph of the vowel pattern. In standard graphics, V1 represents the level of vowels, and the value range is [0, 100]. 100 means the lowest tongue position, and 0 means the highest tongue position. V2

indicates that front and rear of the tongue position, the range is [0, 100]. 0 indicates the most rear tongue position, and 100 means that the most front tongue position.

$$V1 = 100 * \frac{B1_x - B1_{min}}{B1_{max} - B1_{min}} \quad (3)\ [9]$$

$$V2 = 100 * \frac{B2_x - B2_{min}}{B2_{max} - B2_{min}} \quad (4)\ [9]$$

when calculating V1 value, the recorder's lingual vowels also participate in the normalization. B1x refers to the retroflex vowels of B1 value. And B1min, and B1max are respectively the minimum and maximum values of B1 in lingual monophonic and retroflex vowels. It is true of the V2 value calculation. " In this paper, the author plots the V value of the retroflex vowel into a reftrolex vowel acoustic vowel map. Since there are many levels for some native language background, the acoustic vowel map will present more points. To make the diagram clearer, we did not include the lingual vowels in the diagram. The slope values and discrepanciess of K, K1 and K2 of F2 and F3, as well as the V values of F1 and F2, can be referred to in the appendix.

2.1.2 Five-level scoring method

On the basis of the Likert scale, the article sets 5 distance scales between the “二” of learners with different native language backgrounds and the “二” of the S&P. A means "completely consistent"-5 points; B means "very close"-4 points; C means "relatively close"-3 points; D means "not very close"-2 points; E means "completely inconsistent"-1 point. We invite 15-20 students whose native language is a northern S&P dialect and whose Putonghua proficiency is 2-A and above to score the retroflex vowels “二” that we have recorded. If the students think that the sound they hear is exactly the same as that of S&P, select "Completely consistent". Here I just want to mention that because the learners from Bangladesh and Pakistan are beginners, some students cannot clearly distinguish the tones of the “er” sound in those Chinese characters which contain it. In order to avoid the influence of pitch deviation, this experiment only examines the situation of their retroflex vowels. By referring to the tone value of "two" in S&P, we changed the tone of the Mengba students to a falling tone through the Praat software, which is convenient for students to evaluate the pronunciation of the retroflex vowels.

Also, we invite the national Putonghua testers to record the audition, meanwhile, relies on the use of one-way ANOVA to test the accuracy of learners' pronunciation. In order to ensure the quantities of the data sample, Laibin, Cantonese and Vietnamese male and female language learners were not separated by gender for univariate analysis. There are only male language learners in Bangladesh and Pakistan, and only female learners in Thailand, so an ANOVA was performed separately for male and female language learners in S&P by gender.

2.2 Experiment objects

There are 21 volunteer students from Laibin in Guangxi, including 4 male students and 17 female students, 5 students of whom get Certificate of Putonghua Proficiency Level 2-A, 9

Level 2-B, and 7 Level 3-A. There are 4 male students and 5 female students from Guangdong province, among of whom 3 students with 2-A, and 6 students with 2-B. Vietnamese students are all from northern Vietnam, including 12 males and 9 females, 7 of whom are with primary level Chinese proficiency with 2 months Chinese learning experience; 8 of whom are with mid level Chinese proficiency with 2 years Chinese learning experience; 6 of whom are with senior level Chinese proficiency with 4 years Chinese learning experience; There are 12 female students from Thailand with mid level Chinese proficiency with 2 years Chinese learning experience. There are 10 male students from Bangladesh and 6 male students from Pakistan with primary level Chinese proficiency with 2 months Chinese learning experience; In this paper, the Putonghua of the native language learners is referred to as Laipu (it means Putonghua with Laibin accent), the native language of Cantonese learners is abbreviated as Guangpu (it means Putonghua with Cantonese accent), the native language of Vietnamese learners is abbreviated as Yuepu(it means Putonghua with Vietnamese accent), and the native language of Thai learners is abbreviated as Taipu(it means Putonghua with Tai accent). Pakistani learners whose native language is Urdu are abbreviated as Bapu(it means Putonghua with Pakistan accent), and learners whose native language is Bengali are abbreviated as Mongpu(it means Putonghua with Bengali accent).

2.3 Experiment corpus

The sound was recorded by using FiFeng software with the Sound BlasterX-FiSurround 5.1 Pro sound card and microphone recording. We collected the recordings of "二" sound from 4 people who speak S&P and 79 second language learners. The author selected the better pronunciation recording samples and picked out 219 samples of "二" in total. In order to better calculate the V value of "二", we also recorded the lingual unit sounds a, o, i, u, y, e, and ê of each recorder.

3 Analysis of experimental results

3.1 Analysis on S&P retroflex vowel

As it can be seen from Figure 3-4, Appendix and Table 1, according to V value to determine the position of retroflex vowels from Shi Feng and Shi Xiujuan (2007), [9]the retroflex vowels are roughly in the range of the lower central vowel and it is a compound vowel. As it can be seen from Figure 2-3 and Table 1, the F3 slope of S&P are all negative values, and the decline is sharper, while the F2 slopes are all positive values, so the F2 is rising. The difference between men and women are very closer to each other.

Table 1. List of One-Way ANOVA of the Retroflex Vowels of the Learners and the Retroflex Vowels of the S&P

	F1			F2			F2			F3			Difference
							K	K1	K2	K	K1	K2	
level	Starting point	Midpoint	endpoint	Starting point	Midpoint	endpoint	discrepancy						

LB	+	+	+	-	+	+	-	+	+	+	+	+	-
LC	+	-	+	+	+	+	+	-	-	-	-	-	-
LD	-	-	+	+	+	-	-	-	+	-	-	-	-
LE	-	-	+	+	-	-	-	+	+	-	-	-	-
GB	+	+	-	+	+	-	+	+	-	+	+	+	+
GC	-	-	-	+	+	-	-	-	+	-	+	+	-
GD	-	-	-	+	+	-	-	+	-	-	-	-	-
GE	-	-	+	+	+	+	-	+	+	-	-	-	-
YB	+	+	+	-	+	+	+	-	+	-	+	-	-
YC	-	-	-	+	+	-	-	-	+	-	-	-	-
YD	-	-	+	+	+	-	-	+	+	-	-	-	-
YE	-	-	-	+	+	-	-	+	-	-	-	-	-
TD	-	-	-	-	+	-	-	-	-	-	-	-	-
JC	+	-	-	-	-	+	-	+	-	-	+	-	-
JD	-	-	-	-	-	-	+	+	+	-	-	-	-
JE	-	-	-	-	-	-	+	+	+	-	-	-	-
MB	+	+	-	+	-	+	+	+	+	+	+	+	+
MC	+	-	-	+	+	+	+	+	+	-	+	-	-
ME	+	-	-	-	-	+	+	+	-	-	-	-	-

Note. In the table, S means S&P, L means Laipu, G means Guangpu, Y means Yuepu, T means Taipu, J means Bapu, M means Mengpu, B, C, D, E are S&P levels. "+" indicates that there is no discrepancy between the learner and the S&P; "-" indicates that there is a discrepancy. It is consistent with the following, so the author will not repeat it.

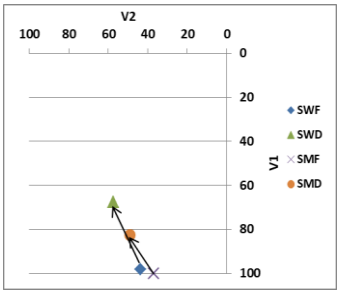

Fig. 3. S&P Acoustic Vowel Plot

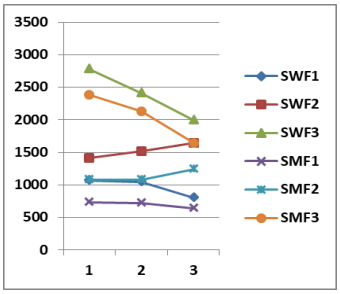

Fig. 4. S&P Formant Plot

Note. W is female, M is male, F is the starting point, D is the ending point. It is consistent with the following, so the author will not repeat it.

3.2 An analysis of the retroflex vowels of the learner's S&P

3.2.1 Analysis on Laipu's retroflex vowels

Based on the experiment result, we divided students' retroflex vowels into four levels: B, C, D, E. There still remain certain gaps with S&P. In terms of acoustic sense, it can be regarded as a "pronunciation defect" according to the S&P. Level E can obviously be judged wrong from the perspective of acoustic sense. According to the Table 1, Appendix and Figures 5-6, the position of retroflex vowels and tongue sliding direction in Level B are almost similar to those of S&P and the discrepancy is significant obvious between Level C, D, E and S&P. According to the ANOVA, there is only one obvious discrepancy between Level B, C and S&P and three obvious discrepancies between Level D, E and S&P.

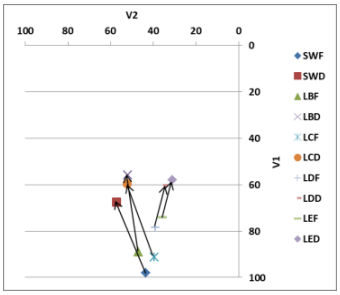

Fig. 5. Laipu female acoustic vowel diagram

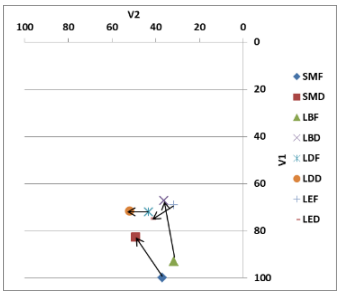

Fig. 6. Laipu male acoustic vowel diagram

It can be seen from Table 1, Appendix and Figures 7-10 that the formant mode of Laipu Level B is the closest one to that of S&P, followed by level C and level D, and level E presents the obvious discrepancy. According to F3 slope ANOVA, the F3 slope of level B is not significantly distinguished from that of S&P; The F3 slopes of level C, D, and E are significantly distinguished from that of S&P; According to F2 slope ANOVA, K value of Level B and E is distinguished from that of S&P. And there are two discrepancies between Level C,D and S&P. Therefore, the author concludes that there exists an obvious discrepancy between four levels and S&P. We can separate different levels by referring the falling slope of

F3. Although the difference is all significantly different from S&P, the higher the rank, the smaller the difference.

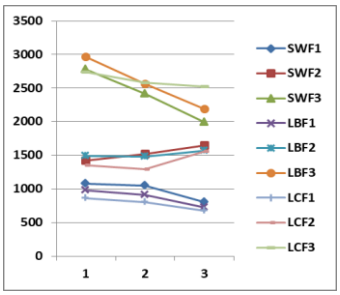

Fig. 7. Female formant diagrams of Laipu level B, levelC

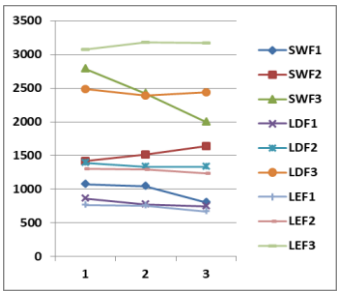

Fig. 8. Female formant diagrams of Laipu level D, level E

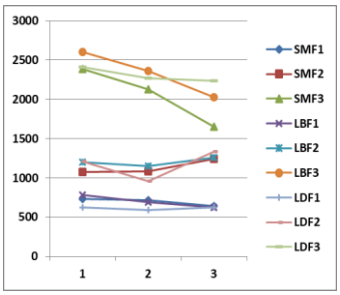

Fig. 9. Formant diagrams of Laipu male level B, level D

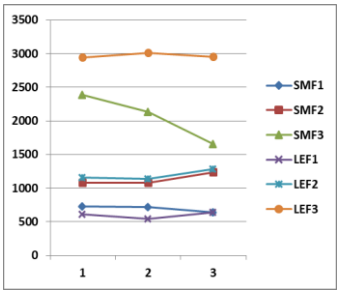

Fig. 10. Formant diagram of Laipu male level E

3.2.2 Analysis on Guangpu's retroflex vowels

According to Table 1, appendix and Figures 11-12, the positions and tongue sliding directions of the retroflex vowels of level B and level C of Guangpu are similar to those of S&P, while the positions and tongue sliding directions of the retroflex vowels of level D and level E of Guangpu are distinguished from those of S&P. According to ANOVA, there exist two obvious discrepancies between Level B and S&P; four obvious discrepancies between Level C, D and S&P; three obvious discrepancies between Level E and S&P.

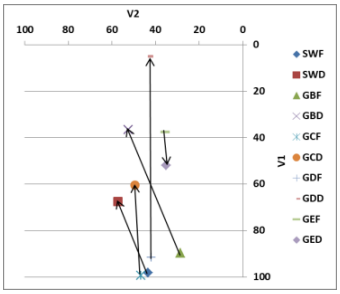

Fig. 11. Vowel diagram of Guangpu female acoustics

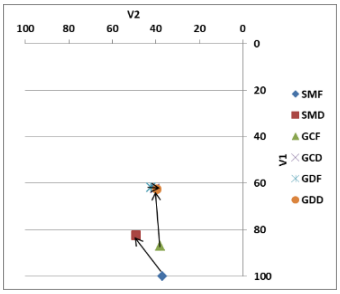

Fig. 12. Vowel diagram of Guangpu male vocals

According to Table 1, Appendix and Figures 13-15, the formant modes of level B are the closest to S&P, followed by level C, level D, and in the end level E. According to F3 ANOVA, there is no discrepancy between Level B and S&P. There is only one discrepancy between Level C and S&P. there is an obvious discrepancy between Level D, E and S&P.

According to F2 ANOVA, there is one obvious discrepancy between Level B, E and S&P; two obvious discrepancies between Level C, D and S&P. And the discrepancy between Level B and S&P is not significant, and the discrepancy between the other grades and S&P is obvious. In conclusion, we can separate different levels by referring the falling slope of F3. It shows that the lower the level, the larger the difference.

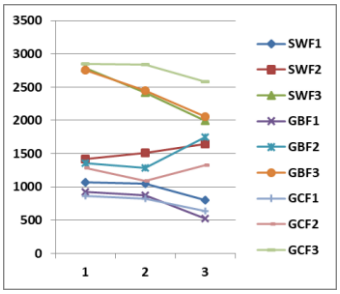

Fig. 13. Formant diagrams of Guangpu female B, C

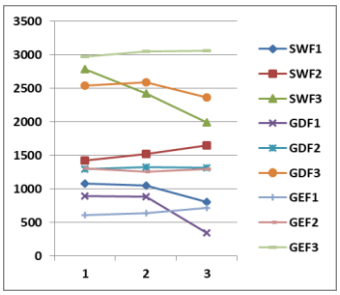

Fig. 14. Formant diagram of Guangpu female D, E

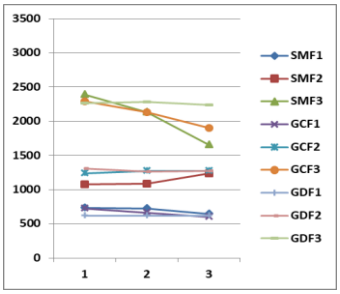

Fig. 15. Formant diagram of Guangpu male C, D

3.2.3 Analysis on Yuepu's retroflex vowels

It can be seen from Figures 16-17 that the retroflex vowels sliding direction is similar to that of S&P. Level B is different from S&P most, followed by Level C and Level D, ending up with Level E. According to ANOVA, there is only one obvious discrepancy between Level B and S&P, four obvious discrepancies between Level C, E and S&P, and three obvious discrepancies between Level D and S&P.

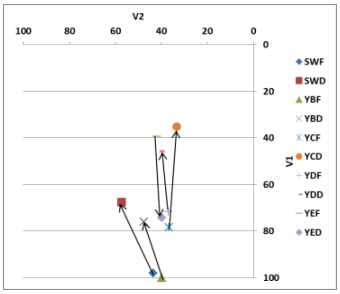

Fig. 16. Vowel diagram of Yuepu females

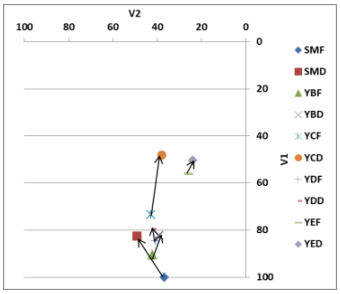

Fig. 17. Vowel diagram of Yuepu males

According to Table 1, Appendix and Figure 18-21,According to F3 ANOVA, there is three obvious discrepancy between Level B and S&P; and the discrepancy between the other grades and S&P is obvious. According to F2 ANOVA, there is one obvious discrepancy between Level B, D and S&P; there is two obvious discrepancy between Level B, D and S&P. Except the large discrepancy between the level C and level D, the other levels show the characteristics of higher levels and smaller discrepancies, which indicates that the gap between F3 and F2 is getting larger and larger. there is three obvious discrepancy between Level B and S&P, which is distinguished from Laipu and Guangpu. After being checked by national S&P testers, it was found that, in Yuepu level B, there are three learners whose pronunciation is the second pattern mentioned by Sun Guohua (1994).[4] For the three female recorders in level B, their starting average value of V3 is 42.37, and its endpoint average value is 0. And the other recorder's starting average value of V3 in level B is 81.85, and its endpoint average value is 0. The starting average value of V3 of S&P is 63.76, and its endpoint average value is 0. The starting average value of Guangpu female recorders level B is 52.9, and the endpoint average value of it is 0. The starting average value of Laipu female recorders level B is 68.84, and the endpoint average value of it is 0. In the second pattern, the starting average value of V3 of the three recorders is lower than that of the other recorder, but V3 shows an overall downward trend, and the decline is smaller than that of the V3 of the level B under other native langauges background, which leads to a significant discrepancy between the level B and K and K2 in S&P. The male Level D recorder roll the tongue at first. The starting average point of F3 is lower, which is almost the same as the target value. But based on the sense of hearing, the other two recorders' pronunciation sounds like the first pattern mentioned by Sun Guohua(1994) with a slight retroflex movement process, but their F3 values don't fall.[4] At present, this situation cannot be explained. Although the slopes of level B, C, and D are

smaller compared to other recorders' slopes from different native language backgrounds, the discrepancy among these levels and S&P is obvious. However, the discrepancy in slopes can roughly reflect the discrepancies among levels.

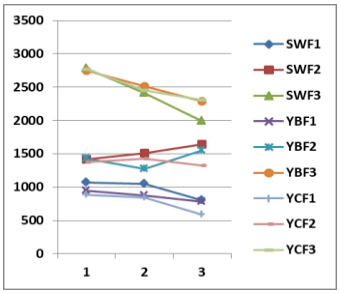

Fig. 18. Formant diagrams of Yuepu females B, C

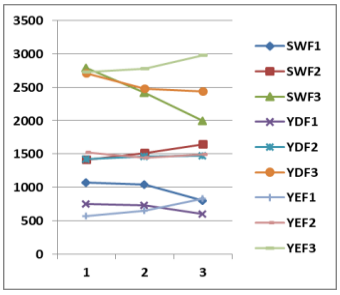

Fig. 19. Formant diagrams of Yuepu females D, E

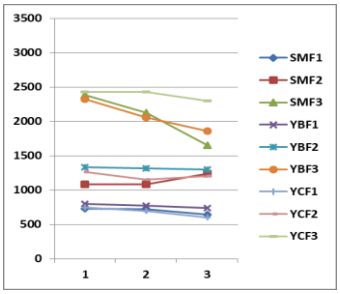

Fig. 20. Formant diagrams of Yuepu males B, C

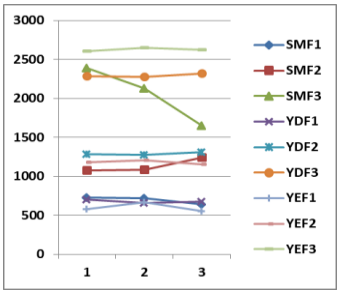

Fig. 21. Formant diagrams of Yuepu males D, E

3.2.4 Analysis on Taipu's retroflex vowels

According to Table 1, Appendix, Figure 22-23, there is only females level D with Taipu. The retroflex vowels position and tongue sliding direction of the retroflex vowels are quite different from those of the S&P. It shows one discrepancy compared to S&P. The discrepancy, K, K1, and K2 of F2 and F3 are significantly different from that of S&P.

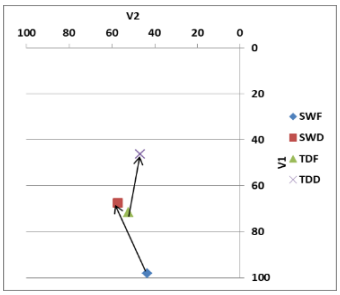

Fig. 22. Taipu female acoustic vowel diagram

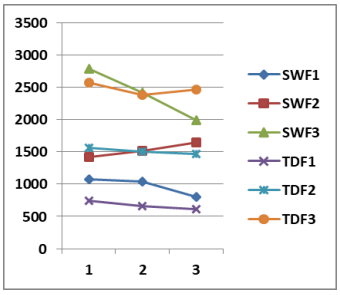

Fig. 23. Taipu female formant diagram

3.2.5 Analysis on Bapu's retroflex vowels

According to Table 1, Appendix and Figure 24-26, the retroflex vowels position and tongue sliding direction of level C, D and E of Bapu is not consistent with S&P, and the tongue sliding movement of level E is also very slight. According to the position of retroflex vowel ANOVA, there are four discrepancies between level C and S&P. Level D and E are

significantly different from S&P. According to F3 slopes ANOVA, there are two significant discrepancies between level C and S&P. Level D and E are significantly different from S&P. According to F2 slopes ANOVA , there are two places where Level C and S&P are significantly different, and there are significant discrepancies between Level D and E and S&P. The discrepancy among the three levels and the S&P is significant. In conclusion, the descent slope of F3 can be used to distinguish various levels. The difference is the smallest such as Level E, and the difference of Level C is close to that of Level D.

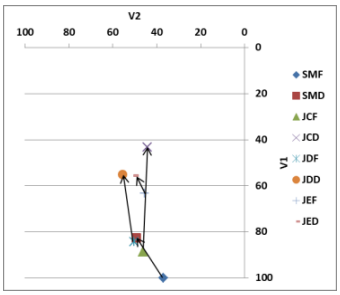

Fig. 24. Males with Bapu acoustic vowels map

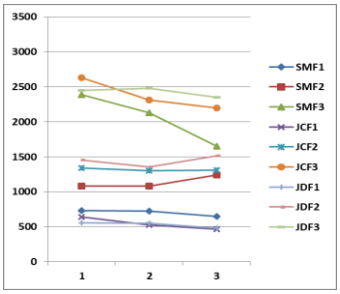

Fig. 25. Formant diagrams of Bapu Male C, D

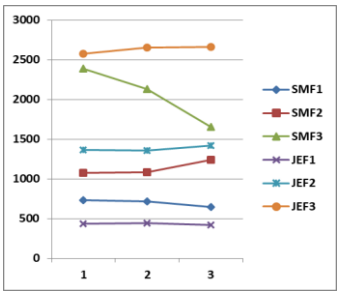

Fig. 26. Formant diagrams of Bapu Male E

3.2.6 Mengpu's acoustic vowel diagram and formant pattern analysis of retroflex vowels

According to Table 1, Appendix and Figure 27-29, The retroflex vowel position and tongue sliding direction of Mengpu level B are closer to those of S&P, followed by Level D and E. According to F3 slope ANOVA, the discrepancy between Level B and S&P is not significant.

K and K2 in Level C are significantly different from S&P. The Level E and S&P are significantly different. According to F2 slopes ANOVA, only K2 of Level E differs significantly from the S&P, and the other levels are not significant. The F3 slope can basically distinguish the levels of Mengpu's retroflex vowels. And their differences are obvious. The lower the level, the greater the difference.

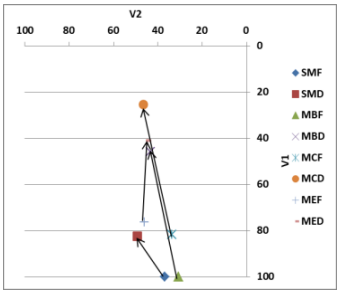

Fig. 27. Mengpu male acoustic vowel map

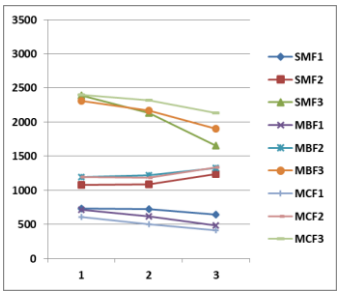

Fig. 28. Formant diagrams of Mengpu male B, C

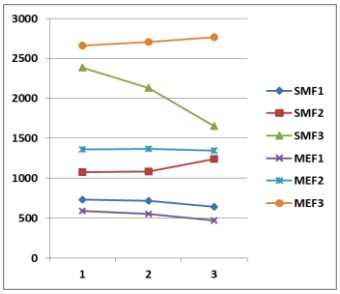

Fig. 29. Formant diagrams of Mengpu male E

4 Conclusion

Experimental results show that: (1) From the perspective F3 of the slope, there is no significant discrepancy between B and S&P except for Yuepu. F3 of the slopes in B are negative values. In C, there are one to two items which are different from the S&P, but they are all negative values. D is significantly different from S&P, and its K1 or K2 is positive, but

K is negative. E is significant different from S&P, its K1 and K2 are all positive, or one of them is positive, but K is positive. Although most of the differences are significantly different from those of the S&P. However, it can also roughly distinguish between four levels according to their differences. It can be seen that when the levels are closer to the S&P, the greater the decline in F3, the smaller the difference, and the closer F3 and F2 are. Therefore, the value of the F3 slope and the value of the difference are the key acoustic features for learners to acquire retroflex vowels.

Appendix

		V1			V2			F2			F3			Difference
Gender	level	Starting point	Midpoint	endpoint	Starting point	Midpoint	endpoint	K	K1	K2	K	K1	K2	
female	S	98.08	95.2	67.52	43.59	49.69	57.32	0.81	0.65	0.96	-2.89	-2.62	-3.15	349.73
	LB	89.05	81.16	55.89	47.08	46.85	51.97	0.28	-0.03	0.59	-2.84	-3.03	-2.65	614.54
	LC	91.21	82.45	59.59	39.44	35.36	52.25	0.82	-0.42	2.06	-0.97	-0.93	-1.01	962.45
	LD	78.37	65.63	60.34	38.97	35.14	34.95	-0.21	-0.41	-0.01	-0.14	-0.66	0.38	1106.7
	LE	74.04	72.55	57.99	35.78	34.77	30.98	-0.28	-0.16	-0.41	0.38	0.8	-0.04	1929.9
	GB	89.67	83.12	36.63	28.61	23.26	52.45	1.28	-0.56	3.12	-2.41	-2.21	-2.62	307.8
	GC	99.3	92.78	60.48	46.63	33.7	49.53	0.15	-1.47	1.77	-1.02	-0.16	-1.89	1251.2
	GD	91.52	91.37	5	41.87	43.21	43.15	0.11	0.14	0.08	-0.85	0.33	-2.02	1044.2
	GE	37.59	41.9	51.97	35.71	32.99	35.24	-0.11	-0.33	0.12	0.26	0.71	-0.18	1761.2
	YB	100	90.19	75.95	39.69	29.39	47.49	0.5	-0.94	1.93	-1.59	-1.53	-1.65	734.53
	YC	78.37	72.18	35.26	36.59	40.17	33.27	-0.25	0.26	-0.76	-1.51	-1.9	-1.12	987.23
	YD	71.62	67.95	45.7	36.82	39.35	40.42	0.26	0.35	0.17	-1.27	-2.09	-0.44	956.8
	YE	39.51	51.42	74.34	41.98	36.18	39.7	-0.17	-0.48	0.13	0.5	0.04	0.96	1484.1
	TD	71.46	56.95	46.35	52.37	48.63	46.81	-0.24	-0.34	-0.15	-0.27	-0.94	0.39	993.7
male	S	100	98.02	82.55	36.93	37.29	49.18	0.55	0.16	0.95	-2.65	-1.8	-3.51	410.15
	LB	93.08	79.02	67.29	31.81	28.14	36.37	0.26	-0.41	0.93	-2.53	-2.13	-2.94	772.03
	LD	72.02	66.62	71.65	43.35	21.98	52.13	0.43	-1.72	2.58	-0.62	-1	-0.23	901.94
	LE	69.19	57.09	74.75	31.84	30.83	42.16	0.56	-0.12	1.25	0.04	0.55	-0.47	1666.2
	GC	87.02	75.25	62.36	38.11	40.71	40.38	0.1	0.28	-0.08	-2.31	-1.83	-2.79	631.96
	GD	62.01	62.53	62.67	42.22	39.27	39.64	-0.13	-0.31	0.04	-0.12	0.11	-0.34	966.48
	YB	90.6	87.81	82.42	42.21	40.73	39.19	-0.33	-0.29	-0.38	-3.73	-4.1	-3.35	559.9
	YC	73.5	64.95	48.02	42.86	34.52	38.14	-0.11	-0.68	0.46	-0.45	-0.06	-0.85	1098.9
	YD	84.63	77.77	79.6	40.69	40	42.44	0.14	0.02	0.26	0.1	-0.01	0.21	1005.7
	YE	56.18	75.21	50.4	25.95	28.35	23.99	-0.08	0.23	-0.38	0.06	0.29	-0.16	1472.7
	JC	88.56	58.42	43.17	46.33	43.84	44.26	-0.09	-0.25	0.07	-1.14	-1.62	-0.66	889.15
	JD	84.43	79.71	54.98	50.18	41.32	55.5	0.26	-0.51	1.04	-0.24	0.38	-0.86	833.84
	JE	63.28	65.88	55.46	45.58	44.91	50.38	0.12	-0.08	0.32	0.22	-0.41	0.03	1235.9
	MB	100	78.03	45.8	30.64	33.5	43.28	0.65	0.21	1.09	-2.11	-1.41	-2.82	579.18
	MC	81.69	53.19	25.35	33.55	32.63	46.49	0.54	-0.03	1.11	-0.86	-0.51	-1.21	790.71
	ME	76.45	65.75	40.9	46.13	46.58	44.91	-0.08	0.03	-0.19	0.39	0.28	0.51	1419.6

Funding

1.2022 National Social Science Foundation Key Project "Systematic Layering Experimental Research on Phonological Features of the Three Major Dialects: Hakka, Cantonese and Min dialect (22AYY010).

2. The key research base of humanities and social sciences of ordinary colleges and universities in Guangdong Province (Chinese Dialect Research Center of Jinan University).

3.Guangxi Philosophy and Sociology Science Planning Research Project "Guangxi Southwest of Mandarin Phonetic Acoustics and Perception Experiment Research" (20FYY018).

4.The Basic Ability Improvement Project for Young and Middle-aged Teachers in Guangxi Universities -National Language Promotion Base of Guangxi Science & Technology Normal University (YWJD2014502).

5.Research and practice of the integrated teaching mode of "exploration-imitation-practice-evaluation" of language expression courses from the Perspective of New Liberal Arts (2022JGA374) Phased results.

Reference

[1] Zhou DianFu, & Wu ZongJi: Mandarin Pronunciation Map. *Beijing: The Commercial Press.* (1963).

[2] Peter Ladefoged, & Lan Maddison, Zhang WeiJia, & Tian FeiYang (translation): The Sounds of The World's Languages. *The Commercial Press*. 390 (2015).

[3] Chen ZhongMin: On Retroflexed Sound. *Minority Languages of China.* (03):3-20(2022).

[4] Sun GuoHua: Acoustic mode and perception of Mandarin retroflex vowel. *Applied Acoustics*. (04):25-29(1994).

[5] Bei XianMing: Study on Mandarin pronunciation acquisition in Guangzhou, Hong Kong and Macao. *China Social Sciences*. (2021)

[6] Liu XinZhong, Liang JiaYing, Zhou YiMin: Determinants of the Pronunciation of Rhotarized (r) Syllables in Mandarin Chinese: Main Vowels and Nasal Endings. *Journal of Guangdong Normal University of Technology*. 43(02): 88-94(2022).

[7] Yu Jue, Li AiJun, & Wang Xia: A contrastive study of Retroflex Vowel between Standard Putonghua and Shanghai Putonghua. *Contemporary Linguistics*, 03. 211-219+285(2008).

[8] Yu Jue, Lin Yuxin, & Jin Qianwen: Acquisition of Mandarin Retroflex Vowel in Children with Cochlear Implants. Chinese Journal of Phonetics, 02, 41-49(2020).

[9] Shi Feng, & Shi XiuJuan: Sampling and Data Analysis in Phonetic Experiment. *Linguistic Sciences.* 02. 23-33(2007).