Advances in Social Science, Education and Humanities Research, volume 213
4th International Conference on Humanities and Social Science Research (ICHSSR 2018)

Modeling Coarticulation of Consonants in Monguor Language

Zhang Hanbin
Northwest Minzu University, Lanzhou, China
zhanghb06@163.com

Keywords: Coarticulation dispersion, Acoustic features, Monguor.

Abstract. The purpose of the study is to discuss which acoustic features of consonants have significant linear relation with coarticulation dispersion in Monguor language. The acoustic features include F1, F2 and F3 of vowels combined with each consonant, the length and intensity of consonants themselves. On the base of acoustic parameter database of the Monguor speech, the study comes to the conclusion that a linear relationship exists only between F2 and the coarticulation dispersion of the Monguor consonants.

1. Introduction

Coarticulation is the way of pronouncing a certain sequence of vowels and consonants, and the individual movements of places of articulation will be intermingled into one smooth whole(Zhang, 2017a). The pronunciation process of each phoneme can be decomposed into the cooperative work of several pronouncing parts. When thoughts are expressed in a coherent sentence, the phonemes of adjacent notes affect each other because of the time limit of the pronunciation and the restriction of the muscle movement mechanism.

The application of acoustic features to coarticulation can expand the scope of coarticulation study. By means of acoustic realizations of a given phonetic segment affected by coarticulation with the preceding and following phonetic context, Schatz et al. (2017) confirm stronger anticipatory than perseveratory coarticulation and stronger coarticulation for lax/short vowels than for tense/long vowels. Körkkö (2015) studies the spectral moments of coarticulation development of fricative /s/ in Finnish-speaking children. The results show that there are significant differences in the effects of vowel context on the acoustics of /s/ between adults and children. There appear salient developmental stages in children’s progress toward adult-like phonetic realizations of fricative /s/.

The combination of acoustics and physiology is another way to discuss coarticulation. Bouchard and Chang (2014) discuss the relationship between the acoustic parameters of vowels and cortical activity. They find that contextual effects on human ventral sensory-motor cortex suggest active control of coarticulation. The appearance of vowels has the influence on the appearance of the preceding consonant, and conversely, the appearance of consonants affects the upcoming vowels. Dehqan et al. (2016) think that second formant (F2) transitions can be used to infer attributes of articulatory transitions. They find that, compared to nonstuttering speakers, the stuttering speakers have greater F2 during transitions, spend longer reaching vowel steady state, have steeper slopes at the beginning of transitions, and have slower speaking rates.

2. Method

2.1 Source of corpus

This study set up an acoustic parameter database of the Monguor speech. The Monguor words recorded were collected from Monguor people of Gansu and Qinhai Provinces of China. The listing 426 words are composed of only monosyllabic words because the internodes of polysyllabic words have mutual influence on coarticulation.

2.2 Speech signal collection

Four male speakers, aged from 40 to 60, read all the listing words. The speakers come from the local primary or junior schools. Their speeches belong to Huzhu dialect (in Qinghai Province). This study employs a Dell Notebook, a microphone of Behringer, and a sound card of YAMAHA Steinberg as the recording equipment. The sound is recorded in a recording studio with a sampling rate 44.1 kHz and resolution ratio 16 bits. The recording is saved with *.wav. Each word is read three times by each speaker.

2.3 Measurement of coarticulation

This study makes use of coefficient of dispersion (COD) to measure the degree of coarticulation, which was applied in previous studies (Zhang, 2017a, 2017b, 2017c, 2017d). The formula of COD is:

$$COD=(\sum|R_i-\bar{R})/(N\times M) \quad (1)$$

in which R is the ratio of $F2_{target}$ to $F2_{onset}$; F2 refers to the second formant; R_i is each ratio value; $\bar{R}$ is the mean of all ratio values; N is the number of the sample; and M is the median of all ratio values.

2.4 Research question

This study will address the following question: In Monguor language, which acoustic features of consonants have significant linear relation with coarticulation dispersion? The acoustic features of consonants include the first formant (F1), the second formant (F2), and the third formant (F3) combined with each consonant, the length and intensity of consonants themselves.

3. Results and discussion

This study collected the acoustic feature data of consonant coarticulation(Table 1). The purpose of this experiment is to test the degree of fitting between acoustic features and coarticulation dispersion.

In this study, F1, F2, F3, length and intensity are independent variables, and coarticulation dispersion is the dependent variable. A regression equation may be set up by the optimal combination of the five independent variables.

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 \quad (2)$$

In the equation, "y" is the dependent variable, namely, coarticulation dispersion; "a" is the constant or intercept; "x_1, x_2, x_3, x_4 and x_5" are F1, F2, F3, length and intensity respectively; "b_1, b_2, b_3, b_4 and b_5" are the regression coefficient or the slope of regression line of each independent variable. This study will judge the validity of the regression equation and find out whether the five independent variables can be used to predict the value of the dependent variable, what factors are found to be significant, and which factors are not significant.

Table 1. The acoustic feature data of consonant coarticulation

	Dispersion	F1(Hz)	F2(Hz)	F3(Hz)	Length(ms)	Intensity (dB)
[pʰ]	0.075	440	1336	2512	106.2	57
[p]	0.098	452	1407	2539	6.6	65.4
[tʰ]	0.144	515	1538	2643	96.4	60.4
[t]	0.128	488	1473	2697	6.2	62.6
[kʰ]	0.216	552	1675	2739	120	61.4
[k]	0.157	523	1588	2730	18.8	60.2
[G]	0.130	498	1522	2544	19	57.6
[ʦ]	0.147	517	1553	2733	79.8	55.8
[ʦʰ]	0.123	480	1462	2630	135.6	59.4
[tʂ]	0.146	516	1547	2635	30.2	64.2
[tʂʰ]	0.129	492	1488	2633	110.6	64.2
[ʨ]	0.140	508	1527	2510	56.6	61.8

[tɕʰ]	0.142	513	1534	2628	140.6	66.4
[f]	0.157	532	1568	2734	150.6	59.6
[s]	0.109	465	1452	2588	206.4	54
[ʂ]	0.148	519	1556	2660	237.8	60
[ʐ]	0.102	460	1454	2569	157	64
[ɕ]	0.201	550	1650	2733	247.6	63.8
[x]	0.114	473	1428	2607	141.8	54.8
[m]	0.093	450	1379	2501	81.4	69
[n]	0.116	476	1436	2620	80	68.2
[ŋ]	0.255	573	1739	2788	152.6	65.6
[l]	0.082	444	1354	2522	121.2	68.6

The analysis of multiple linear regression (Table 2) shows that R is 0.980 (R refers to the correlation coefficient of the fitting model of consonant coarticulation in Monguor language), R square 0.960, adjusted R square 0.948. The overall correlation is strong, and standard error of the estimate is small.

Table 2. Model summary

R	R Square	Adjusted R Square	Std. Error of the Estimate
.980	.960	.948	.010

The ANOVA of the fitting model of consonant coarticulation in Monguor language (Table 3) shows that the model deviation sum of square is 0.040, in which the sum of square about regression embodying the importance of independent variable is 0.038, and the residual sum of square reflecting the experimental error and other factors is 0.002. Accordingly, the variance of the model mainly comes from the sum of square about regression. F value of ANOVA is 80.649, and P value is smaller than 0.05. Therefore, the fitting model of consonant coarticulation in Monguor language is valid.

Table 3. ANOVA of the model

Model	Sum of Squares	df	Mean Square	F	P
Regression	.038	5	.008	80.649	.000
Residual	.002	17	9.314E-5		
Total	.040	22			

The result of regression coefficient test in Table 4 for the fitting model of consonant coarticulation in Monguor language shows that the coefficients of the constant, F1, F2, F3, length and intensity are -0.601, -2.693×10^{-5}, 3.878×10^{-4}, 4.316×10^{-5}, 2.422×10^{-5}, and 0.001respectively. The T-test shows that the constant in the model is valid ($P = 0.000 < 0.05$). Of the 5 independent variables, only F2 is valid ($P = 0.002 < 0.05$), and the P values of the remaining variables are greater than 0.05. The result shows that there is a linear relationship between the independent variable F2 and the coarticulation dispersion of the Monguor consonants in this model. Another model can be obtained by linear regression analysis with only F2 the independent variable.

Table 4. Coefficients

Model	Coefficients	T	P
Constant	-.601	-6.743	.000
F1	-2.693E-5	-.088	.931
F2	3.878E-4	3.586	.002
F3	4.316E-5	.989	.336
Length	2.422E-5	.777	.448
Intensity	.001	1.689	.110

In the second model, R becomes 0.975, and R square becomes 0.950. The ANOVA of the second model (Table 5) of consonant coarticulation in Monguor language shows that the total sum of square is 0.039, in which the sum of square about regression is 0.037, and the residual sum of square is 0.002.

Therefore, the variance of the second model mainly comes from the sum of square about regression. F value of ANOVA is 401.602, and P value is smaller than 0.05. The result proves that the second model of consonant coarticulation in Monguor language is valid.

Table 5. ANOVA of the model

Model	Sum of Squares	df	Mean Square	F	P
Regression	.037	12	.037	401.602	.000
Residual	.002	1	9.262E-5		
Total	.039	22			

The result of regression coefficient test (Table 6) for the second model of consonant coarticulation in Monguor language shows that the coefficients of the constant and F2 are -0.481 and 4.103×10^{-4} respectively. The T-test shows that both the constant and F2 in the model are valid ($P = 0.000 < 0.05$). Accordingly, the following regression equation is established:

$$y = -0.481 + 4.103 \times 10^{-4} x_2 \quad (3)$$

Table 6. Coefficients

Model	Coefficients	T	P
Constant	-.481	-15.567	.000
F2	4.103E-4	20.040	.000

This equation model further proves that the coarticulation dispersion is related to the second resonance peak, and the measurement of F2 coming from the combination a consonant and a vowel can roughly estimate the degree of coarticulation. For example, if the mean of F2 from the consonant [p^h] combined with each vowel is 1336, its coarticulation dispersion is about: $-0.481 + 4.103 \times 10^{-4} \times 1336 \approx 0.067$; if the mean of F2 from the consonant [ɕ] combined with each vowel is 1650, its coarticulation dispersion is about: $-0.481 + 4.103 \times 10^{-4} \times 1650 \approx 0.196$. The degree of coarticulation of [ɕ] is larger than that of [p^h], which means that [ɕ] is more easily affected by vowels than [p^h].

In the scatter plot of the consonant coarticulation dispersion (Fig. 1), the coarticulation dispersion of the consonants of the Monguor language is a longitudinal axis, and the normalized prediction value of the regression is a horizontal axis. The distribution of the numerical values in the map shows a linear trend. It can be seen that there is a significant linear relationship between F2 and consonant coarticulation dispersion.

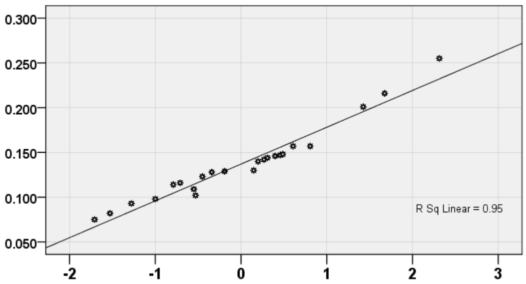

Fig. 1 The scatter plot of the consonant coarticulation dispersion

4. Summary

On the base of acoustic parameter database of the Monguor speech, the study discusses the relationship between acoustic features of consonants and coarticulation dispersion. The study shows that only F2 of the acoustic features of consonants has a linear relation with the coarticulation dispersion in Monguor language. A regression equation is established in this study: $y = -0.481 + 4.103 \times 10^{-4} x_2$. This equation model may be rough, but it can be of much help for the realizations of speech recognition and speech synthesis in Monguor language and other languages. In addition, the main significances of the study lie in language application, protection, and preservation of the precious language resource because Monguor language has become an endangered language.

Acknowledgement

This work was financially supported by State Language Commission of China (Grant No. YB135-48) and Central Government Special Funds for Northwest Minzu University (Grant No. 31920180056).

References

[1] H. Zhang, Acoustic Analysis of Nasal Vowels in Monguor Language, *Materials Science and Engineering*, vol.242, pp. (012132)1-5, 2017a.

[2] T. Schatz, Turnbull, R. Bach, and E. Dupoux, A Quantitative Measure of the Impact of Coarticulation on Phone Discriminability, *Proc. Interspeech 2017*, pp. 3033-3037, 2017.

[3] P. Körkkö, Spectral moments analysis of /s/ coarticulation development in Finnish-speaking children, *The Scottish Consortium for ICPhS*, pp. 1-4, 2015.

[4] K. E. Bouchard, and E. F. Chang, Control of spoken vowel acoustics and the influence of phonetic context in human speech sensorimotor cortex, *Journal of Neuroscience*, vol.34(38), pp. 12662-12677, 2014.

[5] A. Dehqan, F. Yadegari, M. Blomgren, and R. C. Scherer, Formant transitions in the fluent speech of Farsi-speaking people who stutter, *Journal of fluency disorders*, vol.48, pp. 1-15, 2016.

[6] T. Mahr, B. T. McMillan, J. R. Saffran, S. E. Weismer, and J. Edwards, Anticipatory coarticulation facilitates word recognition in toddlers, *Cognition*, vol.142, pp. 345-350, 2015.

[7] H. Zhang, Coarticulatory Dispersion in NCV Sequences in Monguor Language, *MATEC Web of Conferences*, vol. 125, pp. 996-1001, 2017b.

[8] H. Zhang, Acoustic Analysis of Monophthongs in Monguor Language, *Advances in Social Science, Education and Humanities Research*, vol. 146, pp. 70-74, 2017b.

[9] H. Zhang, Coarticulation of Nasal Vowels in Monguor Language, *Advances in Social Science, Education and Humanities Research*, vol. 107, pp. 81-84, 2017d.