Research on Citation Tones and Disyllabic Tone Sandhi in Maoping Town Dialect Based on Deep Learning

Zidan Wang

e-mail: 446144920@qq.com

China National Institute of Standardization, Beijing, China

Abstract—The paper focus on the application of deep learning techniques in the domain of speech processing. Chinese is a tonal language. In a tonal language, pitch information is used to discriminate the meaning of different words. A word spoken with different tones corresponds to different meanings. Chinese has many dialects.Tone related information can play an important role in the Chinese speech recognition systems.Therefore, it's necessary to make sue the tone pitches and the tone categories of a dialect in dialect speech processing database. The paper aims to use the Parrat to study citation tones and disyllabic tone sandhi in Maoping Town Dialect. Through experimental data,We proposed actual pitch value of five citation tones and disyllabic tone sandhi. There are seven patterns of disyllabic tone sandhi in Maoping Town Dialect.

Keywords—Praat; phonetic experiment; tone sandhi

1. Introduction

Deep learning can be understood as "feature learning"or "representation learning". [6]With the expectation of furthering study in this field on the basis of the previous achievements, this paper studies the tones of Maoping town dialect by means of acoustic experiments instead of the traditional speaking-listening method[8].The regional language used in this survey is Maoping Town dialect. Maoping town dialect is used in eight townships in the south of Baihe County, Ankang City, Shaanxi Province, including Sixin township, Taoyuan township, Maoping township, Songjia township, Xiying township, Cangshang township and Xiaoshuang township. [7]Baihe County is located in the southeast of Shaanxi Province. Maoping town dialect is the first language used in communication and life. About 100000 speakers use the dialect. Maoping town dialect is a language system. [4]It's significant to research the the tone system and disyllabic tone sandhi in Maoping town dialect. The pronunciation of this survey is local permanent residents.[5]

2. Phonetic Experiment of Citation Tones

Speech processing has broad application prospects, such as voice retrieval, command control, automatic customer service and machine automatic translation. [2]This paper used Praat software to get experimental datas in acoustic experiments to ascertain actual pitch value of

citation tones. [9]We use the lz-score method to normalize the fundamental frequency datas. Based on normalized data, we can obtain a chart of actual pitch value of citation tones. According to the tone examples in the "Character Table of Dialect Investigation[10]", there are 5 tone categories for citation tones in Maoping town dialect. The following text uses T1, T2, T3, T4, and T5 to represent five tone categories. We record pronunciation of three example words for each tone category. Finally, fifteen example words were selected.

2.1 The Experimental Process

2.1.1 Obtaining experimental data for each word in each category

For example，we can get the fundamental frequency value and pronunciation duration of the example word "Guai" that belongs to T3. A graphical representation of datas is shown in figure 1.
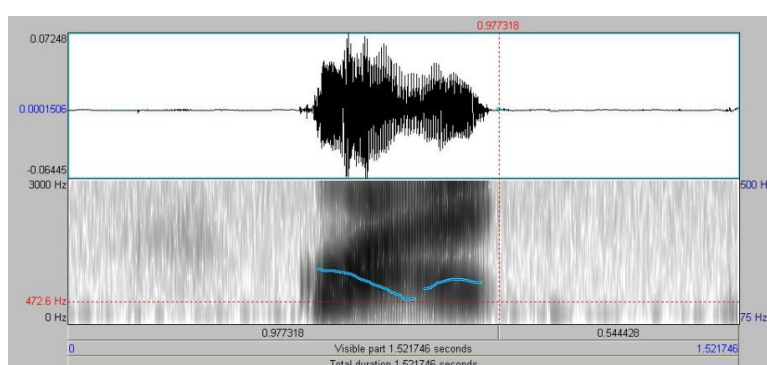

Figure 1 The Graphical Representation of The Example Word "Guai"

2.1.2 Obtaining average fundamental frequency value and average duration of example words in each tone category

In each tone category, we have three example words. Firstly, we record one word to get the fundamental frequency value and duration. Secondly we divided the duration into ten equal parts and got the fundamental frequency value in ten equal-part points. Thirdly, following the above operation, we can obtain the corresponding values for the other two words. Finally, we got the average fundamental frequency value and average duration of example words in each tone category. A graphical representation of datas is shown in table 1.

In table 1, we use P1 to represent the fundamental frequency value of the first equal-part point. We use P2 to represent the fundamental frequency value of the second equal-part point. We use P3 to represent the fundamental frequency value of the third equal-part point. We use P4 to represent the fundamental frequency value of the fourth equal-part point. We use P5 to represent the fundamental frequency value of the fifth equal-part point. We use P6 to represent the fundamental frequency value of the sixth equal-part point. We use P7 to represent the fundamental frequency value of the seventh equal-part point. We use P8 to represent the fundamental frequency value of the eighth equal-part point. We use P9 to represent the fundamental frequency value of the ninth equal-part point. We use P10 to represent the fundamental frequency value of the tenth equal-part point.

Table 1 Average Fundamental Frequency Value and Average Duration of Example Words in Each Tone Category

Tone Category	Average Fundamental Frequency Value in Ten Equal-part Points(hz)										Average Duration(ms)
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
T1	257	252	245	240	234	225	215	204	195	181	198
T2	251	236	235	235	235	235	235	234	233	235	212
T3	227	216	214	211	210	212	218	223	229	238	199
T4	227	206	183	170	136	128	174	201	210	198	242
T5	219	207	206	203	201	199	197	195	193	190	234

2.1.3 Obtaining the logarithmic value of the average fundamental frequency value

According to Table 1, we can calculate the logarithmic value of the average fundamental frequency value. A graphical representation of datas is shown in table 2.

Table 2 Logarithmic Value of Average Fundamental Frequency Value

Tone Category	Logarithmic Value of The Average Fundamental Frequency Value										Average Duration(ms)
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	
T1	2.41	2.40	2.39	2.38	2.37	2.35	2.33	2.31	2.29	2.26	198
T2	2.40	2.37	2.37	2.37	2.37	2.37	2.37	2.37	2.37	2.37	212
T3	2.36	2.33	2.33	2.32	2.32	2.33	2.34	2.35	2.36	2.38	199
T4	2.36	2.31	2.26	2.23	2.13	2.11	2.24	2.30	2.32	2.30	242
T5	2.34	2.32	2.31	2.31	2.30	2.30	2.30	2.29	2.29	2.28	234

2.1.4 obtaining mean value of logarithmic values and standard deviation

According to the data in Table 2, we can calculate that the mean value of logarithmic values is 2.32 and the standard deviation is 0.06.

2.1.5 Normalization of values

We use lz-score method to normalize the values in Table 2. In the following calculation formula, we use E to represent the logarithmic value. We use F to represent the mean value of logarithmic values. We use G to represent the standard deviation. We use M to represent the value of normalization.[3]

$$M=(E-F)/G \qquad (1)$$

A graphical representation of datas is shown in table 3.

Table 3 Values of Normalization

Tone Category	Value of Normalization									
	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
T1	1.46	1.31	1.10	0.96	0.79	0.48	0.17	-0.22	-0.56	-1.12
T2	1.30	0.84	0.81	0.81	0.81	0.82	0.81	0.78	0.75	0.81
T3	0.55	0.18	0.13	0.01	-0.04	0.05	0.24	0.42	0.62	0.89
T4	0.57	-0.15	-1.04	-1.60	-3.20	-3.65	-1.42	-0.35	-0.03	-0.45
T5	0.29	-0.13	-0.18	-0.27	-0.36	-0.43	-0.48	-0.57	-0.63	-0.76

2.1.6 Calculating absolute duration data

According to calculation，we can obtain ten absolute duration data in each tone category. A graphical representation of datas is shown in table 4.

Table 4 Absolute Duration Data in Each Tone Category

Tone Category	Ten Absolute Duration Data(ms)									
T1	20	40	59	79	99	119	139	158	178	198
T2	21	42	64	85	106	127	148	170	191	212
T3	20	40	60	80	100	119	139	159	179	199
T4	24	48	73	97	121	145	169	194	218	242
T5	23	47	70	94	117	140	164	187	211	234

2.1.7 Obtaining tone distribution map

According to the values in table 3 and table 4, we can get the tone distribution map of Maoping town dialect. A graphical representation of datas is shown in figure 2.
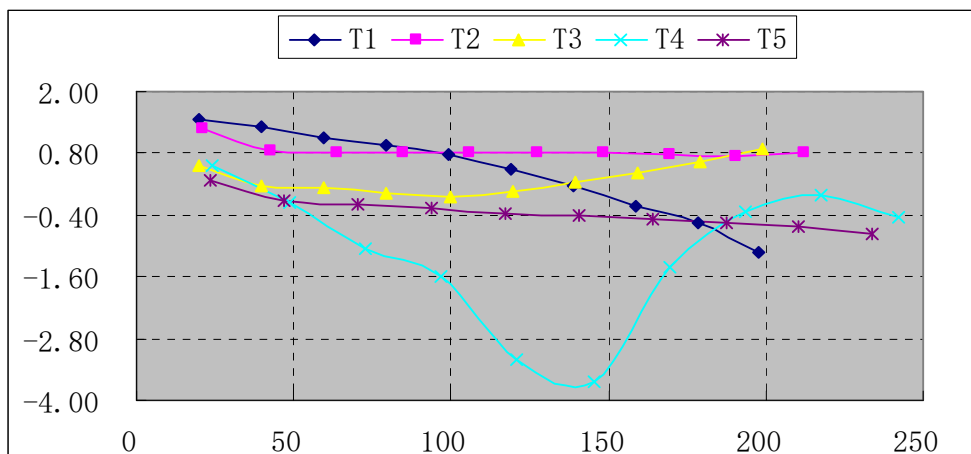
Figure 2 The Tone Distribution Map of Maoping Town Dialect

2.2 The Experimental Result

According to five level tone mark and the tone distribution map, we can ascertain actual pitch value of five citation tones in Maoping town dialect. The actual pitch value of five citation tones is T1(53) and T2 (55) and T3 (45) and T4 (413) and T5(33) .

3. Disyllabic Tone Sandhi

The actual pitch value of the first word or the last word will be change during the process of disyllabic tone sandhi. [1]Based on the comparison with fundamental frequency data of single word, we found there are seven patterns of disyllabic tone sandhi in Maoping Town Dialect.

When the actual pitch value of the first word is T2(55), A graphical representation of disyllabic tone sandhi is shown in table 5.

Table 5 The Disyllabic Tone Sandhi When The Actual Pitch Value of The First Word is T2(55)

Pitch Value of The First Word	Pitch Value of The Last Word	Disyllabic Tone Sandhi
T2(55)	T1(53)	T2(55)+T1(53)
T2(55)	T2(55)	T2(55)+T2(55)
T2(55)	T3(45)	T2(55)+T3(45)
T2(55)	T5(33)	T2(55)+T5(33)
T2(55)	T4(413)	T2(55)+**T4(41)**

When the actual pitch value of the first word is T3(45), A graphical representation of disyllabic tone sandhi is shown in table 6.

Table 6 The Disyllabic Tone Sandhi When The Actual Pitch Value of The First Word is T3(45)

Pitch Value of The First Word	Pitch Value of The Last Word	Disyllabic Tone Sandhi
T3(45)	T1(53)	T3(45)+T1(53)
T3(45)	T2(55)	T3(45)+T2(55)
T3(45)	T3(45)	T3(45)+T3(45)
T3(45)	T5(33)	T3(45)+T5(33)
T3(45)	T4(413)	T3(45)+**T4(41)**

When the actual pitch value of the first word is T4(413), A graphical representation of disyllabic tone sandhi is shown in table 7.

Table 7 The Disyllabic Tone Sandhi When The Actual Pitch Value of The First Word is T4(413)

Pitch Value of The First Word	Pitch Value of The Last Word	Disyllabic Tone Sandhi
T4(413)	T1(53)	**T4(41)**+T1(53)
T4(413)	T2(55)	**T4(41)**+T2(55)
T4(413)	T3(45)	**T4(41)**+T3(45)
T4(413)	T5(33)	**T4(41)**+T5(33)
T4(413)	T4(413)	**T4(33)**+T4(413)

Through experimental data, We can make sure actual pitch value of five citation tones and disyllabic tone sandhi. There are seven patterns of disyllabic tone sandhi in Maoping town dialect. When the structure of tone pitch in two-word group shows T2(55)+T4(413), the tone pitch of the last word changes to T4(41). When the structure of tone pitch in two-word group shows T3(45)+T4(413), the tone pitch of the last word changes to T4(41). When the structure of tone pitch in two-word group shows T4(413)+T1(53), the tone pitch of the fist word changes to T4(41). When the structure of tone pitch in two-word group shows T4(413)+T2(55), the tone pitch of the fist word changes to T4(41). When the structure of tone pitch in two-word group shows T4(413)+T3(45), the tone pitch of the fist word changes to T4(41). When the structure of tone pitch in two-word group shows T4(413)+T5(33), the tone pitch of the fist word changes to T4(41). When the structure of tone pitch in two-word group shows T4(413)+T4(413),the tone pitch of the fist word changes to T4(33).

4. Conclusions

This empirical study on the citation tones and disyllabic tone sandhi suggests an observation method which can be used to analyze and record the tone characteristics of dialects. Firstly, Dialect tone is may be acquired by pitch extraction. It's helpful to build dialect speech processing database for dialect protection. It is evidenced that the local dialect is in danger of dying out in some areas. Secondly, In addition the results of the acoustic analysis of the dialect tone provided clear support for studying the acquisition of a second language and Mandarin.

References

[1] Li, X.F. (2004) The Levels and Types of Tone sandhi in Chinese Dialects. Dialect. J. vol. 2. pp. 32-37.

[2] Lin, S.Y. (2019) Speech Recognition Technology. Digital Design. vol. 4. pp. 182-183.

[3] Zhu, X.N. (2010) Phonetics. Commercial Press. Beijing.

[4] Guo, S.Q. (2013) A Study on the Phonetics of Hakuo Dialect in Southern Shaanxi. China Social Sciences Press. Beijing.

[5] Ming, M.X. (2017) Experimental Research on Chongqing Dialect Tones. Southwest Normal University Press. Chongqing.

[6] Zhou, Z.H. (2015) Machine Learning. Tsinghua University Press. Beijing.

[7] Wang, J.C. (2019) Integration of Shaanxi Dialects. Commercial Press. Beijing.

[8] Guo, J.F. (1993) Explanation and Exploration of Chinese Tone and Intonation.Beijing Language and Culture Institute Press. Beijing.

[9] Wu, Z.j., Lin, M.C. (1989) An Outline of Experimental phonetics. Higher Education Press. Beijing.

[10] CASS. (2004) Character Table of Dialect Investigation. Commercial Press. Beijing.