ISSN 1799-2591 (print); ISSN 2053-0692 (online)
Theory and Practice in Language Studies, Vol. 16, No. 8, pp. 2592-2597, August 2026
DOI: https://doi.org/10.17507/tpls.1608.05

An Acoustic Study of the Stop Consonants in Wu Dialects: Evidence From Tongdong Dialect

Xiaowei Gu
School of Foreign Languages, Tianjin University of Science and Technology, Tianjin, China

Abstract—**This paper presents key findings from an acoustic study on the stop consonants in Tongdong dialect, a language variety in Wu dialects. The research was conducted through voice recording and analysis of data collected from Tongdong native speakers. Recordings of 3240 overall tokens were collected and analyzed using Praat phonetic software. The mean VOT values, vowel lengths, vowel pitches and vowel intensities obtained from Chinese words with both prevocalic stops and intervocalic stops were compared and statistically tested using One-way ANOVA. The research findings show that in Tongdong dialect first, the prevocalic voiced stop has a positive VOT; the voicing begins just after the release of the stop followed by a strong voiced aspiration; the adjacent vowel is significantly longer in length and lower in pitch, which makes the stop sound voiced; and the stop is actually a voiceless stop followed by a voiced aspiration. Second, the intervocalic stop has a negative VOT; the voicing begins before the release of the stop, making a voicing lead in the spectrogram and followed by a less strong voiced aspiration; the adjacent vowel is significantly longer in length and lower in pitch; and the stop is a voiced stop followed by a voiced aspiration. The research presents proof that as a border region in Wu dialects, Tongdong dialect is relatively faster in its voiced stops devoicing. Meanwhile it still keeps the three grades of stops: voiceless unaspirated; voiceless aspirated and voiced aspirated, which is a distinctive feature of Wu dialects.**

Index Terms—**Tongdong dialect, stop consonant, voice**

I. Introduction

Wu dialects, which still keep the ancient voiced initial consonants now, have three grades of stops: voiceless unaspirated, voiceless aspirated and voiced. The three-grade-classification of stops is the distinctive feature of Wu dialects from the other six dialect regions in China, which are Mandarin dialects, Xiang dialects, Gan dialects, Kejia dialects, Yue dialects, and Min dialects (Shi, 1983, pp. 49-83). Nevertheless, the ancient voiced initial consonants have been turning voiceless as time goes by, the change of which is gradual and imbalanced. Till now, the devoicing progress of the so-called voiced initial consonants are quite different from place to place (Zhong & Chen, 2012, pp. 126-129).

Nantong, a city on the north side of Changjiang estuary, locates on the isogloss dividing Jiang Huai Mandarin and Northern Wu dialects. As a result, Nantong has comparatively complex dialects, among which the communities close to Haimen district, like Erjia, Yuxi, Yuanzao, etc. are connected with Qidong district and Haimen district, where Tongdong dialect is used as the language variety. Tongdong dialect, also called Jiangbei dialect by local citizens, is an extension of Wu dialects, and it is also classified into Wu dialects (Yan, 1984, pp. 1-14; Chen et al., 2010, pp. 101-105). As a small dialect region, Tongdong dialect has been existing between Wu dialects and Jianghuai Mandarin for long, and the acoustic features of the stop consonants in Tongdong dialect have experienced changes.

Thus, the aims of the given research are as follows:

1. To identify the VOT values of the stop consonants in Tongdong dialect;
2. To provide the distinctive features, or the devoicing progress of the stop consonants in Tongdong dialect.

II. Literature Review

The three grades of stops in Wu dialects have drawn attention from quite a few Chinese linguists and researchers. In 1925, Liu Fu used a wave recorder in his acoustic experiment to prove that the voiced initial consonants in Wu dialects were not really voiced (Chao, 2002). Chao (2011) also disclosed that the voiced consonants in Wu dialects were actually voiceless consonants followed by a voiced aspiration. Chao (2002) transcribed the voiced initial consonants in Wu dialects by using a voiceless phonetic symbol followed by a [ɦ] to represent a voiced aspiration, such as [$p^{ɦ}$] [$t^{ɦ}$] [$k^{ɦ}$], etc. From then on, the saying of "voiceless initial consonants followed by a voiced aspiration" in Wu dialects has been a popular view in Chinese linguistic circle. In his later research, Chao (1970, pp. 45-56) made a clear statement:

> For simplicity in notation and terminology, we are calling these voiced stops: b, d, g. Actually, they are purely voiced only in unstressed intervocalic positions, whereas when stressed they are voiceless stops followed by a voiced aspiration.

Later researchers made similar discoveries when recording Wu accents. Qian (1992) pointed out that in most part of Wu dialects, the voiced initial consonants were not really voiced, but voiceless consonants followed by a voiced aspiration. The voiced aspiration was an auditory impression caused by the lower pitch at the beginning of these

voiceless consonants, owing to which the bilabial stop and the alveolar stop can be transcribed as [pʱ tʱ], but usually as [b d]. One of the major contrasts between the so-called voiced stops [b d] and the unaspirated voiceless stops [p t] is the closure duration: the so-called voiced stops [b d] have short closure durations while the voiceless stops [p t] have long closure durations. The voiced initial consonants in Wu dialects can be classified as voiceless lax stops. The annotations in *Syllabary of Dialect Pronunciations of Chinese Characters* say in Suzhou dialect, the voiced initial consonants are not voiced at the beginning, but voiceless consonants followed by strong voiced aspirations. *The Overview of Jiangsu Dialect and Shanghai Dialect* also says the voiced initial consonants in Suzhou, Wuxi, Changzhou, Haimen, and Shanghai are not very voiced at the beginning of articulation.

Nevertheless, not all the ancient voiced initial consonants in Wu dialects region are voiceless consonants followed by a voiced aspiration. In some parts, there exist "voiced consonants followed by a voiced aspiration", or "very authentic voiced consonants", so different dialect regions should be dealt with differently while doing researches. In Wenzhou dialect, for example, the voiced initial consonants are produced with a not very strong voiced aspiration, and thus they are voiced consonants followed by a voiced aspiration (Zhong & Chen, 2012). Cao (2022) pointed out that in the cities like Dongyang, Suichang, and Wenzhou, etc., the ancient voiced initial consonants are still voiced now: Wenzhou dialect keeps voiced consonants followed by a not very strong voiced aspiration, and Changshu dialect even keeps very authentic voiced initial consonants.

The initial devoicing is a popular trend in the evolution of Chinese language, as well as the direction of language change in Wu dialects. Nevertheless, language never changes in balance, owing to which we can feel the real existence of voiced consonants and make researches on the different types of devoicing. In some areas of Wu dialects, the ancient voiced initial consonants have totally devoiced, such as Qingyan district in Lishui city, Zhejiang province, the town of Hengdu in Shitai district, Chizhou city, Anhui province, and the town of Lingyang in Qingyang district, Chizhou city, Anhui province. Meanwhile, in some other areas of Wu dialects, the voiced initial consonants have been completely kept till now, like the Wu dialect in Wenzhou, Zhejiang province. Generally speaking, in a dialect region the step of devoicing of the ancient voiced initial consonants is faster in border areas than that in the center areas (Zhong & Chen 2012).

This study aims at conducting an instrumental analysis of the initial stop consonants in Tongdong dialect.

III. Methodology

A. Participants of the Study

The participants of the study are four local citizens in Yuxi district, all of whom speak Tongdong dialect and Chinese mandarin, and their ages are between 48 and 76.

B. Experimental Words

The stop consonants in Wu dialects are [b p pʰ d t tʰ ɡ k kʰ]. The experimental words used in this study are composed of three parts, which are one-character Chinese words with pre-vocalic stops, two-character Chinese words with pre-vocalic stops, and two-character Chinese words with intervocalic stops. Part 1 is 270 one-character Chinese words, composed of 9 sets, with 30 words for each set. The 30 words in each set are composed by one of the nine stop consonants above and 30 rhymes in Chinese Mandarin, like 扒 [bɑ], 牌 [baɪ], 旁 [bɑŋ], etc. Part 2 is 270 two-character Chinese words, with the 9 stop consonants at prevocalic positions, like 扒手 [bɑseɪ], 牌匾 [baɪpɪn], 旁边 [bɑŋpɪn], etc. Part 3 is 270 two-character Chinese words, with the 9 stop consonants at intervocalic positions, like 铁耙 [tɪbɑ], 品牌 [pʰɪnbaɪ], 身旁 [sənbɑŋ], etc.

C. Instruments

The voice recording and analyzing products used in the acoustic recording and processing are as follows.

The device used for the acoustic recording of the participants' voices was Sony Linear PCM Recorder PCM-D100. 44.1kHz/16bit was set on the recorder, and the output file was in "wav" format, which is compatible with Praat software (Gu, 2023).

Processing the recorded voices and measuring the VOT values, vowel lengths, vowel pitches, and vowel intensities were conducted using Praat software version 6.4.61 (Boersma & Weenink, 2025).

The numerical data obtained were further processed with IBM SPSS Version 22.0 (Gu, 2023).

IV. Results and Discussion

A. Prevocalic Stops in One-Character Chinese Words

In this section, the VOT values of the 9 stop consonants and the vowel length, pitch and intensity of the rhymes following will be analyzed as below.

(a). VOT of Stops

The data of the participants' VOT values are given below in Table 1.

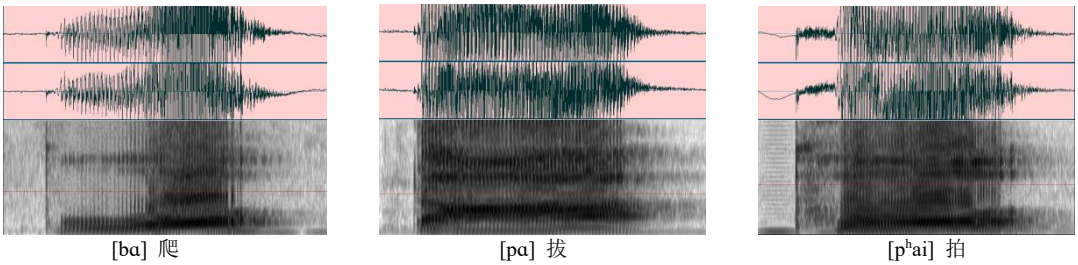
TABLE 1
VOT VALUES OF WORD-INITIAL STOPS IN ONE-CHARACTER CHINESE WORDS (MS)

Stop	b d g	p t k	p^h t^h k^h
Mean	22.66	16.20	113.34
Standard Deviation	9.33	7.64	21.62
Sample Number	360	360	360

The average of VOT values of the voiced stops [b d g] is 22.66 ms, that of the unaspirated voiceless stops [p t k] is 16.20 ms, and that of the aspirated voiceless stops [p^h t^h k^h] is 113.34 ms. This finding shows that in one-character words in Tongdong dialect, the three grades of prevocalic stops are all voiceless stops, with a positive VOT. Classified in VOT values, the stops can only be divided into two types: aspirated voiceless stops [p^h t^h k^h] and unaspirated voiceless stop [b d g p t k]. The aspirated voiceless stops [p^h t^h k^h] have long voicing lags (VOT exceeding around 35ms) while the unaspirated voiceless stops [p t k] and the so-called voiced stops [b d g] have short voicing lags (VOT less than 30 ms) or the plosive is released simultaneously with the start of vocal fold vibration (Borden et al., 1994, pp. 23-33; Gu, 2023, pp. 2910-2917; Kent & Read, 2002, p. 141).

It can be inferred that the voiced stops [b d g] in Tongdong area are actually voiceless stops. Then how can the so-called voiced stops [b d g] and the unaspirated voiceless stops [p t k] distinguish from each other, for the two grades of stops sounds very different?

Figure 1 below is the waveform graphs and the spectrograms for the three prevocalic stops randomly chosen in Tongdong dialect, [bɑ]爬, [pɑ]拔, and [p^haɪ]拍.



[bɑ] 爬 [pɑ] 拔 [p^hai] 拍

Figure 1. The Waveform Graph and Spectrogram of [bɑ] 爬, [pɑ] 拔, and [p^haɪ] 拍

Figure 1 shows that the VOT of the aspirated voiceless stop [p^h] in 拍 is much longer than that of [b] in 爬 and [p] in 拔, thus can be easily distinguished. For the stops [b] and [p], both of them are unaspirated voiceless stops, with a short VOT. The vowel [ɑ] in [pɑ] is a normal voiced sound, while the vowel [ɑ] in [bɑ] shows less intensity, weak formants, and random patterns, all of which are features of a vowel produced with aspiration. This feature is especially obvious from the beginning to the middle of the vowel, which is a clear reflection that when producing [bɑ], there is an aspirated voiced sound passing the glottis, especially in the first half of the vowel. Listeners are able to percept this voiced aspiration while listening. This distinction in the waveform graph and spectrogram verifies that the prevocalic voiced stop in Tongdong dialect is actually a voiceless stop followed by a voiced aspiration.

(b). Length, Pitch and Intensity of the Adjacent Vowel

Though the prevocalic stop is the objective of this study, it is hard to distinguish the different segments within a syllable. In many cases, a consonant can only be said to be a particular way of beginning or ending of a vowel. Thus there is virtually no difference in the sounds during the actual closures of [b d g], and absolutely none at all during the closures of [p t k], for at these moments there is only silence. Each of the stop sounds conveys its quality by its effect on the adjacent vowel (Ladefoged, 1982, p. 182).

Therefore, this study analyzes the length, pitch and intensity of the adjacent vowel to disclose the difference in the different grades of stops. Table 2 below shows the lengths of the adjacent vowels.

TABLE 2
VOWEL LENGTHS IN ONE-CHARACTER CHINESE WORDS (MS)

Stop	b d g	p t k	p^h t^h k^h
Mean	546.42	458.73	462.08
Standard Deviation	53.69	68.40	75.34
Sample Number	360	360	360

As is shown in Table 2, the average lengths of the adjacent vowels after the three grades of stops are 546.42 ms in [b d g], 458.73 ms in [p t k] and 462.08 ms in [p^h t^h k^h]. The average vowel length after [b d g] is longer than the other two groups. A one-way ANOVA reveals that the average vowel length after [b d g] is 87.689* (sig.=0,000) ms significantly higher than that after [p t k], and 84.344* (sig.=0.000) ms significantly higher than that after [p^h t^h k^h], while the difference between the average vowel length after [p t k] and that after [p^h t^h k^h] is only 3.344 (sig.=0.939) ms. It can be concluded that in one-character Chinese words, the vowel length after a voiced stop onset is significantly higher than that after a voiceless stop onset, which is in accordance with our auditory impression. In Tongdong dialect, a

one-character Chinese word with a voiced stop onset sounds longer than that with a voiceless stop onset. This feature is in accordance with the vowel length feature in Suzhou dialect (Shi, 1983, pp. 49-83).

Former researches show that the manner of articulation of a consonant is closely related to the pitch of the adjacent vowel. The adjacent vowel after a voiced consonant has a relatively lower pitch than that after a voiceless consonant (Lehiste, 1970). In a tonal language like Chinese, the vowel pitch is affected by the tone type. In order to reduce the influence from the tone type, the vowel pitch on its locus is recorded for each sample, as Table 3 below shows.

TABLE 3
VOWEL PITCHES IN ONE-CHARACTER CHINESE WORDS (HZ)

Stop	b d g	p t k	p^h t^h k^h
Mean	100.48	158.36	144.22
Standard Deviation	13.44	34.20	29.12
Sample Number	360	360	360

As is shown in Table 3, the average pitches of the vowel locus adjacent to the three grades of stops are 100.48 Hz in [b d g], 158.36 Hz in [p t k], and 144.22 Hz in [p^h t^h k^h]. The average vowel pitch after [b d g] is shorter than the other two groups. A one-way ANOVA reveals that the average vowel pitch after [b d g] is 57.878 * (sig.=0.000) Hz significantly lower than that after [p t k], and 43.744 * (sig.=0.000) Hz significantly lower than that after [p^h t^h k^h]. Meanwhile, the average vowel pitch after [p t k] is significantly 14.133 * (sig.=0.002) Hz higher than that after [p^h t^h k^h]. In Tongdong dialect, the vowels after voiced stops [b d g] are relatively lower in pitch, which is in accordance with vowel pitch rules in English referred to above.

Just like the recording of the vowel pitch, the vowel intensity at its locus is recorded for each sample, as Table 4 shows below.

TABLE 4
VOWEL INTENSITIES IN ONE-CHARACTER CHINESE WORDS (DB)

Stop	b d g	p t k	p^h t^h k^h
Mean	87.11	88.33	87.13
Standard Deviation	2.67	2.13	1.78
Sample Number	360	360	360

As is shown in Table 4, the average intensities of vowel locus adjacent to the three grades of stops are 87.11 dB in [b d g], 88.33 dB in [p t k] and 87.13 dB in [p^h t^h k^h]. Different from the results in the vowel lengths and the vowel pitches, the vowel intensities show very small differences among the three groups. A one-way ANOVA reveals no significant difference among the three averages. The voicing of a stop shows no significant influence on its adjacent vowel's intensity.

The findings in vowel length and pitch are that the vowel length after a voiced stop is significantly longer and the pitch is significantly lower, which are conformant with the findings in Suzhou dialect, an important member in Wu dialects, reported by Shi (1983, pp. 49-83).

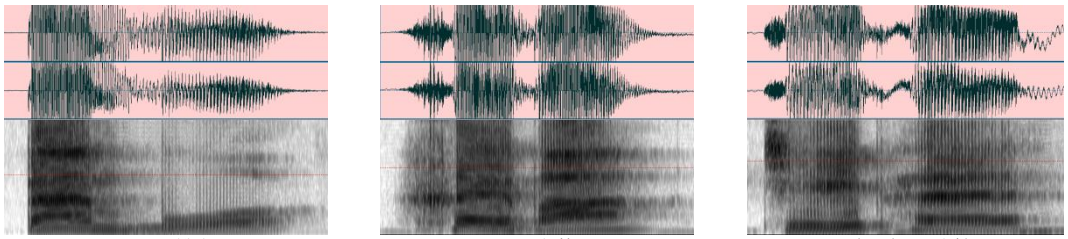
B. Prevocalic Stops in Two-Character Chinese Words

Prevocalic stops in two-character Chinese words have the same position in a syllable with the prevocalic stops in one-character Chinese words, and thus their acoustic features are almost the same, with only a slight difference. The statistic data are not listed here in detail. The VOT values of the voiced stops [b d g] are positive VOT; the average vowel length after [b d g] is significantly longer than that after [p t k] and [p^h t^h k^h]; the average vowel pitch after [b d g] is significantly lower than that after [p t k] and [p^h t^h k^h]. Nevertheless, the average vowel intensity after [b d g] is significantly lower than that after [p t k] and that after [p^h t^h k^h], which is different from the result in one-character Chinese words. Since in addition to consonant voicing, vowel intensity may be easily affected by factors like place of articulation, manner of articulation, adjacent consonants, strength of speaking, etc., the acoustic features of intervocalic stops should be observed next.

C. Intervocalic Stops in Two-Character Chinese Words

In Wu dialects, the acoustic features of the intervocalic stops are different from those of the prevocalic stops. In two-character Chinese words, the stops [b d g] in the second character change from voiceless stops followed by voiced aspirations into unaspirated voiced stops, which means that the voicing begins before the release of the stop. The voiceless stops [p t k p^h t^h k^h] keep voiceless in the intervocalic position (Chao, 2002).

(a). VOT of Stops



[pɑnbɑ] 攀爬 [haɪpɑ] 海拔 [tɕʰjupʰaɪ] 球拍

Figure 2. The Waveform Graphs and Spectrograms of [pɑnbɑ] 攀爬, [haɪpɑ] 海拔, and [tɕʰjupʰaɪ] 球拍

As is shown in Figure 2, in the spectrogram the voiced stop [b] in an intervocalic position in [pɑnbɑ] 攀爬 shows a voicing bar during the closure before the release, which is connected with the voicing bar of the vowel in the former character [pɑn]攀 and that in the later character [bɑ]爬. This voicing bar before the release of [b] is the voicing lead (Lisker & Abramson, 1964, p. 389). Meanwhile, the voiceless stop [p] in [haɪpɑ] 海拔 and [pʰ] in [tɕʰjupʰaɪ] 球拍 don't show this feature. Instead, both of them show a voicing lag after the release. It can be concluded that in Tongdong dialect, [b d g] are only voiced in an intervocalic position, while in a prevocalic position, they are voiceless stops followed with voiced aspirations.

Another distinctive feature of [bɑ] in Figure 2 is that the vowel [ɑ] shows a clear trace of aspiration: in the waveform graphs, the sound wave of [ɑ] in [bɑ] has a lower amplitude than that of [ɑ] in [pɑ] and [aɪ] in [pʰaɪ]; in the spectrogram the first half part of [ɑ] in [bɑ] shows very weak formants with clear random patterns, which don't exist in the spectrogram of [pɑ] or that of [pʰaɪ]. This feature reflects that though keeping the feature of being voiced from the ancient voiced stops, the intervocalic voiced stop [b] produces a voiced aspiration through the glottis, forming a voiced stop followed by a voiced aspiration.

This finding aligns with the findings reported by Zhong (2012) and Cao (2022). In Wu dialects region, not all the ancient voiced consonants have devoiced into voiceless consonants followed by voiced aspirations. Instead, there still exist voiced consonants followed by voiced aspirations, as in Wenzhou dialect, and classical voiced consonants as well, as in Changshu dialect. The intervocalic voiced stops in Tongdong dialect keep the feature of the ancient voiced stops, but are voiced stops followed by voiced aspirations.

(b). Pitch and Intensity of the Adjacent Vowel

The acoustic features of the vowel after an intervocalic stop are conformant with those after a prevocalic stop discussed in Part A. The average vowel length after [b d g] is significantly longer than that after [p t k] and [pʰ tʰ kʰ]; the average vowel pitch after [b d g] is significantly lower than that after [p t k] and [pʰ tʰ kʰ]. Unlike the intensities of the vowels adjacent to prevocalic stops in two-character Chinese words, discussed in Part B, the intensities of the vowels adjacent to intervocalic stops among the three groups show no significant difference, which is the same as the prevocalic stops in one-character Chinese words.

V. Conclusion

The research findings indicate that as a border region in Wu dialects, Tongdong dialect keeps the ancient voiced stop consonants [b d g] in the intervocalic position, and has three grades of stops: voiceless unaspirated, voiceless aspirated, and voiced aspirated, which is a distinctive feature in Wu dialects.

The prevocalic voiced stop has a positive VOT; the voicing begins just after the release of the stop followed by a strong voiced aspiration; the adjacent vowel is significantly longer in length and lower in pitch, which makes the stop sound voiced.

The intervocalic voiced stop has a negative VOT; the voicing begins before the release of the stop, making a voicing lead in the spectrogram and followed by a less strong voiced aspiration; the adjacent vowel is significantly longer in length and lower in pitch.

Located in the border region of Wu dialects, Tongdong dialect is relatively faster in its voiced stops devoicing than other regions. Till now, in Tongdong dialect, there is no typical voiced stop like those in Changshu dialect, but a coexistence of voiced stops with voiced aspirations and voiceless stops with voiced aspirations.

References

[1] Boersma, P., & Weenink, D. (2025). *Praat: Doing phonetics by computers (Version 6.4.61)* [Windows]. Retrieved January 20, 2025, from http://www.praat.org

[2] Borden, G. J., Harris, K. S., & Raphael, L. J. (1994). *Speech science primer: physiology, acoustics, and perception of speech* (3rd ed.). USA: Williams & Wilkins.

[3] Cao, Zhiyun. (2022). *A phonetic research on southern Wu dialects*. Beijing: The Commercial Press.

[4] Chao, Yuenren. (2011). *A study on modern Wu dialects*. Beijing: The Commercial Press.

[5] Chao, Yuenren. (2002). *Linguistic essays by Yuenren Chao*. Beijing: The Commercial Press.

[6] Chao, Yuenren. (1970). The Changzhou dialect. *Journal of the American Oriental Society*, *90*(1), 45-56. Reprinted in 1976.

Aspects of Chinese sociolinguistics. Redwood: Stanford University Press.
[7] Chen, C., Yan, M., & Yu, J. (2010). Research on dialect distribution in Nantong. *Journal of Radio & TV University (Philosophy & Social Sciences)*, *13*(3), 101-105.
[8] Group of Research on Jiangsu and Shanghai Dialects. (1960). *An overview of Jiangsu and Shanghai dialects*. Nanjing: Jiangsu People's Publishing House.
[9] Gu, Xiaowei. (2023). An acoustic study of the stop consonants in lettered-words produced by Chinese mandarin speakers. *Theory and Practice in Language Studies*, *13*(11), 2910-2917.
[10] Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). United Kingdom: Singular Thomson Learning.
[11] Ladefoged, P. (1982). *A course in phonetics* (2nd ed.). New York: Harcourt Brace Jovanovich, Inc.
[12] Language Section, Department of Chinese Language and Literature, Beijing University. (1989). *Glossary of Chinese dialects*. Beijing: Language and Culture Press.
[13] Lehiste, I. (1970). *Suprasegmentals*. Boston: MIT Press.
[14] Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurement. *Word*, *20*(3), 384-422.
[15] Qian, Nairong. (1992). *A study on contemporary Wu dialects*. Shanghai: Shanghai Educational Publishing House.
[16] Shi, Feng. (1983). The acoustic features of the stops in Suzhou dialect. *Studies in Language and Linguistics*, *3*(1), 49-83.
[17] Yan, Yiming. (1984). The northern border of Wu dialects in Jiangsu province. *Dialect*, *6*(1), 1-2.
[18] Zhong, J., & Chen, L. (2012). A comparative study of the articulatory features of the voiced initial consonants in modern Xiang dialects and Wu dialects. *Journal of Hubei Minzu University (Philosophy and Social Sciences)*, *30*(4), 126-129.

Xiaowei Gu was born in Tianjin, China in 1979. She received her master degree in English Language and Literature from Tianjin Normal University, China in 2003.

She is currently an associate professor in the School of Foreign Languages, Tianjin University of Science and Technology, Tianjin, China. Her research interests include phonetics and language testing.