INTERSPEECH 2023
20-24 August 2023, Dublin, Ireland

# The effect of stress on Mandarin tonal perception in continuous speech for Spanish-speaking learners

*Lixia Hao*[1], *Qi Gong*[1,2], *Jinsong Zhang*[3]

[1]Dongguan University of Technology, China
[2]Jinan University, China
[3]Beijing Language and Culture University, China
haolx@dgut.edu.cn, jamesgq36@163.com,jinsong.zhang@blcu.edu.cn

## Abstract

The perception of lexical tones is a well-known challenge for L2 learners especially in continuous speech where the tonal variations are complicated. To investigate whether the stress, by affecting the acoustic manifestations of tones, has an effect on Mandarin tonal perception for L2 learners, we carried out a perceptual experiment based on the Annotated Speech Corpus of Chinese Discourse with 25 native Spanish-speaking participants. The results indicate that: the perceptual accuracy of stressed tones is significantly higher than that of the unstressed ones in general; T3 is the most difficult one to be perceived among the four Mandarin tones (T1-T4) in both stressed and unstressed syllables, and presumably the Spanish-speaking learners' perceptual order of Mandarin tones is T4-T1-T2-T3; the significant interactive effect found between tone and tonal context in continuous speech may lead to great confusion of tonal perception, especially when T2 and T3 are adjacent with one another.

**Index Terms**: tonal perception, stress, Spanish-speaking learners, continuous speech

## 1. Introduction

The perception of Mandarin lexical tones for L2 learners has been a focal topic of much discussion and exploration. By employing monosyllabic/disyllabic words, and/or nonsense bi-tonal sequences, or synthetic tonal continua, previous studies have examined Mandarin Chinese tonal perception with L2 learners of various L1 backgrounds, including English [1, 2], French [3], Korean [4, 5], Japanese [6, 7], Hungarian [8], Spanish [9], etc. The results have shown that: a) L2 learners whose L1 is a tone language have performed better in the categorical perception of Mandarin tones than those whose L1 is non-tonal, which indicates a positive effect of L1 transfer; b) regarding the tonal perceptual difficulties in disyllabic words, the average error rates of preceding tones were reported significantly higher than those of the succeeding tones; L2 learners with non-tonal L1 pay more attention to the features of tones associated with the second syllables rather than the first in disyllabic words [10]; c) among the four lexical tones in Mandarin (see Table 1), the confusion of T2/R tone and T3/L tone was commonly observed [4, 6, 7, 11], basically across all L1 backgrounds; the L2 learners were found to hold a continuous but not categorical perception pattern in discriminating between the synthetic stimuli of a R-L tone continuum; d) the acoustic cues of tonal perception rest on the tone register, pitch range and pitch contour, which are basically F0 related parameters [1, 4, 9, 11, 12].

Table 1: *Four lexical tones in Mandarin Chinese.*

| Tonal category | Phonological description |
|---|---|
| Tone 1 (T1) or H tone (H) | High level |
| Tone 2 (T2) or R tone (R) | (mid) Rising |
| Tone 3 (T3) or L tone (L) | Low dipping |
| Tone 4 (T4) or F tone (F) | (high) Falling |

The previous studies have revealed a tendency and some mechanisms of Mandarin tonal perception for L2 learners and at the same time aroused new questions.

First, given that the primary acoustic cues for L2 learners' perception of Mandarin lexical tones are F0 related, the stress that is mainly manifested by F0 may exert impact upon the tonal perception of L2 learners. Stress will make the high target of Mandarin tones higher, low target lower and the pitch range extended [13, 14]. Then, by sharing the acoustic representation with tones and bringing about tonal variations, whether and/or how stress will affect L2 learners' perception of Mandarin lexical tones required to be further investigated.

Second, if the variations of F0 contour affect L2 learners' tonal perception or even lead to confusions of specific tones, as introduced above, it is hard to ignore the possible influences of tonal context that has been found to cause considerable variations of tones by the carryover and anticipatory effects [15, 16, 17]. Besides, to examine the interaction between stress and tonal context is helpful for further understanding and explaining L2 learners' perceptual difficulties and confusions.

These unsolved issues of tonal perception of L2 learners need to be addressed and further explored.

## 2. Methods

### 2.1. Material

In earlier studies on Mandarin tonal perception, specially designated disyllabic words were often used as materials. On the one hand, it is convenient for controlling irrelevant factors and explaining the results; on the other hand, the isolated disyllabic words might cover the complexity of the issue and have only limited explanatory power in the continuous speech. The present study has therefore selected materials from a discourse corpus, i.e. the Annotated Speech Corpus of Chinese Discourse (ASCCD) [18], in which 18 passages (8762 syllables) were read at normal speech rate. To maintain relatively complete syntactic and semantic structures, we chose phrases or clauses with the designated target bi-tonal combinations consisting of one stressed syllable and a succeeding unstressed syllable as perceptual materials.


10.21437/Interspeech.2023-1662

The factors considered in the experiment were lexical tones, tonal context, stress, break, intonation, length of stimuli and the gender of speakers. Among these, only the former three factors were under direct scrutiny. The control of the other four factors was done by keeping their variations to a minimum. To control the effect of break, phrases or clauses within one prosodic phrase were selected; To control intonation, only declaratives were used; To control the length of stimuli, each stimulus involved no more than seven syllables to avoid memory and processing overload for L2 learners; To control the effect of gender, female and male speakers' materials each took up 50%.

The factors under direct scrutiny were systematically manipulated. The stimuli involve all 16 possible disyllabic tonal combinations (4×4) of the four Mandarin tones (see Table 1). In the bi-tonal combinations, the first syllable is stressed and the second one is unstressed. The material-selection step was based on the segmentation of syllables and annotation of tones, stress and break index in ASCCD (as illustrated in Figure 1) and done through a set of Perl programs written for the current study.

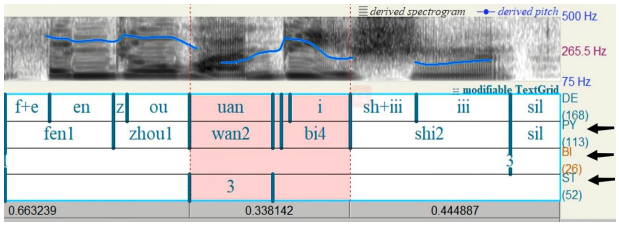


Figure 1: *Illustration of the annotation in ASCCD. The numbers in PY tier refer to Mandarin tone types (T1-T4); the number 3 in BI and ST tier refers to the prosodic phrase and the highest stressed syllable respectively. The selected two syllables exemplify a target bi-tonal combination of T2T4 (RF).*

The total number of stimuli was: 4 (1st stressed syllable with four tones) ×4 (2nd unstressed syllable with four tones) ×5 (different bi-tonal combinations) ×4 (repetition) =320.

### 2.2. Participants

The effect of stress on Mandarin tonal perception, if any, is probably easier to be tested on L2 learners whose L1 is a stress language, such as Spanish. In this study, 25 native Spanish-speaking participants were recruited from the learners of Mandarin open online course, including 20 Ecuadorians, 2 Colombians and 3 Bolivians. All of them have self-reported normal hearing and studied Chinese for about two months (the first month focused on alphabets and tones).

### 2.3. Procedure

The experiment was conducted online through synchronous sound and screen sharing via "Tencent meeting". The stimuli were arranged randomly in Praat with 5 seconds interval between each two of them. The participants' task was to listen and make choice (one out of four) for each tone of the target bi-tonal combinations, and then submit their answers online.

#### *2.3.1. Pilot test*

A pilot study with 5 Ecuadorian learners showed that all the target disyllabic words/phrases have not been studied yet by the participants of elementary Chinese level. The result perfectly meets the requirement of this study that the participants' identification of tones depends exclusively on their tonal perception rather than their knowledge of the words learned. Other feedback received in the interviews after the pilot test includes: a) the one-hour-task makes the participants feel "*sort of exhausted and hard to stay focused throughout the test*", b) the speech rate in ASCCD that is considered "normal" for native Chinese speakers was "*too fast to follow*" for elementary L2 learners. Accordingly, we reduced stimuli for each bi-tonal combinations from 5 to 3 and cut the total number of stimuli down to 192 in the formal test; the speech was lengthened (with the factor of 1.2) in Praat [19] to make it slower while maintaining its naturalness. The texts of 3 stimuli and response choices for HH (T1T1) are shown in Figure 2 as an example of 16 tonal pairs.

| | | | | |
|---|---|---|---|---|
| 1. dou jiang dàgēdà | A.dōu | B.dóu | C.dǒu | D.dòu |
| (All put Big Brother (i.e. mobile phone) ...) | A. jiāng | B.jiáng | C.jiǎng | D.jiàng |
| 2. yí gè jiandu wěiyuánhuì | A. jiān | B.jián | C. jiǎn | D.jiàn |
| (one oversight committee) | A.dū | B.dú | C.dǔ | D.dù |
| 3. zài dongfeng zhìyàochǎng | A.dōng | B.dóng | C.dǒng | D.dòng |
| (in Dongfeng Pharmaceutical Factory) | A.fēng | B.féng | C.fěng | D.fèng |

Figure 2: *Texts of 3 stimuli and response choices for HH. The syllables of target bi-tonal sequences are highlighted in gray.*

#### *2.3.2. Test*

Twenty learners participated in the formal test. Among them, data from 3 participants were excluded because of incomplete submission. The results of 17 participants were analyzed.

## 3. Results

The perceptual accuracy of stressed and unstressed Mandarin tones in various tonal contexts are shown as follows.

### 3.1. The perceptual accuracy of stressed and unstressed lexical tones in continuous speech.

As demonstrated in Figure 3, the average perceptual accuracy of stressed tones is higher than the unstressed ones. The general tendency has been observed in H, R and L tones, but not in F tone. The perceptual accuracy of the unstressed F tone is slightly higher than that of the stressed ones.

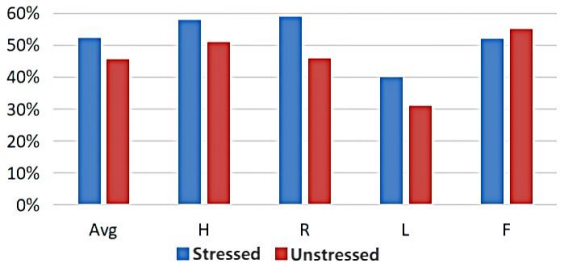


Figure 3: *Perceptual accuracy of stressed and unstressed Mandarin lexical tones.*

### 3.2. The perceptual accuracy of four Mandarin tones.

Figure 3 also shows the difference of perceptual results among the four tones. More than 50% of H, R and F tones, on average, have been perceived correctly while the perceptual accuracy for L tone is less than 40%, which is distinctly lower than other three tones in both stressed and unstressed syllables.

### 3.3. The perceptual accuracy of the Mandarin tones in various tonal contexts.

To examine the influence of various tonal contexts on the perception of Mandarin tones, we have calculated the tonal perceptual accuracy in each of the 16 bi-tonal combinations, as presented in Table 2.

Table 2: *Perceptual accuracy of Mandarin tones in various tonal contexts.*

| Target stressed tones | Succeeding tones H | R | L | F |
|---|---|---|---|---|
| H | 55% | 63% | 63% | 53% |
| R | 55% | 59% | 59% | 65% |
| L | 55% | 33% | 37% | 33% |
| F | 61% | 35% | 55% | 57% |
| **Target unstressed tones** | **Preceding tones H** | **R** | **L** | **F** |
| H | 51% | 47% | 45% | 61% |
| R | 69% | 55% | 27% | 33% |
| L | 49% | 10% | 31% | 33% |
| F | 47% | 61% | 59% | 55% |

The perceptual accuracy of the L tone, in most of the tonal contexts, is lower than 40% (highlighted in gray), which is consistent with what we have seen in Figure 3. Besides, the perceptual accuracy of the stressed F tone succeeded by the R tone, and unstressed R tone preceded by the L and F tone are also found to be lower than 40%.

### 3.4. Significance test of difference

We have performed Three-factor repeated-measures analysis of variance (RMANOVA) on perceptual accuracy with stress (stressed and unstressed), tone (H, R, L and F) and tonal contexts (adjacent with H, R, L and F) as the independent variables. The results are shown in Table 3. S, T and C refer to stress, tone and (tonal) context respectively.

Table 3: *The results of the three-factor RMANOVA.*

| Factor | df | F | Sig. | Factor | df | F | Sig. |
|---|---|---|---|---|---|---|---|
| S | 1 | 15.04 | .001 | S*T | 3 | 1.84 | .186 |
| T | 3 | 7.12 | .004 | S*C | 3 | 1.71 | .211 |
| C | 3 | 2.49 | .103 | T*C | 9 | 3.40 | .049 |
| S*T*C | 9 | 10.53 | .101 | | | | |

It is revealed that the main effects of stress and tone as well as the interactive effect of tone and tonal context are significant at 0.01 or 0.05 level (highlighted in gray). Post-hoc multiple comparison conducted for the factor of tone (df > 1) indicates that, the significant difference is only found between L tone and the other tones (H, R, F >L, $p < .001$) while the difference among H, R and F tones are not statistically significant. Simple effect analysis on the interactive effect of T*C, as depicted in Figure 4, illustrates that: a) for H tone, the effect of tonal context is slight; b) for R tone, the perceptual accuracy is the highest when it is adjacent to H tone and the lowest with L tone; c) for L tone, the perceptual accuracy is the highest when it is adjacent with H tone and the lowest with R tone; and d) for F tone adjacent to R tone, the perceptual accuracy is distinctively lower than that in other tonal contexts.

To sum up, the statistical results demonstrate that the perceptual accuracy between the stressed and unstressed tones in continuous speech is significantly different; the Spanish-speaking learners' performances in the perception of H, R and F tones are significantly better than those of L tone; the lowest perceptual accuracy of L tone and that of R tone is observed when they are adjacent with one another.

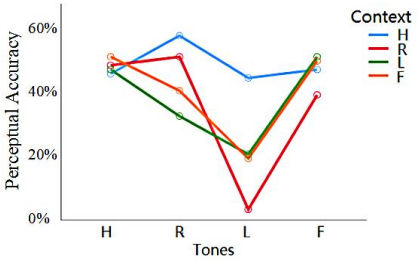


Figure 4: *Simple effect analysis on T*C.*

## 4. Discussion

### 4.1. Spanish-speaking learners' perceptual accuracy of stressed tones is significantly higher than that of the unstressed tones in continuous speech.

*4.1.1. The perceptual difference between stressed and unstressed tones is attributable to the teaching method of tones and the acoustic manifestation of stress.*

For the L2 beginners, the lexical tones are primarily taught in monosyllables in which the acoustic feature of tones are realized completely with canonical F0 forms. When it comes to the continuous speech, the realization of tones will be varied under the influence of adjacent tones and global prosodic events such as stress, break, intonation, etc. According to the model of prosodic strength [20, 21], the stressed tones that produced with big articulatory efforts are more likely to fully realize their pitch targets. In other words, the acoustic manifestation of stressed tones are approximate to that of the monosyllabic tones which are more likely to be perceived correctly. However, tones associated with syllables in weak prosodic positions, i.e. unstressed tones produced with less articulatory efforts, can be commonly seen deviated or even distorted from their target form [22, 23, 24]. The deformation of unstressed tones in continuous speech may cause perceptual difficulties for L2 learners.

*4.1.2. The perceptual result of F tone might be ascribed to the positive L1 transfer and the influence of tonal context.*

Although the perceptual accuracy of stressed F tones reached an average of all stressed tones, the result of the unstressed F tones is better (see Figure 3). The learners' performance on F tone might be affected by their L1 which is a stress language. It is found that the acoustic realization of Spanish lexical stress is similar to that of the Mandarin F tone: both lead F0 to a peak and then fall [9]. Hence, the F tone is probably easier for Spanish-speaking learners to perceive based on the theory of L1 transfer [25, 26], even when it is unstressed.

For the stressed F tones, the perceptual accuracy is distinctively lower when it is succeeded by R tone (35%, see Table 2.) Interestingly, an earlier study conducted on ASCCD has shown that in more than 90% of FR combinations, the valley of F tones (i.e. the lowest point of F0) delay to the syllables of the following R tones; while the percentage as well as the distance and the magnitude of valley delay, are much lower for the F tone succeeded by the other tones [27]. The falling target of a stressed F tone in FR is far from being realized since the falling range is remarkably decreased. The phenomenon of valley delay, as one of the acoustic

manifestations of anticipatory effect [16], may explain the perceptual result of the F tone in FR, which lowered the average perceptual accuracy of stressed F tones.

#### *4.1.3. Comparison of the current and earlier studies*

The perceptual accuracy of stressed tones, as discussed above, is significantly higher than that of the following unstressed tones in general. Unlike what we have observed, the previous studies with Mandarin disyllabic words on Spanish, Hungarian, Swedish and Indonesian learners [8-10], have shown that the average error rates of preceding tones were significantly higher than those of the following tones; and it is argued that L2 learners pay more attention to the second tone but not the first in Mandarin bi-tonal combinations.

What is worth noting is that the stress in most of the Mandarin disyllabic words is found to be associated with the second rather than the first syllable [28, 29]. We therefore assume that it is stress, rather than the position of a syllable in disyllabic words, that exerts underlying influence upon the tonal perception of L2 learners. Without examining the effect of stress, the earlier studies that seem to explained the data of isolated disyllabic words have limited explanatory power in the continuous speech and might obscure the nature of the issue. By taking the effect of stress into consideration, the results of current and earlier studies, albeit seemingly to be contradictory, can be explained in a consistent way.

Based on the findings of the present study that the unstressed tones are more difficult to be perceived accurately, and the fact that they are not rare in both spontaneous and read speech [20, 21], we suggest the L2 teachers and learners of tone languages to pay more attention to the unstressed tones and the influence of different stress patterns on tones.

### 4.2. The difference of Spanish-speaking learners' perceptual accuracy of the four Mandarin tones.

The perceptual accuracy of L tone is significantly lower than that of the other three tones in both stressed and unstressed syllables, as demonstrated in 3.2 and 3.4. Mandarin L tone was well discussed for its instability [30-34]. To be more specific, the L tone associated with a monosyllable or appearing at the end of an utterance is considered as a low-dipping tone; the L tone precedes another L tone will be turned into a R tone; the L tone preceding tones other than L tone will be turned into a low-falling or a low-level tone with a creaky voice occasionally. Due to its complexity, the fact that the L tone has long been regarded as the most difficult one among the four Mandarin tones [35] explains its lower perceptual accuracy observed in this study.

As discussed in 4.1.2, the perception of F tone is easier for Spanish-speaking learners because of the positive L1 transfer. It is in agreement with the previous studies on Mandarin disyllabic words with learners whose L1 is stress languages such as Spanish and Hungarian [8, 9]. The perceptual accuracy for H tone is comparatively high and barely affected by the tonal context; while for a R tone adjacent with L and F tones (see Table 2 and Figure 4), it is lowered. According to Krashen's "Natural Order Hypothesis", L2 learners may follow a particular order that they initially acquire some items before others [36]. Given the above, the Spanish-speaking learners' natural perceptual order of the four Mandarin lexical tones (from easy to difficult) presumably is F-H-R-L. This order is basically identical with the teaching order of Mandarin tones (H/F-R-L) proposed by Chao and Wang on the basis of their teaching and research experiences [37, 38]. The only difference is that, in this study, F tone is given more priorities than H tone for Spanish-speaking learners.

### 4.3. The interactive effect of tone and tonal context on the Spanish-speaking learners' perception of Mandarin tones.

The interactive effect of tone and tonal context are significant at 0.05 level (see Table 3); the perceptual accuracy of L tone is the lowest when it is adjacent to R tone, and vice versa (see Figure 4). For L2 learners, the confusion of Mandarin R tone and L tone is widely reported in both speech perception and production, basically across all L1 backgrounds; the L2 learners are found to hold a continuous rather than categorical perception pattern in discriminating the synthetic R-L tone continuum [4, 6, 7, 11, 12]. The confusion of R tone and L tone might lead to the higher error rates when they are adjacent with one another.

The FR pair is also worth noting, since the perceptual accuracy of both stressed F tone and unstressed R tone in the combination was observed to be low (see Table 2). The anticipatory effect, or valley delay specifically, might explain for the perception of F tone in FR, as we discussed in 4.1.2; while the carryover effect, or down-step specifically, might explain for that of R tone. Downstep refers to the phenomenon that in a pitch sequence of HLH, the F0 height of the second H under influence of L is presumably lower than that of the first H. It has been found in many languages including African tone languages [39, 40], English [41], Mandarin [42], Spanish [43], etc. In the FR that contains a pitch sequence of HLLH, the high offset of R tone is lowered with the effect of downstep, i.e the rising target of R tone is uncompleted, which makes it more difficult to be identified by L2 learners.

## 5. Conclusions

The present study examined the effect of stress on Mandarin tonal perception in continuous speech for Spanish-speaking learners. The results have shown that stress exerts significant effect on the perception of Mandarin lexical tones; L tone is the most difficult one to be perceived among four Mandarin tones, and the Spanish-speaking learners' natural perceptual order of Mandarin tones is assumed to be F-H-R-L; the significant interactive effect between tone and tonal context on tonal perception may lead to great confusion, especially when the L tone and R tone are adjacent with one another.

The findings have suggested that the effect of stress on L2 learners' perception of lexical tones needs to be taken into serious consideration in both teaching and conducting researches of tone languages. The findings of this study need to be tested with more L2 learners of diverse L1 backgrounds as well as in different tone languages. It also awaits further study upon how the interactions between stress and the other prosodic events such as break and intonation exert influences on L2 learners' tonal perception in continuous speech.

## 6. Acknowledgments

This work is supported by the Philosophy and Social Science Planning Project of Guangdong Province (GD23WZXY02-01), the International Talent Exchange Fund of Guangdong Higher Education Association (GJJL202106), and the Basic Research fund of Beijing Language and Culture University (the Fundamental Research Funds for the Central Universities) (23YJ08008). Jinsong Zhang is the corresponding author.

# 7. References


[1] Y. Wang, "Study on Mandarin tonal acquisition by Americans," *Language Teaching and Linguistic Studies*, no.3, 126-140, 1995.

[2] Y. Xu, J. Gandour and A. Francis, "Effects of language experience and stimulus complexity on the categorical perception of pitch direction," *Journal of the Acoustical Society of America*, vol. 120, no. 2, pp. 1063-1074, 2006.

[3] P. A. Hallé, Y. Chang and T. C. Best, "Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners," *Journal of Phonetics*, vol. 32, no. 3, pp. 395-421, 2004.

[4] Y. Wang and M. Li, "The perception of Tone 2 and Tone 3 in Mandarin by Korean native speakers," *Language Teaching and Linguistic Studies*, no. 1, pp. 17-25, 2011.

[5] Y, Liu, "Analysis of Mandarin tones produced by Japanese and Korean students," *Chinese Teaching in the World*, no. 1, pp. 95-100, 1998.

[6] T. Zou, J. Zhang, W. Cao, "A Comparative study of perception of Tone 2 and Tone 3 in Mandarin by native speakers and Japanese learners," *Proceedings of ISCSLP2012*, Dec, Hong Kong, 2012, pp. 431-435.

[7] P. He, "Talk about the elementary Chinese phonetic teaching the Japanese learner," *Language Teaching and Linguistic Studies*, vol. 3. 1997.

[8] Y. Wang, "Quantitative analysis of Hungarian students' Tone labeling of Chinese disyllabic words," *Chinese Teaching in the World*, no. 2, pp. 91-98, 1998.

[9] G. Wang, "On Perception of Tone in double syllables of Mandarin Chinese for foreign students of Spanish-speakers," *TCSOL Studie*s, no. 3, pp. 29-40, 2015.

[10] M. Chen, J. Wang, "On the development of Chinese disyllable tone of CSL learners," *Journal of Yunnan Normal University (Teaching and Research on CSL)*, no. 4, pp. 24-32, 2010.

[11] Y. Wang, M. Li, "The effects of tone pattern and register in perceptions of tone 2 and tone 3 in mandarin," *Acta Psychologica Sinica,* vol. 42, pp. 899-908, 2010.

[12] Y. Wang, X. Qin, "The identification and discrimination of Tone 2 and Tone 3 in Mandarin: The influence of experimental design on categorical perception of tones," *Linguistic Sciences* vol. 14, no.4, pp. 337-352, 2015.

[13] A. Li, "Acoustic representation of prosodic features in Mandarin Dialogue" *Zhongguo Yuwen (Chinese Linguistics), no.* 6, pp. 525-535-575, 2002.

[14] W. Cao, *ThePprosodic Realization of Focus and Stress in Chinese.* Peking: Beijing language and Culture University Press, 2010.

[15] X. S. Shen, "Tonal articulation in Mandarin," *Journal of Phonetics*, vol. 18, pp. 281–295, 1990.

[16] Y. Xu, "Contextual tonal variations in Mandarin," *Journal of Phonetics*, vol. 25, pp. 61–83, 1997.

[17] L. Hao, W. Zhang, Y. Xie, J. Zhang, A preliminary study on tonal coarticulation in continuous speech. in *Proc. INTERSPEECH 2018-19th Annual Conference of the International Speech Communication Association,* Hyderabad, India, September 2-6, 1849-1853, 2018.

[18] A. Li, M. Lin, X. Chen, et.al, "Speech corpus of Chinese discourse and the phonetic research," in *Proc. 6th ICSLP 2000,* Beijing, China, Oct. 16–20, pp. 42–47, 2000.

[19] P. Boersma and D.Weenink, "PRAAT, a system for doing phonetics by computer. Report 132," Institute of Phonetic Sciences of the University of Amsterdam, 1996.

[20] C. Shih, "Understanding phonology by phonetic implementation." in *INTERSPEECH 2005– Eurospeech, European Conference on Speech Communication and Technology,* Lisbon, Portugal, September, 2005, DBLP, 2469-2472.

[21] G. Kochanski, C. Shih, and H. Jing, "Quantitative measurement of prosodic strength in Mandarin," *Speech Communication*, vol. 41, no.4, pp.625–645, 2003.

[22] G. N. Clements and K. Ford, "Kikuyu tone shift and its synchronic consequences," *Linguistic Inquiry*, vol. 10, pp. 179–210, 1979.

[23] M. Y. Liberman and P. Alan, "On stress and linguistic rhythm," *Linguistic Inquiry*, vol. 8, no. 2, pp. 249–336, 1977.

[24] M. Y. Chen, "Metrical structure: evidence from Chinese poetry," *Linguistic Inquiry*, vol. 10, no. 3, pp. 371–420, 1979.

[25] R. Lado, *Linguistics Across Cultures*. Ann Arbor: University of Michigan Press, 1957.

[26] R.Ellis, *Understanding Second Language Acquisitio*n. Shanghai: Shanghai Foreign Languages Education Press: pp. 19-41, 1999.

[27] L. Hao, W. Zhang, Y. Xie, J. Zhang, Reanalyze Fundamental Frequency Peak Delay in Mandarin. in *Proc. INTERSPEECH 2017 – 18th Annual Conference of the International Speech Communication Association,* Stockholm, Sweden, August 20-24, 3186-3190, 2017.

[28] S. Xu, "Volume analysis of disyllabic words," *Language Teaching and Linguistic Studies*, no. 2, pp. 4-19, 1982.

[29] M. Lin, J. Yan, G. Sun, "A preliminary experiment on normal stress of disyllabic sequences in Beijing dialect," *Dialects,* no. 1, pp. 57-73, 1984.

[30] Y. R. Chao, "Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese," *In For Roman Jakobson* (M. Halle, editor), pp. 52-59, The Hague : Mouton, 1956.

[31] L. Wang, "Several isuues in phonological analysis of Modern Chinese," *Zhongguo Yuwen (Chinese Linguistics), no. 4, 1979.*

[32] M. Y. Chen, *"Tone Sandhi: Patterns Across Chinese Dialects,*" Cambridge: Cambridge University Press, 2000.

[33] Y. Chen, Y. Xu, "Production of weak elements in speech-evidence from F0 patterns of neutral Tone in standard Chinese," *Phonetica*, vol. 63, pp. 47-75, 2006.

[34] W. Cao, "Mandarin tone perception: A report on level F0 contours," *Zhongguo Yuwen (Chinese Linguistics)*, no. 6, pp. 536-543+576, 2010.

[35] Y. Liu, G. Liu, "On the issue of teaching Chinese phonetics to English speaking students," *Language Teaching and Linguistic Studies*, no. 3, 1977.

[36] S. D. Krashen and T. D. Terrell, "The natural approach: language acquisition in the classroom," *Modern Language Journa*l, vol. 19, no. 3, 1983.

[37] J. Chao, "Language teaching methodology based on specific languages," *Chinese Teaching in the World*, vol. 28, no. 04, pp. 554-565, 2014.

[38] A. Wang, "On teaching the features of Chinese tones," *Language Teaching and Linguistic Studies*, no. 03, pp. 70-75, 2006.

[39] J. M. Stewart, "Key lowering (downstep/downglide) in Dschang," *Journal of African Languages Linguistics* 3, pp. 113-138, 1983.

[40] Y. Laniran, "*Intonation in tone languages: the phonetic implementation of tones in Yorùbá*," Doctoral dissertation, Cornell University, 1992.

[41] J. Pierrehumbert, "*The phonology and phonetics of English intonation,*" Doctoral dissertation , MIT, Cambridge , MA. 1980.

[42] C. L. Shih, "Tone and intonation in Mandarin," *Working Papers, Cornell Phonetics Laboratory 3*, pp. 83-109, 1988.

[43] P. Prieto, C. L. Shih, H. Nibert, "Pitch downtrend in Spanish," *Journal of Phonetics*, vol.24, pp. 445-473, 1996.