Utilizing Praat to Identify Neutral Tone Features of Mandarin Learners with L1 as Longshan Dialect

Yao Peng(✉) and Xinrui Li

Faculty of Language and Linguistics, University of Malaya, Kuala Lumpur, Malaysia
{s2017961,s2016504}@siswa.um.edu.my

Abstract. The neutral tone of Mandarin is well-known for its unique acoustic features, so foreign Mandarin learners and Chinese dialect speakers presumably have difficulty in producing the neutral tone. At present, some second/foreign language acquisition studies have investigated the pronunciation of the neutral tone by foreign learners [1], but there is a dearth of scholars who touch upon the production of this prosodic feature by Chinese dialect speakers. The present study aims to use an acoustic software to identify the features of the Mandarin neutral tone by Chinese dialect speakers. Specifically, the author used Praat to compare the production by two Longshan dialect speakers with that by two native Mandarin speakers. Praat clearly demonstrated that dialect speakers could not correctly produce the neutral tone. Their vowels bearing the neutral tone featured excessively long duration, non-standard pitch variation and range, as well as absent vowel centralization. The findings showcased the effective role of Praat in identifying the distinctive prosodic features of Chinese dialect speakers. Thus, the authors recommended the use of Praat in Mandarin teaching, which is conducive to the promotion of Mandarin nationwide.

Keywords: Neutral Tone in Mandarin · Pitch · Longshan Dialect · Praat · Software in Language Learning

1 Introduction

Mandarin has been widely known as a tonal language, which has 4 tones to distinguish word meanings. The four lexical tones include high tone, rising tone, low-dipping tone, and falling tones, which are generally referred to as tone 1, tone 2, tone 3 and tone 4. A numeric notation system has been developed to describe these four contrastive tones and indicate low to high pitches from with one to five. Accordingly, the pitch values including “55”, “35”, “214”, and “51” have been widely used to respectively represent tone 1, tone 2, tone 3 and tone 4. Nevertheless, when following a contrastive tone, certain syllables in Mandarin Chinese might lose the original lexical tones and sound weak and short [2]. Some grammatical markers (such as the particle la following a verb) and diminutives (such as the second mei in meimei) are typical examples. On another note, Longshan dialect, spoken by the local residents in Longshan County of Hunan province,

C. F. Peng et al. (Eds.): EIMT 2022, AHSSEH 3, pp. 88–97, 2023.
https://doi.org/10.2991/978-94-6463-012-1_11

has a tone system distinct from Mandarin Chinese. The major difference lies in the pitch values of the four lexical tones [3]. Namely, the syllable corresponding to the Mandarin syllable characterizes a different tone contour. Notably, the pitch values of tones 1 to 4 in this dialect are respectively "44", "22", "31", and "14" [4]. Typically, all the Mandarin syllables bearing tone 4 are assigned with a rising pitch contour by this dialect. What's more, neural tone has not been observed in this dialect [5].

However, most Mandarin teachers do not undergo phonetic training and might find it daunting to notice the tone discrepancies between the Longshan dialect and Mandarin. Thus, Mandarin teachers cannot provide constructive guidance for dialect speakers, who presumable fail to naturally pronounce Mandarin neutral tone. The researchers assumed the acoustic software, Praat, could be employed by Mandarin teachers to enhance their tone perception competence and to effectively illustrate tone features to dialect speakers in classrooms.

1.1 Research Objectives

Because of the tonal distinction between the widely used Mandarin and rarely noted dialect as well as the prospective contribution of Praat for Mandarin learning and teaching, the authors attempted to showcase the procedures to identify the Mandarin neutral tone features of Chinese dialect speakers and to inform the Mandarin teachers of effective teaching approaches of Mandarin prosody. Specifically, with focus on disyllabic words, we examined if the software can showcase the neutral tone produced by the dialect speakers vary from that displayed by native Mandarin speakers. In addition, we also explore if the native speakers of Longshan dialect fail to produce native-like Mandarin neutral tone based on the acoustic data extracted from Praat.

1.2 Research Questions

In response to the research aim and objectives, this study is guided by the following question: How do the acoustic features of the neutral tone produced by dialect speakers differ from that produced by native Mandarin speakers? In this study, we focus on three variables, including duration, pitch value, as well as vowel quality?

2 Literature Review

Some studies in the field of second language learning have touched upon Mandarin prosody because of its distinctiveness, but most of them are concerned with the four lexical tones. Considering that previous studies sorely look at monosyllables, Tu et al. [6] incorporated disyllabic words in an empirical study to explore the error patterns of the tones produced by Japanese Mandarin learners. Based on the findings, the most errors lied in tone 2 and tone 3 of the initial syllable, and the third tone was also nonstandard on a frequent basis. Such result indicated negative transfer from Japanese prosody. Similarly, Fan et al. [7] conducted a production study involving twenty Korean learners and eight disyllabic words. These scholars observed the most frequent errors in tone 2 and tone 3 of the initial syllable, and no obvious variation was found among the four contrastive

tones falling on the final syllable. As shown in these two studies, foreign Mandarin learners possibly encounter huge challenge when pronouncing the four lexical tones of Mandarin. The neutral tone might pose a greater challenge for them, which highlights the urgency and necessity to closely examine this prosodic feature. Considering this point, Deng and Zhu [1] promptly investigated the production and perception of the neutral tone by forty 40 Mandarin learners from Europe. The empirical results suggested that the participants judged if a tone is lexical or neutral mainly by relying on duration and could not master the duration and pitch when producing the neutral tone. This timely piece of study can serve as a reminder for the teachers of Mandarin to place emphasis on neutral tone.

On the other hand, the production of Mandarin neutral tone by native Mandarin speakers has been extensively investigated from varied perspectives. In a quantitative study [8], the researchers unpacked how lexical frequency interacted with tonal neutralization of Mandarin. Their findings indicated that if a syllable with the neutral tone is frequent, weaker intensity, lower pitch, and shorter duration can be observed. Woefully, the study suffered from a small sample size with only six participants, and its findings run counter to another study [9]. The latter study was conducted under three stimuli conditions of manipulated pitch, manipulated duration, as well as manipulated pitch plus duration. The scholars explored the identification of the neutral tone by 18 native speakers. With the aid of Generalized Linear Mixed Effects Modeling, they found that the indispensable feature to recognize the neutral tone included F0 and duration, and no frequency effect was observed. Although the majority of scholars only focused on how the neutral tone varies from the corresponding full tone in terms of phonetic characteristics, Fan et al. [10] examined the potential effect of neutral tone on the preceding tone. The scholars adopted minimal pairs consisting of neutral-toned words and corresponding canonical words to check the variables of duration, F0, and intensity. Via statistical analysis of paired t-tests and repeated measures ANOVA, the researchers did not observe any significance difference between the initial syllables of these minimal pairs. However, they observed intensity effect in the tonal combination of T4T4 as well as the wider pitch range of the initial syllable in comparison to the other corresponding word combination.

Although the studies reviewed have afforded valuable insights, scare studies have touched upon the challenge that Chinese dialect speakers might face in producing the neutral tone of Mandarin. Among the few scholars, Liu and Gu [11] compared the production of the Mandarin neutral tone by Hong Kong Cantonese and native Mandarin speakers in isolation and sentences. According to the findings, the neutral tone produced by the Cantonese participants was too long, and they could not produce correct neutral tone under the conditions of broad or narrow focus. Drawing on this study, we aim to expand the research and explore how the speakers from other dialect regions produce the neutral tone of Mandarin in two-syllable words, which has occupied a huge proportion of Chinese words. The authors aim to examine how dialect speakers vary from native Mandarin speakers in terms of vowel quality.

3 Methods

To answer the research question, the authors conducted a quantitative study. A reading task (involving four disyllabic words) was implemented to record the production of the participants. All of the participants are female and similarly aged, without any residence experience in any country or region where any non-Chinese language is dominant. All of them cannot speak any non-Chinese language fluently. In the process of reading task, the speakers were requested to insert the four words in a carrier sentence (see Table 1). Each sentence was read three times, and the recordings were checked for the sake of high acoustic quality. The study is limited in research scope because of experimental words and participants, but we covered all of the tonal combinations in Mandarin.

3.1 Participants

The authors adopted convenience sampling because of speedy and cost-effective data collection. The participants of this study include four female speakers, aged from 20 to 29. Particularly, the first group (Group 1) consists of two speakers from Dalian city of Liaoning province (a city in which Mandarin is the medium of communication for the majority of residents). They have achieved 90 scores at minimum in the National Mandarin Proficiency Test. The second group (Group 2) consists of two dialect speakers with a high school degree. They rarely communicate in Mandarin. To ensure anonymity, the two speakers of the first group is referred to as P1 and P2. On the other hand, the participants of the second group include P3 and P4. Although the sampling method might be involved with bias, the information provided herein suggest that Group 1 and Group 2 can respectively represent the two targeted populations under investigation.

3.2 Stimuli

The study covers four disyllabic words (see details in Table 1), of which the initial syllable bears four lexical tones and the final syllable bears the neutral tone. These two-syllable words are part of the Neutral-Toned Words for National Mandarin Proficiency Test, which consists of frequently used Mandarin phrases with the neutral tone. To exclude expounding variables, the study only covers the vowel [a] with the neutral tone.

Table 1. The Stumili

Chinese sentences	Translation
T1T0: Wo qingchu dasheng de du DALA.	T1T0: I read DALA clearly and loudly.
T2T0: Wo qingchu dasheng de du LIBA (T2T0).	T2T0: I read LIBA (T2T0) clearly and loudly.
T3T0: Wo Wo qingchu dasheng de du (T3T0).	T3T0: I read LABA (T3T0) clearly and loudly.
T4T0: Wo qingchu dasheng de du (T4T0) XIABA.	T4T0: I read XIABA (T4T0) clearly and loudly.

To avoid any influence from the adjacent phonological environment and maximize the naturalness of the disyllabic syllables, all the words will be embedded into the sentence, wo qingchu dasheng de du XX (I read XX clearly and loudly).

3.3 Data Collection

The participants recorded the stimuli in a quiet room. They first went through practice with non-test words to familiarize themselves with the procedures. In the process of formal recording, the words were randomly ordered and presented to the participants via computer. If one sentence displayed, the speakers immediately clicked the recording button and recorded the relevant sentence with Audacity. Finally, the audio files were stored in wav format.

3.4 Data Analysis

Duration and pitch have been recognized as distinctive and reliable features of the neutral tone, which is oftentimes accompanied by vowel centralisation [12]. Therefore, we focused on the three variables, i.e., F0, duration, and quality. They were mainly identified by Praat [13]. In this study, we only looked at the neutral-toned vowels in the final syllable of the disyllabic words. This decision is based on the view of Ye [14] that the nucleus is the loudest section of a syllable and can determine syllable duration. Therefore, we first annotated the vowel of each recorded sentence (see Fig. 1 for details). The authors checked the annotated audio files based on the spectrogram and the sound. We also manually edited the wrong pitch values. A native Mandarin speaker was also invited to examine the consistence between the corrected pitch contours and the recordings. Subsequently, we exacted the pitch values of 11 equidistant points and duration of the syllables with neutral tone. With regard to vowel quality, we extracted the formant values from Praat and drew a vowel chart which can clearly showcase the height and backness of each vowel pronounced by the speakers. This vowel chart aims to check whether vowel centralization exists. This design is supported by Ladefoged and Disner who claim that a plot is the most effective way to show the values of first and second formant for each vowel [15].

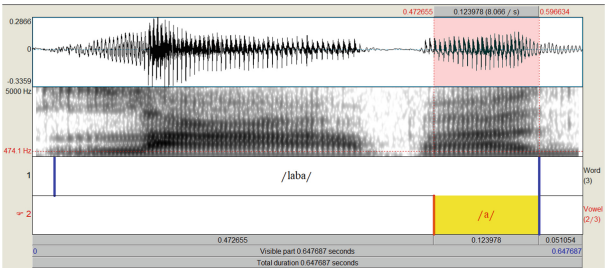

Fig. 1. Annotation example of LABA

4 Findingds and Discussions

4.1 Duration

As show in Fig. 2, the neutral tone produced by dialect speakers was longer than that produced by Mandarin speakers. Particularly, when the first syllable bears tone 1 and 4, the two groups varied a lot in duration, with Group 2's value almost doubling that of the other group. Given the fact that the neutral tone is shorter than full lexical tones [10], it can be inferred that the third and fourth participant could not sustain the neutral tone for a certain duration assigned to the Mandarin neutral tone. The findings are consistent with a few neutral tone studies, including those targeting foreign Mandarin learners [16] as well as Chinese dialect speakers [11].

4.2 Pitch Value

According to Fig. 3, the groups' pitch patterns are different. A few scholars [17] have agreed that the preceding tone can influence the pitch trajectories of the following neutral tone. According to Wang [18], the neutral tone exhibits a downward trend when the first syllable bears tones 1, 2, and 4, but remains level (or level-downward or level-upward) if the preceding tone is tone 3. According to the data, Group 1 has produced the standard pitch trend. However, the pitch contour of P3 and P4 was inconsistent with the corresponding trajectory of the native Mandarin speakers in almost every tonal combination. Notably, the pitch trend of the third participant increases all the time in T1T0 and maintains almost level when the first syllable bears tone 2. In the tonal combination of T3T0, this participant suddenly increased the pitch in the process of production and exhibited almost unchanged trend at the beginning and final section. Furthermore, she did not sustain the falling trend during the whole process of vowel production when the first syllable bears tone 4. Meanwhile, when the first syllable bears tone 1, P4 first displayed an unnoticeable decrease and then moderately increase the pitch. When the initial syllable is tone 3, she exhibited upward trend all the time. Meanwhile, Zhao [19] argues the pitch value of any lexical tone should be wider than that of the neutral tone. In terms of pitch range, P3's pitch range appeared to be quite wide when the neutral tone follows tone 1, tone 3, and tone 4. When the initial syllable bears tone 1, tone 2 and tone 3, P4 also exhibited unexpected wide pitch range.

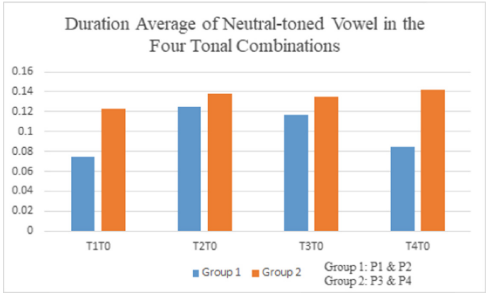

Fig. 2. The average duration of the vowels with neutral tone

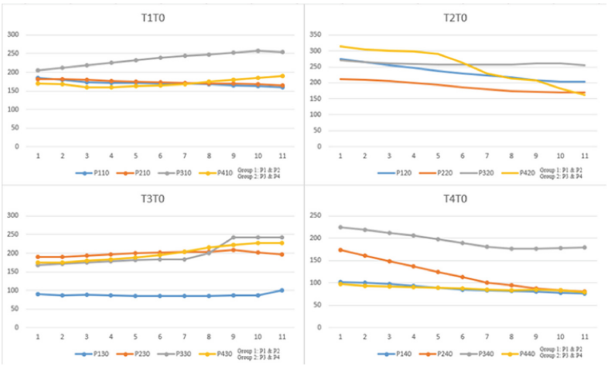

Fig. 3. Pitch trajectories of the vowels with neutral tone

Table 2. The average of formant values

	F1 (Bark)	F1 (Hz)	F2 (Bark)	F2 (Hz)
Group 1	7.74	886.40	11.48	1564.41
Group 2	8.21	950.95	11.97	1626.11

The unexpected pitch performance of Group 2 may be explained by L1 transfer [20]. In other words, the tones of the dialect detrimentally influences the standard pronunciation of Mandarin neutral tone. As noted above, the dialect lacks the neutral tone, and the two dialect speakers have not obtained sufficient Mandarin exposure. Thus, these two dialect speakers could not produce the neutral tone with the pitch contour and range. These findings are consistent with Huang's study [21], in which it was found that the mother tongue of Thai Mandarin learners prevented them from correctly pronouncing F0 contour of the neutral tone. It is worth noting that Thai is also a tonal language with pitch values different from those of Mandarin (55, 35, 214, and 51). Meanwhile, the pitch range of the second participant was unexpectedly wide in the T4T0 combination, which is possibly concerned with the assimilation by the initial tone. The preceding tone 4 has the widest pitch range among the four contrastive tones of Mandarin [20].

4.3 Vowel Quality

As shown in Table 2, the first and second formant values of Group 1 are lower than those of Group 2. As suggested by the formant frequency model, the first formants can mirror vowel quality or position. F1 determines height, and a higher F1 value means the vowel is quite low. In contrast, a higher F2 value suggests that the pronounced vowel is quite fronted. Thus, we can infer that the vowels of Group 2 are lower and more fronted than those produced by Group 1.

The findings accord with the formant plot as shown in Fig. 4. In detail, if the first syllable bears tone 1, the vowel of P3 is far fronted in comparison to other participants. When the initial syllable bears tone 2, this figure suggests that both dialect speakers and

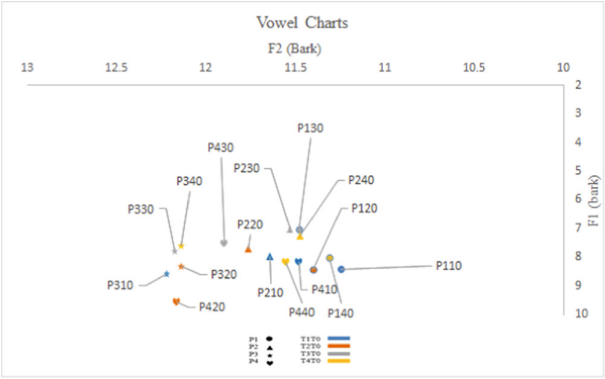

Fig. 4. The formant plot of vowels produced by speakers

the native Mandarin speakers have exhibited a mixed high-low pattern. However, the vowels of P1 and P2 are backwards to some degree. If the preceding tone of this vowel is tone 3, the participants of the first group resemble each other in vowel high-low and front-back dimensions. This figure also shows P3 and P4 markedly vary from the other group in frontness, with vowels slightly lower and markedly fronted. In case of T4T0, P4 resembles the first and second speaker in vowel openness and frontness, whereas the second formant of P3 remains quite high.

According to Lin and Yan [22], neutral-toned monophthongs in Mandarin such as [a] and [i] resembles [ə], which is called as central vowel. This phenomenon occurred in the vowel produced by the native Mandarin speakers, but the Longshan dialect speakers failed to show the case of vowel centralization. Accordingly, we can come to the conclusion that the dialect speakers could not correctly and naturally pronounce vowels with the neutral tone. Vowels are centralized mainly because of shorter pronunciation period and less energy [23]. This study does not cover intensity, but the duration data concurs with the conclusion that the dialect speakers encountered problem when they produce the neutral tone.

5 Conclusions

In general, the Praat can effectively show that Longshan dialect users are unable to produce the Mandarin neutral tone appropriately. Firstly, they produced the neutral-toned syllables for a longer duration than the Mandarin speakers. The Praat data also revealed that the performance of dialect speakers in pitch contour, pitch range and vowel centralization was not standard and natural. Meanwhile, the researchers have described the procedures for extracting neural tone characteristics using Praat (including duration, pitch, and vowel formants) and showcased the effectiveness of this computer software in identifying speakers' neutral tone features, which lends support to previous studies [24, 25]. Future studies are expected to investigate the direct effects of Praat in teaching neutral tone to Chinese dialect speakers. Nevertheless, effective measurement of the data within the sample is constrained by the limited participants and the target words.

Acknowledgments. The authors highly appreciate the speakers for recording. Sincere thanks also go to the reviewers for revision suggestions.

References

1. Deng, D., and L. Zhu. 2019. The perception and production of Mandarin neutral tone on L2 learners. *Yuyan Jiaoxue Yu Yanjiu* 2019 (5): 13–24.
2. Chao, Y.R. 1968. *A grammar of spoken Chinese.* University of California Press.
3. Cao, Z. 2009. Special phonetic phenomena in Xiangxi dialect. *Dialect* 2009 (1): 18–22.
4. Cao, F. 1985. The Longshan county annals. Longshan County Annal Office.
5. Chen, P., and J. Zhou. 2011. The characteristics of Mandarin neutral tone. *Journal of Chongqing Radio & TV University* 23 (6): 60–63.
6. Tu, Y.T., Y.W. Hsiung, M.D. Wu, and Y.T. Sung. 2014. Error patterns of Mandarin disyllabic tones by Japanese learners. In *Proceedings of the 15th Annual Conference of the International Speech Communication Association.*
7. Fan, S.S., A.J. Li, J. Gao, and A. Chen. 2016. The prosodic effect of the neutral tone to the preceding tone. In *Proceedings of the Speech Prosody 2016*.
8. Kong, H.F., and S.Y. Wu. 2019. Frequency effect and neutralization of tones in Mandarin Chinese. *Journal of Quantitative Linguistics* 26 (2): 1–21.
9. Fan, S.S., and A. Chen. 2015. The recognition of neutral tone across acoustic cues. In *Proceedings of the International Conference Oriental COCOSDA Held Jointly with 2015 Conference on Asian Spoken Language Research and Evaluation.*
10. Chen, Y., and Y. Xu. 2006. Production of weak elements in speech: Evidence from F0 patterns of neutral tone in Standard Chinese. *Phonetica* 63 (1): 47–75.
11. Liu, L., and W.T. Gu. 2018. Experimental analysis of neutral tone in Mandarin speech produced by Hong Kong Cantonese learners. *Journal of School of Chinese Language and Culture* 2018 (4): 116–123.
12. Yu, X.M. 2019. A brief review of the researches on the properties and acoustic correlations of the neutral tone in Mandarin. *The Science Education Article Collects* 2019 (454): 70–72.
13. Boersma, P., and D. Weenink. 2021. Praat: doing phonetics by computer [Computer program]. Version 6.1.49, retrieved 12 June 2021 from http://www.praat.org/
14. Ye, J. 2007. The Study on the rhythms of modern Chinese. Shanghai Century Publishing Group.
15. Ladefoged, P., and F. Disner. 2012. *Vowels and Consonants.* Wiley.
16. Li, W. 2020. An experimental study on the acquisition of neutral tone by Korean students [Unpublished master's thesis]. Hebei Normal University.
17. Lin, M.C., and J.Z. Yan. 1980. The acoustic properties of the neutral tone in Beijing Chinese. *Dialect* 1980 (2): 166–178.
18. Wang, W.J. 1996. Study on the pitch of neutral tone. *International Chinese Teaching* 1996 (3): 50–53.
19. Zhao, Y.R. 1979. A grammar of spoken English (Lv, S. X. Trans.). The Commercial Press.
20. Wang, G.P., X.B. Zhou, and A.J. Li. 2009. Pitch analysis of disyllabic neutral tone in standard Chinese spoken by foreign students. *Applied Linguistics* 2009 (4): 113–121.
21. Huang, Q. 2013. The Chinese neutral tone analysis on Thai students [Unpublished master's thesis]. Yunnan University.
22. Lin, M.C., and J.Z. Yan. 1990. Neutral tone and stress in Mandarin. *Language Teaching and Study* 1990 (3): 88–104.
23. Liu, J.L. 2002. The review on Mandarin neutral tone. *Chinese Language* 2002 (3): 43–47.

24. Gorjian, B., A. Hayati, and P. Pourkhoni. 2013. Using Praat software in teaching prosodic features to EFL learners. *Procedia-Social and Behavioral Sciences* 84: 34–40.
25. Chen, M. 2022. Computer-aided feedback on the pronunciation of Mandarin Chinese tones: Using Praat to promote multimedia foreign language learning. *Computer Assisted Language Learning*, 1–26.