

ANALYSIS OF MANDARIN EXCLAMATIONS: A COMPARISON OF THE PRONUNCIATION OF “EN” IN REGIONS*

Jiaying Li
York University

This study focuses on two phonetic differences of the pragmatic particle *en* (嗯 in Chinese character): tonal pitch and vowel length in four representative Chinese cities: Guangzhou, Shanghai, Beijing and Shenyang. To compare the use of the particle in two meanings: agreement and disagreement, the method of sociolinguistic interview was used to collect data from speakers, and the software Praat was used for the acoustic measurement. From the results, Guangzhou participants use the vowel duration contrast, but they use the falling tone in both meanings. Shanghai participants may rely on both. They used the falling tone for the agreement meaning, but they used the level tone for the disagreement interpretation. Beijing speakers used falling tones in both meanings, so they only rely on the vowel duration contrast. For Shenyang speakers, they may rely on both the duration contrast and the tone contour difference. They used the falling tone for the agreement interpretation, but they produced the level tone with the particle of the disagreement meaning. Generally, the duration difference in terms of the meaning contrast occurs in all four cities, but participants from only two cities appear to rely on the tone shape difference.

1. BACKGROUND INFORMATION

1.1 *EN in Mandarin*

Tang (2016) provides a brief introduction on the particle *en* in the prescriptive grammar. Indeed, the particle “en” is an exclamation word, and it is usually placed at the beginning of the response during a casual conversation. The conversation pair (1) is a typical example of the usage of *en* because the particle is at the beginning of the response from Y:

(1) X: 我今天好累。

Y: 嗯, 我看出来了。

(X: wo jintian hao lei)

(Y: en, wo kanchulaile)

[X: I am so tired today.]

[Y: en, I feel that.]

Depending on the intended meaning, the pronunciation varies. Theoretically, there are two pragmatic meanings for “en”: the agreement and the disagreement (which may also indicate doubt). Three pairs of conversations in Mandarin (also with English translation)

* Acknowledgement: Thanks Dr. Emily Elfner from York University to be my supervisor for this paper.

provide the common usage of the two pragmatic meanings:

(2) Agreement

A: 米饭很好吃。

B: 嗯, 我同意。

(A: mifan henhaochi)

(B: en, wo tongyi)

[A: Rice is tasty.]

[B: en, I agree on you.]

(3) Disagreement

A: 你挺喜欢玩牌的吧。

B: 嗯, 我从不玩的。

(A: ni tingxihuan wanpaideba)

(B: en, wo congbu wande)

[A: You really like playing cards.]

[B: en, I never play it.]

(4) Doubt, but in this project it is recognized as *disagreement*

A: 你明天要去旅游了吧。

B: 嗯, 你说的啥?

(A: ni mingtian yaoqu lvyouleba)

(B: en, ni shuodesha)

[A: You will go for journey tomorrow.]

[B: en, what did you say?]

When the speaker tends to express his/her agreement, *en* will be pronounced with the falling tone and a short vowel duration. On the other hand, the particle will be pronounced with the rising tone and a long vowel duration if the speaker wants to express his/her doubts or disagreements.

1.2 Previous Research on *En*

Based on Tang (2016) and my observations, it is possible that there is a regional difference in the pronunciation of this particle between southern and northern China. Southern speakers tended to use the pitch contrast: falling vs. rising, while speakers from the north appear to rely on the syllable duration to indicate the difference between the agreement and disagreement meanings. On the basis of this observation, I piloted a study on this particle with Professor Schertz at University of Toronto in 2019. The expectation at that time was, “there is a systematic difference between southern and northern usage on this particle”. Specifically, southern speakers were predicted to use the falling pitch tone when they tended to express their agreement. However, northern speakers will lengthen their vowel duration as they would say their doubts or disagreements. Then, we recruited 20 speakers to examine. 11 speakers were from the southern provinces of China, and 9 of them were from the northern areas. All of participants were second language speakers of Mandarin

because they learned it after the age of 6. They were asked to read the sentences presented on a Power-Point. On the slides, there were two sentences: an opinion and a response. At first, they interpreted the opinion, and then they would read the response. The particle was always placed at the beginning of the response prompt.

The result of the piloting research did not match the expectation. Northern speakers did not really use the vowel duration contrast, and southern speakers also did not rely on the pitch difference. Therefore, it was very difficult to draw any conclusions on the particle usage in different regions.

During the year of 2020 in Professor Ruth King's class, I re-analyzed my previous data with L2 speakers, and I had a new finding. The proposed variation was shown in some individuals. For instance, a northern speaker really used the vowel length contrast when he read the response prompt. On the basis of the re-analysis, I revised my methodology of the study from reading sentences to having sociolinguistic interviews. The reason for the revising is that interviews can ensure the casual speech style when the particle usually occurs. Additionally, it was difficult to categorize speakers based only on the general geographical distinction between the south and the north, so I only focused on the speech in one city: Guangzhou. Five female speakers from that city were recruited to have a 10-minute casual interview. All of them were the L1 speakers of Mandarin though some of them could speak Cantonese fluently. The result of that study indicated that they preferred using the falling pitch in saying the agreements; however, they relied on the vowel duration contrast when they said their disagreements. Individually, none of them used the rising tone in the agreements, but some of them would use it in saying the disagreements. From their linguistic and social backgrounds, most of them had contact with speakers from Shanghai, Beijing and the northeast China. Thus, it was possible to explain their variation was due to the social interaction (Milroy, 1978) with those speakers. Therefore, to examine the speech variation on the particle *en* in those cities is necessary, and it will be the main focus of this paper.

From my previous work on the particle *en*, L2 speakers were not likely to show the variation on the usage of the particle whereas L1 speakers tended to. L1 female speakers from Guangzhou used the falling tone in saying the particle to express their agreements, but they shifted to lengthen the vowel as they would disagree on the interviewer. Although it is possible to conclude that the reason for the variation was the social contact with speakers from three cities, it is required to examine how speakers from cities of Shanghai, Beijing and the northeast China (Shenyang is the representative) behave. This is the topic of the current paper.

1.3 Guangzhou, Cantonese

Leung (2008) conducted both production and perception experiments on the tonal assimilation on L2 Mandarin speakers whose L1s were Cantonese. Her results on the production side indicated that Cantonese speakers had difficulty producing rising tones, but in terms of their perception, they had trouble perceiving Tone 4 (falling tone) in Mandarin.

In terms of the syllable duration, Kong (1987) provided thorough work on the relation between the f_0 and syllable duration. Kong (1987) separated tones into three groups: the longest, middle and the shortest. The first group (longest) includes tones 1 and 3. Then in

the middle group, the low-rising tone (Tone 5) and Tone 3 (mid-level tone) are included. In the last group, the shortest the one, other two level tones (tones 1 and 6) are involved. Generally, Kong (1987) argued that the high rising tone (Tone 2) has the longest duration on syllables. In fact, Kong (1987)'s generalization occurs in Mandarin too because the high-rising tone in Mandarin is also the longest.

Therefore, Cantonese speakers tended to have trouble producing the rising tones, and the rising tones are always the longest. In terms of the project of this report, it is predictable that speakers of Cantonese-accented-Mandarin¹ might omit the tone production then shift to syllable duration only if they would express the particle *en* for doubts or disagreements.

1.4 *Shanghai, Shanghai Dialect*

On the side of syllable durations, Duanmu (1994) should give sufficient background information. Duanmu's work (1994) was a comparison of syllable length between Mandarin and the Shanghai Dialect. Theoretically, the rhyme of Mandarin consists of two vowels, but in the Shanghai Dialect, there is only one vowel. Thus, the rhyme of the Shanghai Dialect is lighter than the rhyme of Mandarin. Duanmu's analysis showed the statistical significance of the syllable duration between the Shanghai Dialect and Mandarin. The length of the former is definitely shorter than the syllable duration in Mandarin. The average length of a Mandarin syllable is around 215ms, but in the Shanghai Dialect, it is around 162ms. The other feature considered by Duanmu (1994) was the coda. In Mandarin, only the alveolar and velar nasal can occur in coda position, but in the Shanghai Dialect, both the nasal sounds and the glottal stop can occur the coda position. The final result indicated that none of the syllable duration was not affected by adding codas.

As the consequence of their single mora in the rhyme, it is possible to examine if speakers from Shanghai prefer the vowel duration contrast because they can simply add one mora to their rhyme.

1.5 *Beijing, Beijing Mandarin*

Beijing Mandarin is the variety widely spoken in the city of Beijing. Although the Beijing variety is somewhat similar to the standard variety, there are differences between them (Zhang, 2005). Therefore, the variety of Mandarin in this section focuses on the Beijing variety instead of the standard variety.

Dell (2004) on the tone sandhi of Beijing Mandarin suggests that unstressed syllables have neutral tones, and their duration is much shorter than full syllables. Dell (2004), like Duanmu (2000), proposes that full syllables are with two mora while reduced ones are with one mora only. In addition to the specific features of the neutral tone, Lee (2003) indicates that the fundamental frequency f_0 of the neutral tone is alternative, and it depends on the tone that follows in the previous syllables. In terms of its duration, the range is around 100 to 120ms. Therefore, there is an alternative way to predict the performance of *en* in Beijing variety: the neutral tone. Theoretically, grammatical particles are usually produced as unstressed, and as previously mentioned, *en* is a grammatical particle, so *en* may be produced with the neutral tone. If *en* is pronounced with the neutral tone, native speakers of the Beijing variety are expected to use vowel duration to distinguish the agreement and

¹ Some of them are native speakers of Cantonese, but all of them learned Cantonese-accented-Mandarin.

disagreement.

1.6 *Shenyang, Northeastern Mandarin*

The last city to examine is Shenyang which is the capital city of Liaoning Province. Liaoning Province is located at the northeastern part of China. The Northeastern variety of Mandarin is widely spoken in the city of Shenyang. According to Cui and Kuang (2020), the tonal pattern of the Northeastern variety is similar to the Beijing variety except for Tone 1 (high-level tone). Therefore, it is predictable that Northeastern speakers will rely on the vowel duration to distinguish the meanings of *en* because of the similarity of tonal patterns between the Beijing and Northeastern varieties.

1.7 *Hypothesis*

The previous works on the particle *en* indicated that it is a response token used in casual conversations. It usually occurs at the beginning of the response to indicate the agreement or the disagreement towards the other speaker's opinion. There are two phonetic dimensions that are of interest on this token: tone pitch (f0) and vowel (syllable) duration.

Indeed, the main focus of this study is still about the difference in using the exclamation word *en* in different regions. Specifically, this study aims to test which dimension is favored in producing responses between southern and northern speakers. Furthermore, previous researches suggest that there may be regional differences as based on sources discussed above. Based on Chen (1991), Guangzhou speakers are likely to rely on the tone difference because of the tone sandhi of Cantonese. However, other sources pointed out that they had issues in pronouncing the rising tones, so that is also possible to predict their usage of vowel duration contrast in saying the disagreement. Shanghai speakers are expected from their high reliance on the vowel duration contrast. On the side of the Beijing variety, they are expected to rely on the vowel duration contrast. Finally, Northeastern Mandarin speakers are predicted to rely on the vowel duration contrast.

2. METHODOLOGY

2.1 *Participants*

Forty participants were recruited in total. Specifically, there were ten speakers from each city: Guangzhou, Shanghai, Beijing and Shenyang. Inside those speakers, five of them were males, and five of them were females. They were born between 1990.1.1 and 2002.3.1. Thus, all of them were young speakers around ages from 18 to 31 at the time of recording. Participants did not have to be born in the cities above, but they must have lived in those cities for at least 10 years. They did not have to go to the high school in the city or even in China, but they must have finished their primary and secondary studies in the city shown above. Also, they did not have to acquire Mandarin from their parents, but they must have acquired Mandarin by the age of 6, or they must have acquired it when they were in Grade 1 of their primary school studies. Some of them learned their regional dialects from their family members (e.g. Cantonese, Shanghai Dialect, Beijing Mandarin and Shenyang Mandarin). All participants were compensated for their time. Some of them were my

friends, and some of them were introduced by my friends too. Others were introduced by my parents' close friends.

Twenty participants were from the southern part of China, and other twenty of them were from the north. Geographically, the boundary to distinguish the south and the north is the Qinling-Huaihe band (Lin, 2009). Cities or provinces from the northern side of the border are recognized as northern places; similarly, places located in the southern side of the boundary are treated as the south. Referring Figure 2, cities of Guangzhou and Shanghai are on the southern side of the border, and other cities (Beijing and Shenyang) are on the northern side.

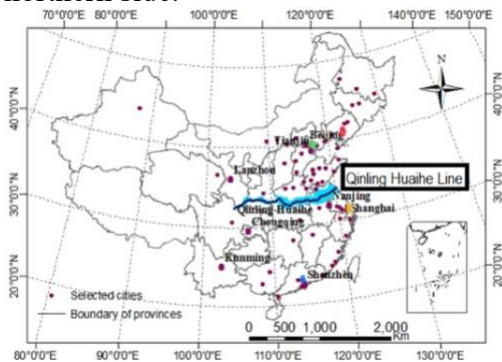


Figure 1: Qinling-Huaihe band extracted from *Chinese National Geography*. The blue point indicates the location of Guangzhou, the yellow dot is Shanghai. The green and red points mean the places of Beijing and Shenyang respectively.

2.2 Method — Sociolinguistic Interview

2.2.1 Virtual Interview and Questions

Participants were asked to have a casual interview with the researcher. The reason for choosing the sociolinguistic interview is the particle *en* commonly exists in the casual speech. The interview creates a circumstance which is similar to the context of a casual conversation. Thus, participants are likely to use the particle in their responses.

The structure of questions was also designed carefully. At first, there were a lot of tag questions (e.g. *You sleep at home, don't you*) in the interview because Oralova (2016) suggested that the particle is likely to occur in answers when the tag questions are asked. In addition, there are different types of tag questions in Mandarin, and Oralova (2016)'s suggestion of tag questions is *duiba* (is it right) while in the previous interview questions, the tag questions were *bushima* (isn't it). However, many interviewees indicated that they did not feel natural towards them, and indeed, they felt odd. One of the interviewees informed me that he believed those questions are common in English whereas they are seldom in Mandarin. Thus, questions involving a particle *ba* at the end were used because Oralova (2016) also indicated that the particle could exist in responses followed by questions involving *ba*. At this time, none of the participants felt odd in the interview; instead, they were relaxed. Therefore, the tag questions were used carefully.

2.3 Data Analysis

2.3.1 Measurements and Analysis with R-Studio

Acoustic measurements are done to examine both the pitch and duration of each tokens. The software Praat was used for the acoustic measurement. A Praat script was used involved for the pitch measurement, it tracked both the maximum, minimum, and mean pitch of each particle. Furthermore, it also tracks the pitch progressively across time to indicate the change in pitch across the particle.

R-Studio (RStudio Team, 2015) is a software program for statistical data analysis. For this project, we used R-Studio to analyze the data in four ways: individuals, cities, regions and the whole data set. For individuals, we analyzed the *en* particle's durations and pitches. Semitones were used to normalize the tones. In addition, the time was also normalized. The vowel in the particle *en* was divided into four quadrants: 0.25, 0.5, 0.75 and 1.

3 RESULTS AND DISCUSSION FROM THE GENERAL VIEW

3.1 Results from the Whole Participant Group

3.1.1 Number of Particles

Table 1 displays the number of tokens of *en* provided by the whole speaker group.

	Agreements	Disagreements	Total
Tokens	646	371	1017

Table 1: Number of tokens from each discourse labels.

According to the result shown in Table 1, speakers used the particle with the agreement meaning more than the disagreement meaning. In terms of their proportions, 63.5% of the total tokens are the agreement meaning. Conversely, 36.4% of the tokens are used as the disagreement meaning. In general, Mandarin speakers preferred using the particle *en* to express their agreements rather than their disagreements in this study. In other words, *en* appears to be more favored to show the pragmatic meaning of agreements instead of disagreements.

3.2 Results from Two Regions: South vs. North

3.2.1 Duration

The results for both regions are statistically significant in terms of the duration of token *en*. Figure 2 is the result for the comparison between two regions.

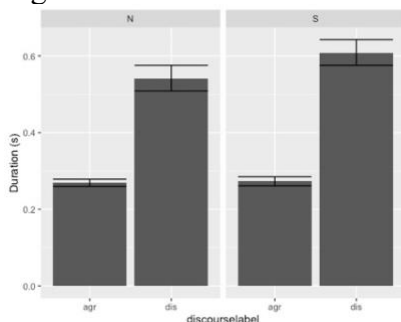


Figure 2: Bar plot of the duration comparison between the north and south

The mean duration for the agreement tokens from the south is around 270ms while the length for the disagreements is longer and it reaches 601ms. Both of the average values are consistent with Oralova's (2016) range. From the results about the north, the average length of disagreement tokens is longer than the agreement tokens: the duration of the disagreement tokens is around 530ms, but the duration of the agreements is about 260ms.

3.2.2 Pitch Process Tracking

The pitch tracking in Figure 3 can provide deep information for the tone shapes of both regions.

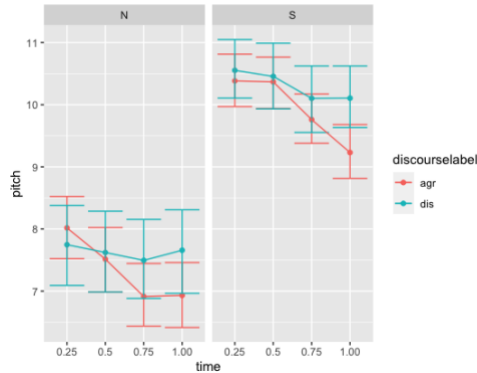


Figure 3: Tone shapes of regions.

By analyzing two lines in Figure 3, lines are parallel in the southern participants while there is a cross in the north. The tone shape of the agreement tokens in southerners is the high falling tone, but the shape of the disagreement tokens may be the Tone 1 (the high-level tone). Even though the tone of the disagreement tokens drops in the first two divisions, it eventually levels in the last two parts. Additionally, the range of the f0 decrease is small, so the tone shape of the disagreements is similar to a high-level tone, Tone 1. In terms of the northern speakers, their tone shape of the disagreement tokens is level while the tone shape of the agreements is falling. Specifically, there is a slight rise for the tone shape of the agreements at the last two divisions. However, the variation range of the rising is small. Hence, the whole range of change for the agreement tone is relatively smaller than the southerners, so the tone shape of the agreement tokens in the northern region can be recognized as the low-level tone. Generally, the overall pitch of the northern speakers is much lower than the pitch of the southern participants. Therefore, the tone shape is important for both regions while it is more obvious for the southern group between the two pragmatic meanings. In terms of the northern group, the difference between the two meanings is not so clear as the southern group whereas there is a contrast between the high and low level tones.

3.2.3 Pitch Process Tracking in Gender

Figure 4 provides the pitch tracking of two regions into gender. Females in both regions do separate the particle meanings into tone difference. They used the high-falling tone to represent the meaning of agreement while they tended to use the low-falling tone to show their disagreements on the particle. On the other hand, the tone shape of males on two pragmatic meanings is different between the north and the south. Northern males both

used the low-level tone to involve the different meanings on the particle, whereas males from the south used the falling tone on the agreement meanings but the low-level tone on the disagreement meaning.

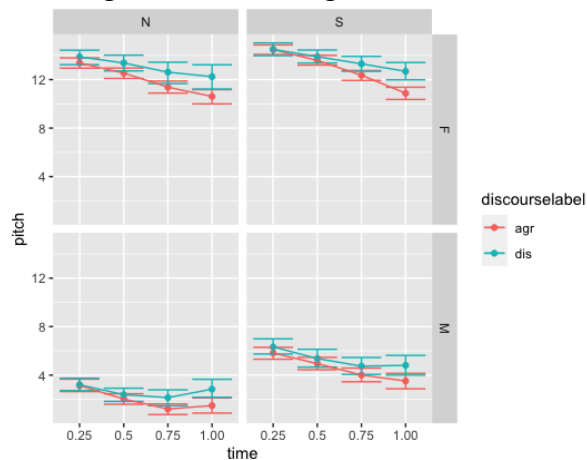


Figure 4: Pitch tracking of regions into gender.

4. RESULTS AND DISCUSSIONS FROM CITIES

4.1 Guangzhou

4.1.1 Duration

The result of the duration of the particle from Guangzhou participants is illustrated in Figure 5.

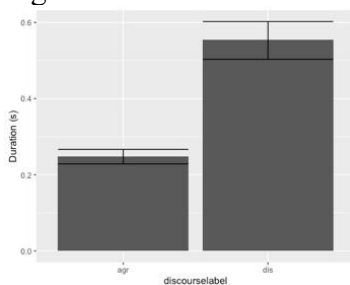


Figure 5: Duration result of Guangzhou participants.

According to the results, the duration of the disagreement tokens is much longer than the duration of the agreement tokens. Numerically, the length of the agreement particles is approximately 250ms while the duration of the disagreements is around 500ms.

4.1.2 Tone Shape Tracking (Including Gender)

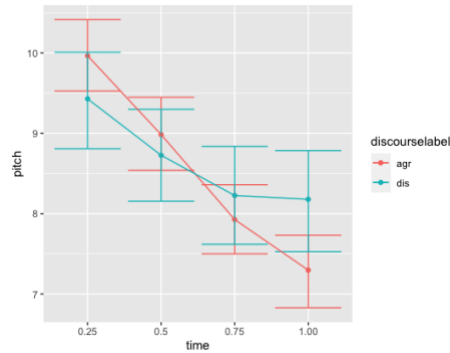


Figure 6: Exact tone process of Guangzhou participants.

From the lines in Figure 6, the slope of the agreement tokens is larger than the slope of the disagreement tokens. There is no interruption or change of the tone-line direction (agreement meaning) during the tone pitch change, and the tone is progressively dropping through four divisions. Thus, it can be recognized as the high-falling tone. However, the tone of the disagreements drops more slightly than the tone of the agreements, so, the tone shape of the disagreement tokens is likely to be the low-falling tone (which exists in Cantonese but not in Mandarin) because it debuts at a relatively lower pitch range to drop.

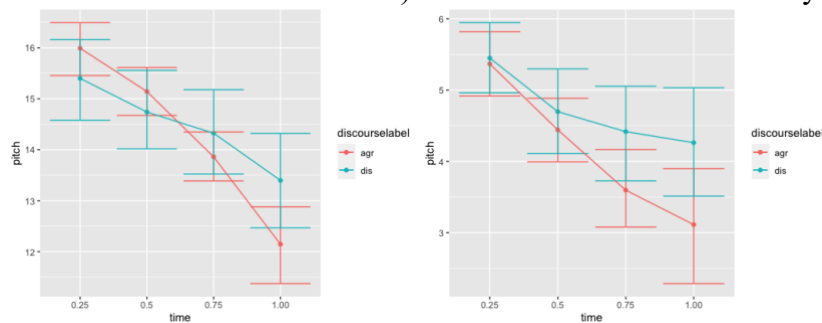


Figure 7 (on the left): Female tone shape from Guangzhou.

Figure 8 (on the right): Male tone shape from Guangzhou.

From results in Figures 7 and 8, the tone shape between women and men is different in terms of the disagreement meaning. For the agreement meaning, both groups used the high falling tone. In contrast, women used the low-falling tone for the disagreement meaning, but men tended to choose the level tone. Specifically, the pitch of females is consistently dropping through all quadrants, while the pitch of men is relatively level when it reaches the third quadrant.

4.2 Shanghai

4.2.1 Duration

The result for duration of the Shanghai group is shown in Figure 9.

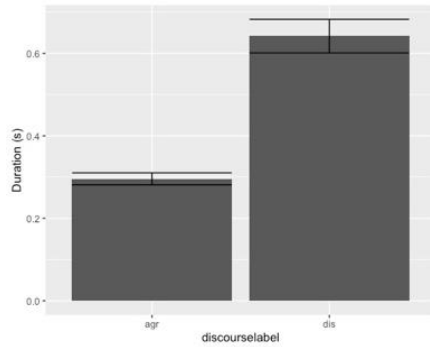


Figure 9: Durations of Shanghai participants in bar chart.

According to the result in Figure 9, the duration of the disagreement particles is much longer than the duration of the agreements. The length of the agreement tokens is 300ms while the duration of the disagreement tokens is around 620ms. Although both duration values are under the duration range given by Oralova (2016), the duration of the agreement particle is slightly longer than the duration in the Guangzhou group.

4.2.2 The Pitch Tracking (also with Gender)

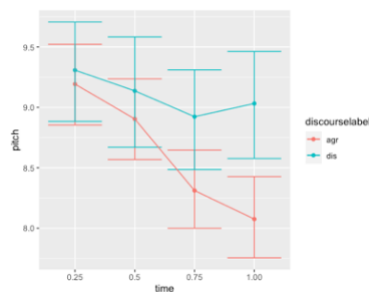


Figure 10: Pitch tracking of the Shanghai group.

According to the result shown in Figure 10, the agreement tokens appear to use the falling tone. In contrast, the tone of the disagreement tokens debuts at a similar pitch value of the agreement tone, but its slope is not steady as the slope of the agreement tone, and it rises slightly at the fourth quadrant. The change of pitch values for the agreement tokens is much larger than the disagreement tokens, and the small range of variation for disagreement tokens can be recognized as the level tone. Thus, the tone shape of the agreement tokens is the falling tone.

Gender is important for the pitch tracking, and the results of females and males from the Shanghai group are shown in Figures 11 and 12.

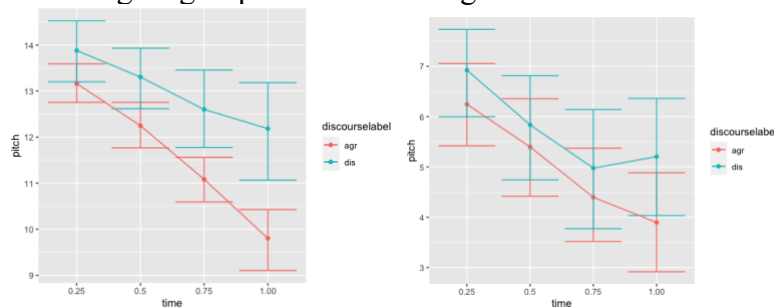


Figure 11 (on the left): Results of women from Shanghai in pitch tracking.

Figure 12 (on the right): Results of men from Shanghai in pitch tracking.

From the position of the error bars in Figure 11, the result of pitch tracking of Shanghai women may be statistically significant; they used the high falling tone to represent the agreement meaning, but with the disagreement meaning, they used the low falling tone, or the level tone. In contrast, Shanghai men used the falling tone in both meanings, and the error bar position in Figure 12 does not impact any statistically significant results.

4.3 Beijing

4.3.1 Duration

The results of the duration from Beijing speakers are presented in Figure 13.

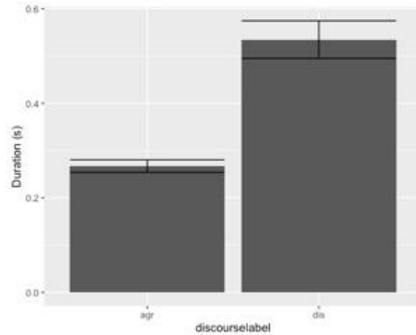


Figure 13: Durations on bar graph of Beijing participants.

Beijing participants definitely rely on the duration difference. When the duration is long, the particle means the disagreement; conversely, the token will mean the agreement as it is short. Recalling back to the hypothesis mentioned in Section 2.7, we predict that the duration is reliable for Beijing speakers, and our significant result confirms the prediction.

4.3.2 Tone Shape Tracking (with Gender)

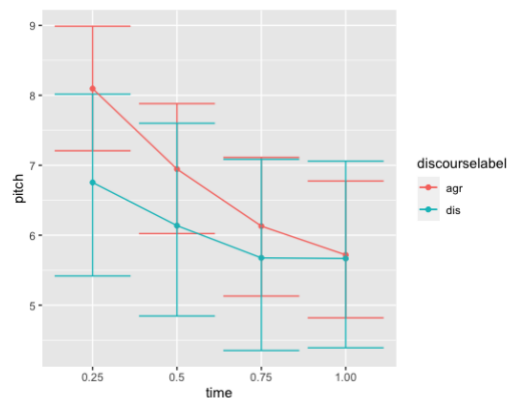


Figure 14: Exact tone shape of Beijing participants.

In Figure 14, the tone shape of the agreement tokens can be recognized as the falling tone while the tone of the disagreement particles is the level tone. The two lines are quite parallel, and there is no crossing between them until the final quadrant. Generally, the range of change in the agreement tone is larger than the range of the disagreement tone. Thus, the tone shape of the agreement tokens is the falling tone and the level tone is the shape of the disagreement tokens.

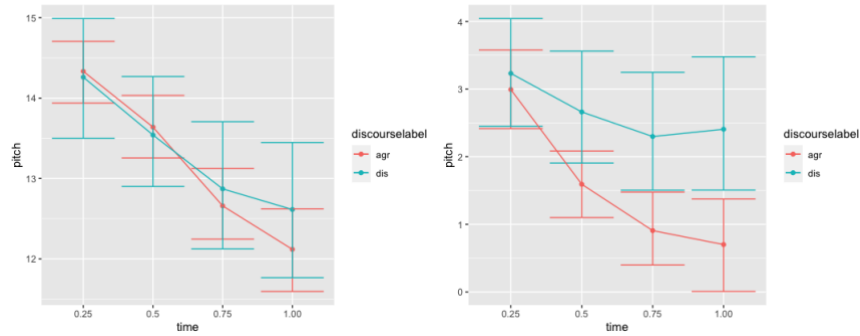


Figure 15 (on the left): Exact tone shape of Beijing female participants.

Figure 16 (on the right): Exact tone shape of Beijing male participants.

The pitch tracking result of gender in Beijing participants is the contrast of the Shanghai group. From the results in Figure 15, high falling tones are used by Beijing women with tokens of both pragmatic meanings. However, the separation of error bars in the second quadrant in Figure 16 indicates the statistical significance of tone shapes to Beijing males. Beijing men used the high falling tone on the tokens with agreement meaning, but they leveled their pitch when they had the disagreement tokens.

4.4 Shenyang

4.4.1 Duration

As previously mentioned in Section 2.8, Shenyang participants are supposed to rely on the length difference to distinguish the pragmatic meanings. The result is shown in Figure 17.

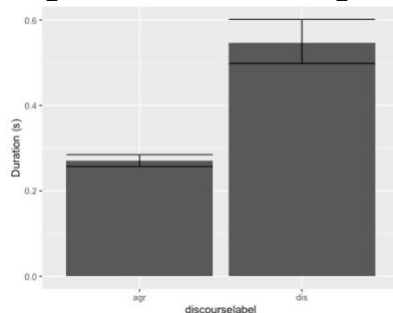


Figure 17: Duration result for the Shenyang group.

4.4.2 Tone Shape Tracking (Including Gender Difference)

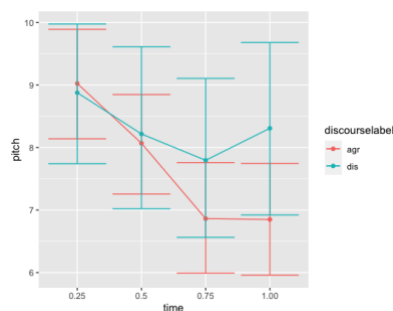


Figure 18: Tone tracking for Shenyang participants.

In terms of the line chart (Figure 18), the tone line of the agreement particles starts at the value of 9Hz and it drops into 6.8Hz in the third part. Eventually, it maintains steady around 6.8 Hz. The other line is the tone of the disagreement tokens. The pitch starts at a similar value to the agreement tone, and it is around 8.9Hz. Then, the blue line of the disagreement tone drops into 7.8Hz at the third division. Later, it raises to 8.3Hz at the end. Therefore, the value of pitch change in the agreement tone is much larger than the pitch change in the disagreement tone. Even though the pitch remains level, the tone shape of the agreement tone may be considered to be the falling tone (Tone 4). However, the tone shape of the disagreement tone might be the level tone (Tone 1) though it increases slightly at the end.

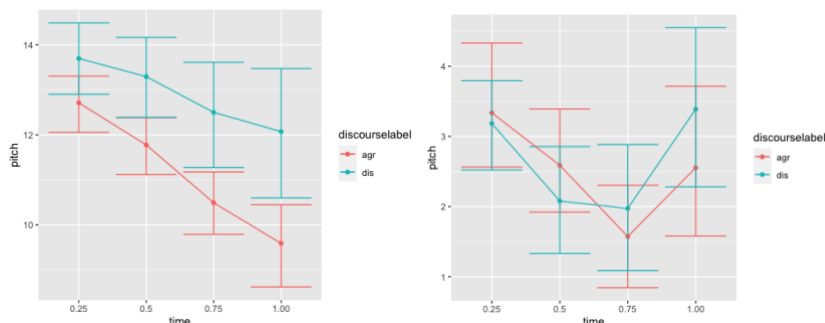


Figure 19 (on the left): Tone tracking for Shenyang female participants.

Figure 20 (on the right): Tone tracking for Shenyang male participants.

Results from Figure 19 indicate the statistical significance of the tone shape difference from Shenyang females. They had the high falling tone (Tone 4) when they said the token with the agreement meaning, whereas the tone shape of the tokens with disagreement meaning is leveled or fell slightly. In contrast, Shenyang men used the same tone with both pragmatic meanings on token *en*. From lines in Figure 20, their tones of two meanings both drop significantly until reaching the third quadrant. After that, their tones rise dramatically to the fourth quadrant. Therefore, tone shapes of two meanings from Shenyang men are similar to Tone 3 in Mandarin tones.

5. CONCLUSION

In summary, this paper reports on quantitative research investigating on a Mandarin particle *en*. The particle has two pragmatic interpretations in casual conversation. From some sources, it can mean the agreement and the disagreement (including the doubt). When the speaker expresses his/her agreement, the particle is produced with a falling tone and a short vowel (the duration of the vowel is around 100 to 300ms). On the other hand, the particle uses a rising or level tone and a longer vowel when the speaker tends to express his/her disagreement or doubt.

On the basis of two pilot studies and some previous studies, four cities were selected to examine further: Guangzhou, Shanghai, Beijing and Shenyang. Cities of Guangzhou and Shanghai are located in the south, but cities of Beijing and Shenyang are in the north part of China. The methodology of sociolinguistic interviews is selected to have a 15-to-30-minute casual conversation with forty participants. From the results of Guangzhou speakers, they rely more on the vowel duration differences. Hence, Guangzhou males may rely on the tone shape difference while females may not. Shanghai speakers were possible to rely on both duration and tone shape. Men from Shanghai are the opposite of women because

the tone shape result of men does not indicate a possible statistical significance. Beijing speakers only use the vowel duration contrast. In terms of gender, Beijing women distinguish the tone shape with different pragmatic meanings, but men use the same falling tone. Shenyang speakers also rely on the duration, but they may rely on the tone shape. Women from Shenyang rely on the tonal difference, but men are the opposite. Therefore, one city from each region shows the reliance on both duration and tone shape, so it is difficult to conclude that the difference is due to the region. Further research on more cities and more participants will help to confirm the conclusions of this paper.

References

- Boersma, P., and Weenink, D. Praat, Version 6.0.16. Retrieved from <https://www.fon.hum.uva.nl/praat/>.
- Chen, M. (1991). An Overview of Tone Sandhi Phenomena Across Chinese Dialects. *Journal of Chinese Linguistics Monograph Series, No. 3, Language and Dialects of China*. 111-156.
- Cui, P., and Kuang, J. (2020). Pitch Declination and Final Lowering in Northeastern Mandarin. *Proc. Interspeech 2020*, 1928-1932.
- Duanmu, S. (1994). Syllabic Weight and Syllabic Duration: A Correlation between Phonology and Phonetics. *Phonology*, 11(1), 1-24.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.
- Dell, F. (2004). On Recent Claims about Stress and Tone in Beijing Mandarin. *Cahiers de Linguistique-Asie Orientale* 33(1), 33-63.
- Kong, Q. M. (1987). Influence of Tones upon Vowel Duration in Cantonese. *Language and Speech*, 30(4), 387-399.
- Lee, W. (2003). A Phonetic Study of the Neutral Tone in Beijing Mandarin. In *Proceedings of the 15th International Congress of Phonetic Sciences (ICPHS 2003)*. Barcelona. 1121-1124.
- Leung, A. S. (2008). Tonal Assimilation Patterns of Cantonese L2 Speakers of Mandarin in the Perception and Production of Mandarin Tones. In *Proceedings of the 2008 CLA Annual Conference*. 1-15.
- Lin, Z. (2009). 中国南北自然分界是“带”不是“线”. Chinese National Geography. From the website: <http://cng.dili360.com/cng/jcix/2009/06082227.shtml>.
[The Natural Distinction of Southern and Northern China is a Band instead of a Line.]
- Milroy, J., and L., Milroy. (1978). Belfast: Change and Variation in an Urban Vernacular. *Sociolinguistic Patterns in British English*, ed. by Peter Trudgill, 19-36. London: Edward Arnold.
- Oralova, G. (2016). *Minimal Response Token en in Mandarin Conversation*. MA Thesis on University of Alberta. 1-111.
- RStudio Team. (2015). *RStudio: Integrated Development for R*. RStudio, Inc.. Boston, MA. URL: <http://www.rstudio.com/>.
- Tang, H. (2016). 谈汉语口语中“嗯”的多义性. *Journal of Language and Literature*, 69-71.
[Talking about the diverse meaning of “en” in spoken Mandarin]
- Zhang, Q. (2005). A Chinese Yuppie in Beijing: Phonological Variation and the Construction of a New Professional Identity. *Language in Society* 34:431-466.

References of Figures

- Figure 1: Lin, Z. (2009). 中国南北自然分界是“带”不是“线”. Chinese National Geography. From the website: <http://cng.dili360.com/cng/jcix/2009/06082227.shtml>.
[The Natural Distinction of Southern and Northern China is a Band instead of a Line.]