Word count: 8503

The perception of lexical tones in emotional speech by Dutch learners of Mandarin

Yachan Liang

Utrecht University

MA thesis

Prof. Dr. Aoju Chen

August 2018

Author's note

This thesis is a required component of the graduate research programme in Linguistics at Utrecht University. The paper has been reviewed and graded by supervisor and first reader prof. dr. Aoju Chen and second reader Dr. Koen Sebregts.

This research is supported by personal funding to supervisor prof. dr. Aoju Chen. Special thanks to prof. dr. Aoju Chen for her support and supervision of this thesis.

Correspondence regarding this thesis should be addressed to Yachan Liang at: y.liang@students.uu.nl.

Acknowledgements

I would like to thank everyone who helped me during the whole process of designing the experiment and writing the thesis.

Firstly, I would like to express my gratitude to my supervisor, Aoju Chen, for her guidance and supervision throughout all stages in my internship and the thesis writing process. Thank you for the invaluable feedback on my thesis, and the grant for my experiment.

Secondly, my special thanks go to all my participants from Chinese Education Center Netherlands, Leiden University, and Utrecht University. Further, I would like to thank Shuangshuang Hu, who helped me to record the stimuli for the experiment.

Last but not least, I would like to thank Jacco van Elst, Maarten Duijndam, and Iris Mulders who gave me endless support in my experiment.

Utrecht, 2018

Abstract

The current study examined the perception of Mandarin tones in different emotions by Dutch learners of Mandarin. As a tone language, Mandarin has both lexical tones and emotional intonation, which interact with each other. An interesting question arises as to how the co-existence of lexical tones and emotional intonation influence the perception of lexical tones in emotional speech by learners of Mandarin with a non-tone language as their L1. Past work on the acquisition of Mandarin tones has exclusively focused on lexical tones produced in neutral emotion. Consequently, current knowledge of the learning of Mandarin tones is confined to the citation forms of the tones and may not be generalizable to the perception of tones in natural speech where speakers often express an emotion other than the neural emotion. Against this background, we have investigated development in tonal perception in emotional speech in a second language learning setting, with the goal of putting forward a revised model of tonal acquisition from a perceptual perspective. In the current study, Dutch learners of Mandarin (beginners and advanced learners) were recruited as the experimental group, and native Mandarin speakers were also included in this study as the control group. All participants were asked to take part in the tonal perception task. In line with the previous studies, this study found that emotions affect the perception of Mandarin tones by Dutch learners of Mandarin. Additionally, other factors such as tones, and language proficiency also influence the perceptual performance to some extent. However, tonal context affected different proficiency groups differently, influencing advanced learners more than beginners.

Keywords: Mandarin, Dutch, lexical tones, emotional intonation, perception

Contents

Contents

List of Figures and Tables

1. Introduction

Over 70% of languages in the world are tone languages, which use tone to distinguish lexical meanings (Yip, 2002).

Mandarin has both lexical tones and emotional intonation, which interact with each other. Prosodic parameters such as pitch and duration are used to realize both lexical tones and emotional intonation. The same lexical tone can have very different surface forms in different emotions, compared to its citation form (Chao, 1968). The situation is different from that in non-tone languages, such as Dutch, which uses pitch and duration primarily to express post-lexical attributes, such as emotions, illocutionary force and information structure. The co-existence of both lexical tone and emotional intonation in Mandarin can pose big challenges for the acquisition of the lexical tones by non-tone language learners.

Notably, previous studies on the acquisition of lexical tones by non-tone language learners mainly focus on the perception of lexical tones in isolation or in carrier sentences. No perception research in this area has thus taken emotions into consideration. Against this background, this study aims to investigate the perception of Mandarin tones in emotional speech by Dutch learners of Mandarin.

This paper contains seven sections. In section 2, “literature review”, previous research on the perception of Mandarin tones is reviewed. The research questions and hypotheses are formulated in section 3. In section 4, the experimental design and procedure are described. This is followed by a report of the results in section 5. In section 6, we discuss the findings of the current study. In the final section, we propose suggestions for further research.

2. Literature review

Before we propose the research questions and hypotheses, we first discuss the perception of Mandarin lexical tones by non-tone language learners, the interface between lexical tones and emotional intonation in Mandarin, and the emotional intonation of Dutch in detail in this section. In this study, we will use pitch when talking about production and perception, and use f0 when describing the acoustic features of the sound.

2.1 Previous research on the perception of tones by non-tone language learners of Mandarin

It is generally accepted that tones are of great difficulty for non-tone speakers who learn Mandarin (Wang et al., 1999, 2003; Francis et al., 2003). In the past few decades, plenty of studies have examined tonal perception by both tone and non-tone language speakers (Krishnan, Xu, Gandour & Cariani, 2005; Xu, Gandour & Francis, 2006). Both behavioral and neuroscientific studies have found that non-tone speakers process pitch information differently than do tone language speakers (Francis, Ciocca, & Ng, 2003; Halle, Chang, & Best, 2004). In Gandour's experiment (1983), he compared tonal perceptual patterns of English, Mandarin and Cantonese speakers and found that English speakers paid more attention to the height and less to the direction dimension than did listeners from tone languages (Mandarin and Cantonese). Further, Content & Noémi (2011) found that tone language speakers such as Mandarin, Cantonese, Thai, and Vietnamese showed a categorical perception (CP) effect when perceiving lexical tones. Braun and Johnson (2011) found that Mandarin and Dutch listeners attend to the same pitch movements with different locations on a segmental string. Mandarin speakers were more sensitive to the rising and falling pitch contours on the initial and the final syllables in a disyllabic non-word, which signal two different lexical tones in Mandarin (i.e. the lexical rising and falling tones). In contrast, Dutch speakers were more attentive to pitch movements in the final position than in the initial position in tonal perception, since a Dutch final pitch movement serves as a

salient cue for paralinguistic meanings, such as question vs. statement (Van Heuven & Krisner, 2004). Prior studies show that tone and non-tone language speakers rely on different phonetic cues in distinguishing lexical tones. For example, the primary cue for tonal contrasts is f0 contour for native Mandarin speakers (Xu, 1997; Liu & Samuel, 2004). Thus, Mandarin speakers pay more attention to the f0 contour, while English speakers seem to emphasize the height of Mandarin tones (Wang et al., 2003). With respect to f0 height and contour, research shows that the perceptual weights of these two dimensions are related to listeners' linguistic experience (Gandour, 1983). Therefore, Gandour (1983) suggested that the reason why English listeners made more use of height is that there are no tones in their native language. Hence, it is proved that listeners' perception of pitch movements is shaped by the way pitch information is used in their native language. According to the previous studies, it is found that tone language speakers perform better than non-tone language speakers in tonal perception tasks (Wayland & Guion, 2004).

In terms of difficulty in perceiving certain tones, Leather (1987) compared the identification of Mandarin Tone 1 and Tone 2 by native English and Dutch speakers with Mandarin listeners. The result indicates linguistically inappropriate perceptual weighting of f0 by native Dutch and English listeners compared with Mandarin speakers, in line with Gandour's (1983) finding. These experiments found that tone language speakers were better in discriminating tones in terms of accuracy rate and speed, than were English speakers, which indicated that tone language speakers acquire the ability of tonal discrimination. Therefore, to some extent, the ability of tonal perception relies on the linguistic function of pitch in their native language.

Besides, non-native tonal perception is subject to the influence of linguistic context and sentence position. Broselow et al. (1987) examined English listeners' perception of Mandarin tones when tones were presented in isolation as well as in the context of

two to three syllables. They found that the identification of Tone 4 varied according to its position in the sentence. For example, Tone 4 was the most easily identified tone when presented in isolation, and in the final position of a sequence of two or three syllables. However, its identification declined dramatically in non-final positions. The authors claimed that the reason is the interference from English intonation. Tone 4 is acoustically similar to the unmarked pattern of declaratives at the end of a sentence in English since both involve a falling pitch. Thus, it is believed that Tone 4 is easier to perceive when it occurs sentence-finally than in other positions.

These experiments found that non-tone listeners' perception of Mandarin tones is affected by their native intonation system. This is consistent with the findings of other studies showing the influence of English stress on the perception of Mandarin tones. White (1981) found that English listeners tended to perceive the Mandarin high tones as stressed and the low tone (e.g. Tone 3) as unstressed.

Two theories have been proposed to explain this difficulty. According to the "level of representation" account, non-tone language speakers are unable to associate lexical tones with their L1 linguistic categories because of the differences between their L1 and L2. In contrast, the "category assimilation" account attributes this difficulty to difficulty in relating L2 tone categories to native ones. According to Ladd (1992, 1996), non-tone language speakers might process tones as intonation in their L1 because of the mismatch between their native intonational language and L2 tone categories.

2.2 Prosodic expression of emotions

Ekman (1992) proposed that emotions are internal feelings of individuals. Verbal communication implies that expressions are made intentionally to send a message. Emotional meanings are encoded along a set of benefit-oriented bio-informational

dimensions which involve both segmental and prosodic aspects of the vocal signal (Xu, Kelly & Smillie, 2013: 33). Anger, happiness, sadness and neutrality are among the most basic emotions and are relevant to the current proposal.

Murray & Arnott (1993) found that people can distinguish these emotions in terms of pitch range, intensity range, and duration across languages. Anger and happiness are the easiest to distinguish. Previous studies found that there are some similarities in emotions in terms of f0 cross-linguistically. For example, both anger and happiness contain higher pitch than neutrality, but there is no agreement as to which of these two has higher pitch. However, anger has a steeper pitch fall than happiness across languages, and sadness is characterized by a reduced pitch range, pitch variability, intensity, and speech rate in different languages (Ellgring & Scherer, 1996).

2.3 The interface between lexical tones and emotional intonation in Mandarin

In this section, I will present an overview of the tonal system of Mandarin, and its interaction with emotions in natural speech.

2.3.1 Lexical tones in Mandarin

As mentioned earlier, Mandarin has four lexical tones. For example, the syllable *ma* has different meanings when it is pronounced in different tones namely *mā* (妈) (mother), *má* (麻) (hemp), *mǎ* (马) (horse), *mà* (骂) (scold). The f0 of a syllable varies in one of the following ways: flat, rising, falling and rising, and falling. According to the five-point scale system proposed by Chao (1968), the tone value of Tone 1 is 55 (high tone), Tone 2 is 35 (rising tone), Tone 3 is 214 (falling-rising tone), Tone 4 is 51 (falling tone). Figure 1 shows the f0 of the four Mandarin tones, each spoken as the syllable *ma* produced by a native Mandarin speaker.

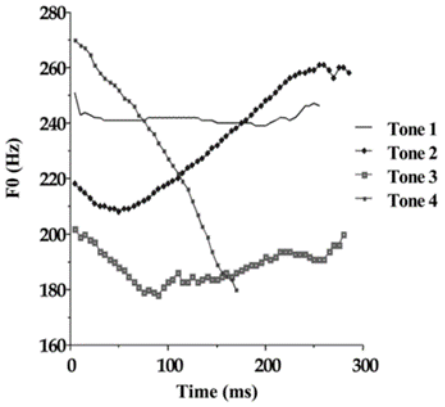

Figure 1. The f0 contours (Hz) for each of the four Mandarin Chinese tones for the segmental context *ma* spoken in isolation by a female speaker (Moore & Jongman, 1997: 1865)

Mandarin lexical tones are widely believed to be predominantly distinguished from each other by the pitch contour (Duanmu, 2007). Native Mandarin speakers use the combination of f0 and durational cues to distinguish the four tones (Xu & Wang, 2001; Yip, 2002). Additionally, the duration of syllables, the intensity, and the voice quality also provide information about the tone category of a syllable (Blicher et al., 1990; Cabrera et al., 2014). As is shown in Figure 2, the duration differs among the four tones, the longest being Tone 3, and the shortest being Tone 4 (Lin, 1965; Chuang et al., 1972; Howie, 1976; Nordenhake & Svantesson, 1983). The relative duration of tones may change according to the function of their position in the sentence (Nordenhake & Svantesson, 1983). The location of the turning point is the primary cue to the distinction between Tone 2 and Tone 3. According to previous studies, the turning point and f0 together typically specify tonal categories. It has been proved that the turning point alone can change listeners' perception of Tone 2 and Tone 3 within a certain range of f0 (Shen & Lin, 1991; Moore & Jongman, 1997).

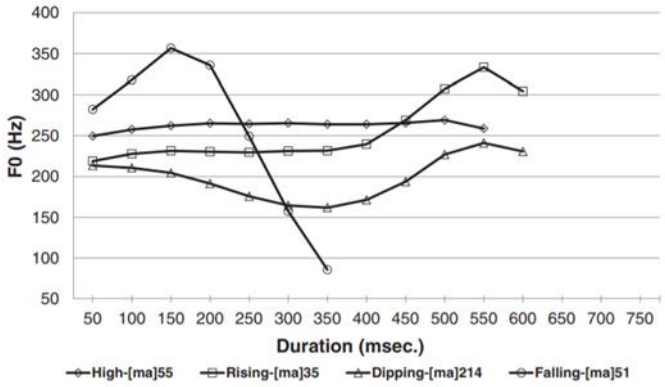

Figure 2. The duration of four Mandarin tones (Burnham, et al. 2014: 1461)

In addition to lexical tones, Mandarin exhibits tone sandhi, which is the phonological process governing the pitch patterns of tones in some tonal contexts (Yip, 2002). There are three main tone sandhi rules in Mandarin. Firstly, Tone 3 is realized with a rising contour similar to Tone 2 when followed by another Tone 3. Secondly, Tone 3 changes to a low falling contour when followed by a different tone, such as Tone 1, Tone 2, Tone 4 or neutral tone. Thirdly, Tone 2 changes to a high level tone if it is preceded by either a high level or other rising tone (Li & Thompson, 1981).

2.3.2 The interplay between lexical tones and emotional intonation in Mandarin

The emotional intonation of tone languages is an autosegmental element independent of lexical tones. However, these two elements are expressed via the same acoustic parameters with f0 being the most important one. The encoding mechanism of emotional speech in Mandarin is the interplay between lexical tones and emotional intonation. According to Chao (1968), Mandarin uses the successive addition tone as well as the edge tone to realize the emotional intonation. The successive addition tone refers to a rising or falling tone which is added successively after the completed lexical tone instead of added simultaneously to the lexical tone, which is shown in the following formula. The left column is for the rising additions and the right one is for

the falling additions. “1” - “5” are tonal values in 5 tone-letter scale. “6” refers to the extra high pitch. For example, the tone value of Tone 1 (55) becomes “56” after a rising tone is added to it.

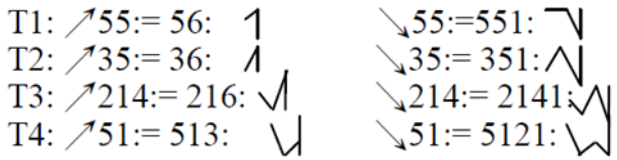

Figure 3. The successive addition tones in Mandarin (Chao, 1933: 122)

Li et al. (2011) investigated emotional intonations by analyzing monosyllabic utterances produced by one female speaker and one male speaker. They found that the tonal patterns are consistent with the two speakers for all emotions except for “anger”, where the male speaker has a much shrunken pattern in amplitude than the female speaker. They also found that the tonal space, the edge tone and duration differed significantly among seven emotions (“neutrality”, “sadness”, happiness”, “fear”, “surprise”, “anger”, “disgust”). As can be seen in Figure 4, “disgust” and “anger” are expressed via the use of an additional “falling” tone, while “happiness” and “surprise” are expressed via an additional “rising” tone. For Tone 4, the rising addition tone changes the slope of the fall in “happiness” and “surprise”. Besides this, there is a very tiny rising tail by the end of the boundary tone in Tone 4 in “happiness” and “surprise”. For the three other tones with high ending, there are additional rising tones. In terms of pitch range, “happiness” and “surprise” are realized with a higher pitch range and a higher register, while “sadness” and “disgust” are realized via a reduced pitch range and a lower pitch register. “Surprise” has higher bottom pitch, which has a slightly narrower range than “happiness”.

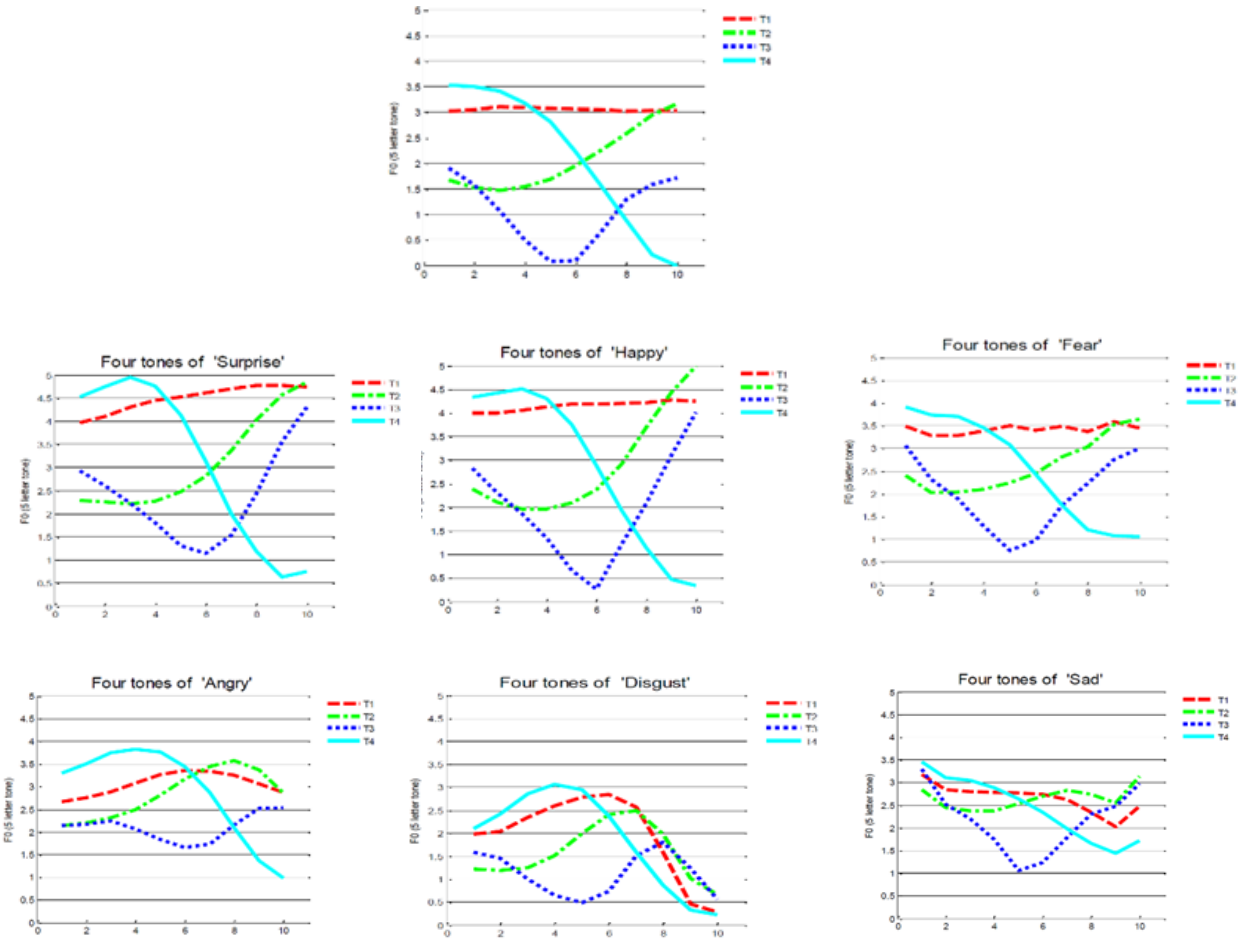

Figure 4. The f0 of seven emotional monosyllabic utterances by male with four tones in 5-tone scale (Li et al. 2011: 1199).

Li et al. (2011) also examined the expression of the seven emotions in multisyllabic sentences. They used a Chinese sentence "打高尔夫" (*da3 gao1 er3 fu1*) (translation: play golf) as an example to illustrate the interface between lexical tones and the prosodic expression of emotions. Compared with neutral emotion, the f0 patterns of the other six emotions differ significantly in terms of tonal registers, tonal range and tonal contours, as can be seen in Figure 5.

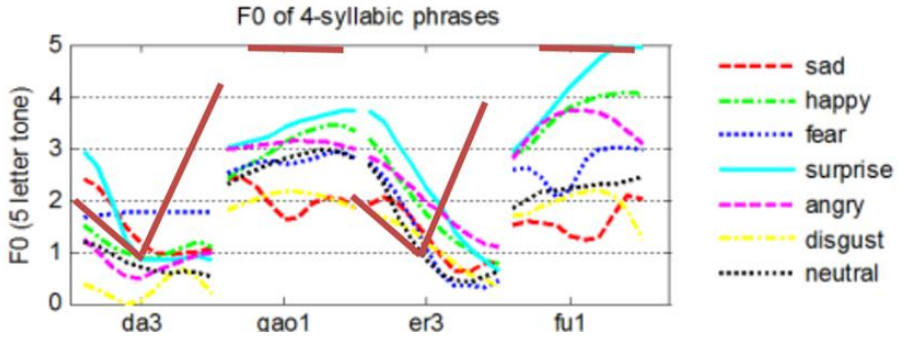

Figure 5. The f0 contours for the same sentence *"da3 gao1 er3 fu1"* (play golf) in seven emotions (Li et al., 2011:4). The dotted lines in colors other than black in figure 5 depicts changes in f0 due to the expression of emotions.

The red solid lines in Figure 5 show the lexical tone of each syllable at the lexical level (e.g. *da3*: Tone 3, 214; *gao1*: Tone 1, 55; *er3*: Tone 3, 214; *fu1*: Tone 1, 55). Different from the monosyllabic utterances, the lexical tone of each syllable changes slightly due to tone sandhi, though it is spoken in neutral emotion, as shown by the black dotted line.

Similar to the realization of tones in different emotions in monosyllabic utterances, the emotional intonation of the six emotions ("sadness", happiness", "fear", "surprise", "anger", "disgust") are expressed via the successive addition tone. However, the tonal space, edge tone and duration differ greatly across these seven emotions. "Happiness" and "surprise" are expressed by a rising successive addition tone (Chao, 1933), while "disgust" and "anger" are conveyed via a falling successive addition tone. In general, the positive emotion "surprise" has the highest pitch and uses the rising successive addition tone. Likewise, "happiness" also shows higher pitch than other emotions and the final tone is rising. The negative emotions such as "sadness, fear, anger" and "disgust" usually have lower pitch and use a falling successive addition tone. As for tonal register, "anger" and "happiness" have higher tonal register, while "sadness" and "disgust" have lower register.

2.4 The emotional intonation in Dutch

Dutch is a stress-accent language, which has a rather restricted pitch range (Hart et al. 1990). Dutch has ternary or binary trochees, with the main stress on the rightmost branching foot or on the final monosyllabic foot (Guessenhoven, 1993). As an intonational language, Dutch employs prosodic variations at the sentential level to convey communicative intentions. Mozziconacci (1995) examined the f0 of the Dutch sentence *"Het is bijna negen uur"* (*It is almost nine o'clock*) spoken by a male speaker in seven emotions, namely, "neutrality", "joy", "boredom", "anger", "sadness", "fear" and "indignation". These seven emotions were measured by frequency (Hz) and anchor points from 1 to 6 which includes: onset (1-2), two peaks (2 & 5), two values in the intermediate valley (after the first peak: 2-3, and before the second peak: 4-5) and offset (5-6). As shown in Figure 6, "indignation" and "fear" have the highest pitch register. While "neutrality" and "boredom" have the lowest pitch register. The other three emotions ("sadness", "anger", and "joy") fall between the highest and lowest pitch registers. In terms of pitch range, "indignation" and "joy" have the largest pitch range, while "boredom" and "neutrality" have the smallest pitch range.

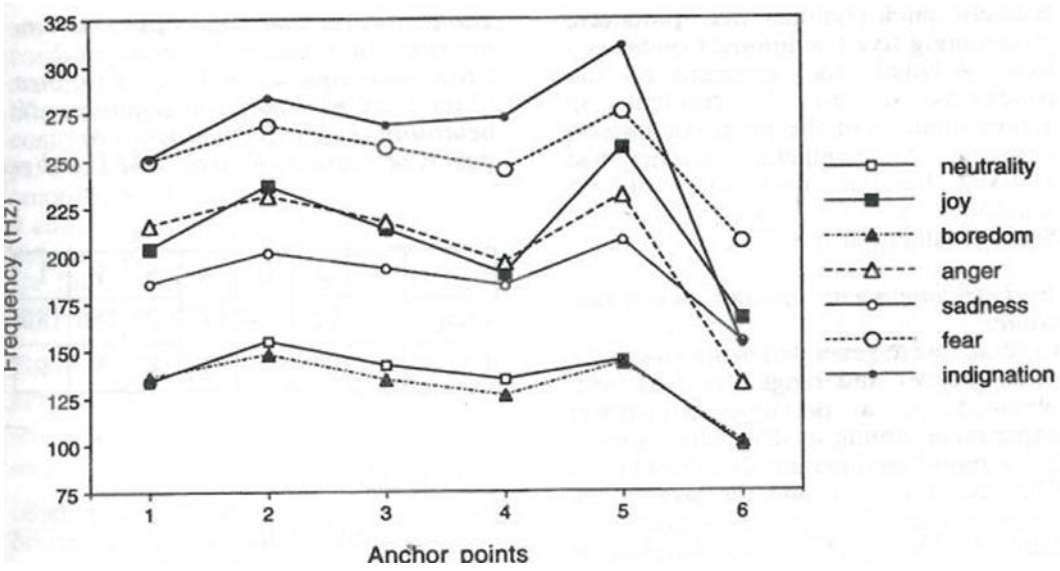

Figure 6: The f0 values at six anchor points in the sentence *Het is bijna negen uur* "It is almost nine o'clock". This is a copy of Figure 6 in Mozziconacci (1995: 179).

Comparing the f0 of emotional utterances in both Mandarin and Dutch, we found that emotional intonation changed the f0 in terms of pitch range, pitch contour and pitch register. In both languages, “neutrality” has the lowest pitch register and smallest pitch range. “Happiness” and “anger” have higher pitch registers and the smallest pitch range compared to that of “neutrality” in both Mandarin and Dutch. The other emotions falls between “neutrality” and “happiness” (or “anger”). However, compared with Dutch, Mandarin uses the successive addition tone to express the emotional intonation, which results in added tails when the lexical tones are finished.

2.5 Similarities and differences in Mandarin and Dutch in rises and falls

Every language uses pitch variation to convey paralinguistic meanings at a sentential level. In other words, rises and falls are common pitch patterns across different languages though pitch plays different roles in tone and non-tone languages. For non-tone languages, such as Dutch, pitch information can be used to indicate the prominence of words in a sentence (Birch & Clifton, 2002; Welby, 2003; Xu & Xu, 2005), and signal sentence types, such as statements and interrogatives (Pierrehumbert & Steele, 1989; Van Heuven & Haan, 2002). Generally, a falling pitch pattern indicates a statement, a rising pattern suggests questions. In addition, other paralinguistic information such as emotions (happiness, sadness, surprise) can be encoded with pitch movements as well (Chen, Gussenhoven, & Rietveld, 2004; Chen, 2005). In contrast, tone languages such as Mandarin, use pitch variations namely tones to distinguish lexical meanings as well as emotions. Chen et al. (2014) investigated the language specific phonetic realization of rises and falls in a production experiment by Chinese and Dutch speakers. This research found language specific patterns for the rise, but not for falls. For rises, both peak and valley were aligned later among Chinese speakers compared to Dutch speakers. In tone languages like Mandarin Chinese, f0 can function both lexically and post-lexically. However, in non-tone languages, such as Dutch, pitch mainly serves as the post-lexical function.

2.6 The current study

As has become clear by now, plenty of prior studies have investigated the L1 influence on tonal perception in L2 Mandarin. However, the impact of emotions on the perception of Mandarin tones is not well understood. In particular, the questions as to how and to what extent emotions affect the perception of Mandarin tones by non-native tone speakers have not been addressed. To bridge the gap, we examine the perception of Mandarin tones in multisyllabic sentences with four different emotions (i.e. neutrality, happiness, sadness, anger) by both beginning and advanced Dutch learners of Mandarin, compared to native speakers of Mandarin. Among these four emotions, "neutrality" is regarded as no emotion, which acts as the point of comparison for other emotions. The other emotions (e.g. happiness, sadness, anger) have been extensively studied in prior research, and are arguably the basic emotions in daily communication (Darwin, 1872).

3. Research questions & hypotheses

The research questions and hypotheses that this study examines are summarized as follows.

Question 1: Does tonal perception differ in different emotions in L2 Mandarin, compared to L1 Mandarin?

Based on the literature review above, the emotional intonation changes the f0 of lexical tones in emotional speech in terms of pitch range, pitch register and duration, compared to utterances in neutrality, which results in challenges for Dutch speakers.

Hypothesis 1: We hypothesize that emotions will affect the perception of tones in emotional speech by Dutch learners of Mandarin.

Native speakers of Mandarin may perceive tones similarly accurately across emotions, while L2 Mandarin learners may be less accurate in tonal perception in non-neutral emotions, especially those that lead to bigger changes to the surface form of the tones, than in neutral emotion. It is predicted that L2 learners may be least accurate in the angry emotion, compared to the neutral emotion.

Question 2: Does tonal perception in emotional speech differ in different tones in L2 Mandarin, compared to L1 Mandarin?

Previous studies showed that learners of Mandarin with English as their L1 can perceive Tone 1 more accurately than the other tones, followed by Tone 4, Tone 2 and Tone 3 (Yue, 1986; Wang, 1995). For the perception of Mandarin tones, Wang et al (1999) conducted a tonal identification task on beginning English learners of Mandarin. They found that participants showed an identification accuracy rate of 69% with a prevailing Tone 2-Tone 3 confusion, compared to an accuracy of 87% for Tone 1-Tone 4. Besides, Hao (2012) found that English and Cantonese learners of Mandarin had difficulty in distinguishing T2 and T3, followed by distinguishing T1 and T2. The T2-T3 confusion was the major problem for both native English speakers and native Cantonese speakers. The confusion can be the influence from listeners' first language, and probably stems from the acoustic similarity of these two tones, which share similar pitch contours and there is overlap in their pitch ranges (Moore & Jongman, 1997). This confusion has also been found for adult native Mandarin speakers (Zhang, Samuel & Liu, 2012).

Hypothesis 2: We thus hypothesize that tonal perception can vary in different tones in emotional speech and predict that the perception of Tone 2 and 3 will be more difficult than the perception of Tone 1 and Tone 4.

Question 3: Does tonal context influence tonal perception in emotional speech in L2 Mandarin?

Previous studies found that tonal context influences tonal perception (Xu, 1997; Moore & Jongman, 1997). When produced in isolation, different tones are realized with stable and distinctive pitch contours. However, tones can be affected by the neighboring tones and undergo substantial acoustic variations, which leads to changes of f0 realizations. Tone 3 has the most variants among these tones in connected speech, which has a low-falling shape (Shih, 1997: 82). Tone 3 is realized with a rising contour when followed by another Tone 3. It will change to a low falling contour when followed by a different tone (e.g. Tone 1, Tone 2, Tone 4, and neutral tone).

Hypothesis 3: We hypothesize that the tonal context will influence the perception of tones and predict that the perception of Tone 3 will be most sensitive to the tonal context.

Question 4: Does proficiency in Mandarin affect tonal perception in emotional speech in L2 Mandarin?

Studies on the perception of non-native tones found that linguistic experience gained from listeners' native languages facilitates their perception of non-native tones substantially (Burnham & Francis, 1997; Gottfried & Suiter, 1997; Wayland & Guion, 2004). More specifically, native tone language speakers have more experience in using pitch variations (or tones) in their native languages than do speakers of non-tone languages. Some studies suggested that speaking a tonal L1 may make it easier for listeners to perceive L2 tones. Wang et al. (1999) examined the tonal identification

performance by listeners who were beginning learners of Mandarin (six weeks of learning) and intermediate learners of Mandarin (one year of learning). The experiment found that beginners performed worse than intermediate learners in the task. Wayland and Guion (2004) trained native English and Mandarin Chinese speakers to discriminate and categorize a pair of Thai tones. While the Chinese speakers showed apparent improvement after training, in contrast, the English group did not display any improvement in discrimination and showed only very slight progress in categorization. They concluded that speakers of tonal L1s might have an advantage over speakers of non-tone L1s in learning to perceive tones in an L2. Besides, Lee et al. (1996) compared Mandarin, Cantonese, and English speakers' ability to discriminate both Mandarin tone pairs. They found that Mandarin speakers performed better than the Cantonese speakers, and Cantonese speakers performed significantly better than the English speakers. These results suggest that Cantonese speakers' experience with tones facilitates their processing of tonal information.

Hypothesis 4: It is hypothesized that language proficiency would affect the perception of tone.

Participants with higher Mandarin proficiency perform better than those with low proficiency in the tonal perception task. It is predicted that native Mandarin speakers would perform significantly better than Dutch learners of Mandarin, and advanced Dutch learners of Mandarin will perform better than beginning learners of Mandarin.

4. Method

This study was approved by the Ethical Assessment Committee Linguistics (ETCL) of the Utrecht Institute of Linguistics OTS under ETCL reference number 5580242-01-2018.

4.1 Participants

Twenty-four adult Dutch learners of Mandarin at two different proficiency levels (beginners and advanced learners, 12 males, 12 females, age: M = 25.5 , SD =5.3) participated in the experiment. Eight of them were from the Chinese Education Center Netherlands, and the rest of them were students from the Chinese Studies programme in Leiden university. They started to learn Mandarin Chinese at the age of 18 in the Netherlands. None of the them have studied tone languages other than Mandarin prior to this study, and they were not from Limburg, Netherlands. The participants from experimental group have taken the Chinese proficiency test (HSK) at different levels prior to the experiment. They were recruited via the education center and the Leiden university website. They were divided into two groups according to their Mandarin proficiency, and each group had 12 participants. Specifically, participants who passed HSK level 1 & 2 were grouped as beginners, and participants who passed HSK level 3 & 4, and two participants who passed HSK 5 were regarded as advanced learners. In addition, 12 native speakers of Mandarin (mean age: 28,0) participated in the experiment as the controls. They were master and PhD students from Leiden University and Utrecht University, who were not from the Linguistics programme. None of the participants reported having dyslexia or any hearing defects.

4.2 The tonal perception experiment

A tonal perception experiment was carried out in this study, which was designed to investigate the relations between emotions, lexical tones, tonal contexts, language proficiency, and the perception of tones. In the perception task, the participants were asked to seat themselves in front of a computer screen. There were four choices, namely Tone 1, 2, 3, 4. Sound was played over headphones. Participants were asked to choose the correct lexical tone of the target word in each sentence by pressing one of the buttons in the button box (see Appendix C) after they listened to the recordings. The perception experiment was conducted to investigate how well Dutch learners of Mandarin identify lexical tones in different emotions. This experiment was set up

using the software programme ZEP (Veenker, 2017), which is a system for running Psycholinguistic experiments. ZEP delivers auditory and visual stimuli, which also interface with external hardware (i.e. response boxes, eye-trackers). The responses were recorded by ZEP, and the accuracy of the responses were measured at the end of the experiment.

4.3 Stimuli

4.3.1 Experimental materials

The stimuli were 256 short sentences, containing four monosyllabic pseudo words, i.e. two CVC and two CV pseudo words (i.e. *mong, ging, ra, pü*), embedded in four carrier sentences and spoken in each of the four lexical tones and each of the four selected emotions (i.e. "neutrality", "happiness", "anger" and "sadness") (see Appendix A). These pseudo words have a common syllable structure and are legal in terms of phonological rules, but do not exist in Mandarin Chinese (Farinas & Pellegrino, 2001). However, they are pronounceable for native speakers of Mandarin. We chose to use pseudo words as target words because they do not have lexical entries in the mental lexicon and this can in turn minimize familiarity effects in tonal perception.

The target words were embedded in four carrier sentences, adopted from the carrier sentences used in Francis et al. (2008), as shown in example (1). The carrier sentences were semantically neutral and thus could not provide any semantic cues to the tonal perception. Given the fact that lexical tones may change in context due to tone sandhi, we designed four different tones (T1-T4) preceding the target word in each sentence, which were Tone 1 (*chu1*) in carrier sentence (a), Tone 2 (*du2*) in carrier sentence (b), Tone 3 (*xie3*) in carrier sentence (c), Tone 4 (*lian4*) in carrier sentence (d). This allowed us to investigate the effect of tonal context on L2 tonal perception. Additionally, it is generally accepted that f0 changes at the initial and final positions

in a statement. Tones at the initial position have the highest f0 level. In contrast, tones at the final position have “sentence-final tone lowering effects”, which refers to a final fall in the contour of the last syllable (Vance, 1976). In order to avoid the influences of changes in f0 level for tones at the initial and final positions by intonation, I placed the target word in the medial position in each sentence in this study.

(1) a. 指出__这个字。

Pinyin (with tone) zhi3 chu1 __ zhe4 ge4 zi4.

English translation Please point the word __ out.

b. 我会读__这个字。

Pinyin (with tone) wo3 hui4 du2__zhe4 ge4 zi4.

English translation I can read the word __.

c. 我会写__这个字。

Pinyin (with tone) wo3 hui4 xie3__zhe4 ge4 zi4.

English translation I can write the word __.

d. 我想练__这个字。

Pinyin (with tone) wo3 xiang3 lian4__zhe4 ge4 zi4.

English translation I can practice the word __.

In order to minimize boredom, we divided the 256 stimuli into four lists according to a Latin square design. Each list contained 64 stimuli, including 16 representations of

each target word in each emotion. Each list contains four different emotions, four target words, four tonal contexts, and four different lexical tones. Each stimuli list was presented in a randomized order in ZEP according to the following constraints: (1) the same tones should not appear more than two times in a row; (2) the same carrier sentence should not appear more than two times in a row; (3) the same syllable should appear more than two times in a row; (4) the same emotion cannot appear more than two times.

Due to the fact that there were two groups of participants whose native languages were different, we designed two versions of instructions, (i.e. a Dutch version and Chinese version, see Appendix B). The control group and the experimental group received instructions in their native language.

4.3.2 Recording procedure

After the stimuli were designed, a professional female voice actor who was a native Mandarin speaker recorded all the stimuli. The speaker sat in a sound-attenuated booth with printed speech materials at the Linguistic Laboratory of the Utrecht Institute of Linguistics. In order to put more emphasis on the target word, the sentences were recorded in a dialogue setting, in which we designed a question which corresponded to each sentence, as shown in example (2). Each sentence was listed with a number in front of it. The stimuli were printed on A4 sheet of papers with a font size of 28 point, which were clearly readable. Recordings were made with a large membrane microphone at a sampling frequency of 44.1kHz with 16 bits resolution. Two persons assisted in the recording. One was in charge of the technical control. The other was to provide suggestions on how to speak and make the judgement. Recording did not stop until satisfactory stimuli were obtained. The speaker was asked to simulate the emotion based on her own experience and was assisted with some materials such as the photos of her favorite superstars, etc. Different emotions were

recorded separately, with a break in between to make sure that the speaker could adjust herself from one emotion to another. During the recording process, the speaker was encouraged to make additional attempts if she was not satisfied with her first attempt or if the author, who was present during the recording, noticed inconsistency in her emotion. In the case of multiple attempts, the last attempt of recording was selected and saved as an individual .wav file using Praat (Boersma & Weenink, 2014). After the recording, all stimuli were checked by a native Mandarin speaker in a perception test. The listener rated 100% accuracy on the emotions conveyed in each sentence. The high accuracy rate demonstrated that the emotions conveyed in each sentence were clearly recognized. Then all utterances were annotated in Praat. The voiced segment of each sentence was identified visually from an amplitude waveform and a wideband spectrogram display.

(2) a. 指出哪个字给你看？

Pinyin (with tone) zhi3 chu1 na3 ge4 zi4 gei3 ni3 kan4?

English translation Which word do you want me to point out for you?

指出__这个字。

Pinyin (with tone) zhi3 chu1 __ zhe4 ge4 zi4.

English translation Please point the word __ out.

b. 你会读哪个字？

Pinyin (with tone) ni3 hui4 du2 na3 ge4 zi4?

English translation Which word can you read?

我会读__这个字。

Pinyin (with tone) wo3 hui4 du2__zhe4 ge4 zi4.

English translation I can read the word __.

c. 你会写哪个字？

Pinyin (with tone) ni3 hui4 xie3 na3 ge4 zi4?

English translation Which word can you write?

我会写__这个字。

Pinyin (with tone) wo3 hui4 xie3__zhe4 ge4 zi4.

English translation I can write the word __.

d. 你想练哪个字？

Pinyin (with tone) ni3 xiang3 lian4 na3 ge4 zi4?

English translation Which word do you want to practice?

我想练__这个字。

Pinyin (with tone) wo3 xiang3 lian4__zhe4 ge4 zi4.

English translation I can practice the word __.

4.4 Procedure

The participants were tested individually (by the author) in three different places according to the location of the participants, which were the sound-attenuated booth at the Phonetics Laboratory at the University of Utrecht, the library at the University of

Leiden, and the quiet room at the Chinese education center. Prior to the experiment, the participants were asked to read an information letter and sign a participation approval form if they agreed to participate. They received financial compensation of 2 euro for their participation as per the standards of the UiL OTS research lab. They were semi-randomly assigned to a stimulus list so that there was an approximately equal number of participants tested using each list. Each participant was tested individually in the phonetic lab with 64 trials with the auditory stimuli presented through headphones. Before the experiment, each participant was first presented with written instructions according to their native language. In the instructions, the tasks and procedure of the experiment were explained. They were asked to seat themselves in front of a computer screen and put on headphones. They were instructed to listen carefully to the recordings and indicate the tone of the underlined word (for example: 指出 *_hing_* 这个 字.) by pressing one of the four buttons in the bottom box which referred to Tone 1, Tone 2, Tone 3, and Tone 4. Before the real experiment, there was a practice trail which helped the participants get familiar with the experiment.

The timeline of an experimental trial was as follows.

Each trial started with the presentation of a fixation cross ("+"), which lasted for a 1000ms. Then the fixation cross disappeared and the stimulus sentence in Chinese characters and Pinxin, its Dutch translation appeared on the screen, which lasted for 2400ms, as shown in Figure 7. Then a sound icon (🔊) popped out below the target word (i.e. *mong*) for 400ms. After that the sound of the stimulus sentence was played to the participants via a headphone set. At the end of the sound file, the target word, the image of the button box, and the four response boxes appeared, representing the four lexical tones, as shown in Figure 8. The appearance of the response boxes acted as a signal to the participant that he or she could indicate the tone of the target word by pressing the corresponding button of the button box. The buttons had the same arrangement as the response boxes. The experiment lasted on average about 10

minutes. The inter-trial interval was specified at 500ms. The participants' perceptual judgments and response time starting from the end of the sound file were automatically registered in ZEP. Only the perceptual judgments were relevant to this current study.

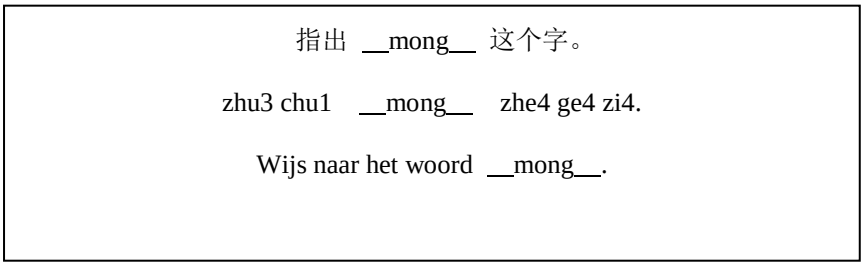

Figure 7. The experimental stimuli

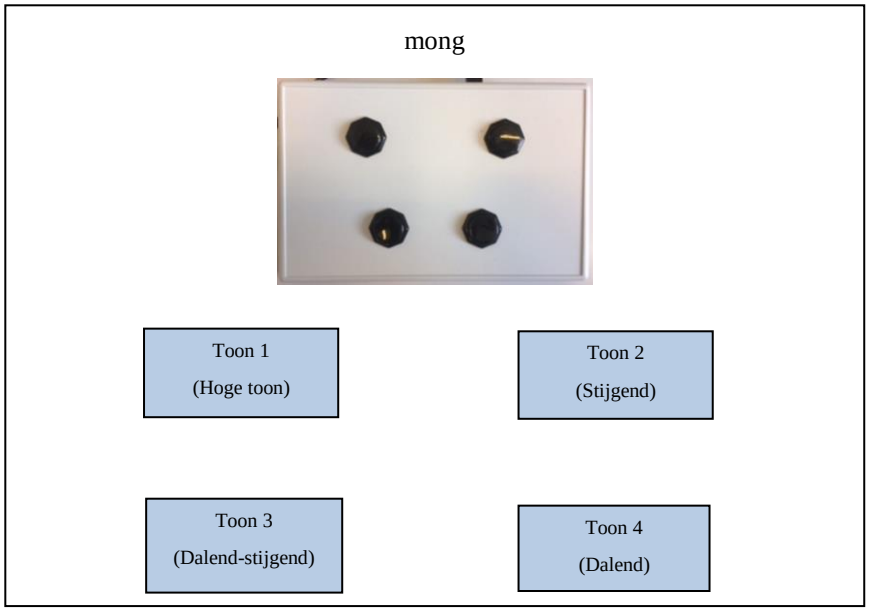

Figure 8. The target word, the button box and the four lexical tones

5. Statistical analysis and results

In the current study, a mixed-effect binary logistic regression in SPSS (IBM SPSS version 25) was used to investigate the influence of emotions, tonal contexts, tones, and proficiency on the perception of tone. The predictors included four factors:

emotions (i.e. neutrality, happiness, sadness, anger), tonal context which refers to the lexical tone of the word preceding the target word (i.e. context 1, context 2, context 3, context 4), tone (the lexical tone of the target word, i.e. Tone 1, Tone 2, Tone 3, Tone 4), and proficiency (proficiency 1: beginning Dutch learners of Mandarin, proficiency 2: advanced Dutch learners of Mandarin, proficiency 3: native Mandarin speakers). In addition, there were four two-way interactions, namely emotion by proficiency, tonal context by proficiency, tone by proficiency, and emotion by tone. Also we included participants and words as the random factors. The outcome variable was the perception of tone (i.e. correct vs. incorrect), with "incorrect" as the reference category.

Our modelling yielded three significant main effects, (i.e. emotion, tone and proficiency), and three significant interactions, (i.e. emotion by tone, tone by proficiency, and tonal context by proficiency). The summary of the results can be seen in Table 1.

Regarding the main effect of emotion with "neutrality" as the reference category, in general the participants were less accurate in tonal perception in the three non-neutral emotions, compared to the neutral condition (happiness: *coefficient* = -0.062, sadness: *coefficient* = -1.130, anger: *coefficient* = -1.815). However, only the difference between anger and neutrality reached statistical significance (*coefficient* = -1.815, *SE* = 0.558, t = -3.254, $p < .05$). Regarding the main effect of tone with Tone 1 as the reference category, the coefficients suggested that the participants performed less well in Tone 2 (*coefficient* = -1.947, *SE* = 0.628, t = -3.100, $p < .01$), Tone 3 (*coefficient* = -1.243, *SE* = 0.633, t = -1.965, $p < .05$), Tone 4 (*coefficient* = -2.038, *SE* = 0.635, t = -3.208, $p < .01$), compared to Tone 1. Participants did worst in Tone 4 and less worst in Tone 2. However, the interaction of emotion by tone revealed that the participants were more accurate in the perception of Tone 4 than the perception of Tone 1 in the

angry emotion, suggesting that tonal perception can differ between tones in the same non-neutral emotion.

With respect to the main effect of proficiency with "proficiency 3" as the reference category, the participants from proficiency 1 (beginning) and proficiency 2 (advanced) did not differ from the native controls from proficiency 3. This is the case across tones and tonal contexts. However, when we ran the analysis again with "proficiency 1" as the reference category, we found that the participants from proficiency 2 (*coefficient* = 2.670, *SE* = 0.996, *t* = 2.681, $p < .01$) were significantly more accurate in tonal perception than the participants from proficiency 1. This advantage of a higher level of L2 proficiency was primarily driven by more accurate perception of Tone 3 (*coefficient* = -1.728, *SE* = 0.823, *t* = -2.099， $p < .05$), compared to the perception of Tone 1 and more accurate perception of tones when preceded by Tone 4 in the carrier sentence (*coefficient* = -1.130, *SE* = 0.509, *t* = -2.222， $p < .05$), as shown by the interaction of tone by proficiency and tonal context by proficiency.

Table 1. Summary of results of the mixed model linear regression analysis. The reference category was "incorrect" for the outcome variable, "neutrality" for the fixed factor emotion, "Tone 1" for the fixed factor tone, "proficiency 3" (native) for the fixed factor proficiency, "tonal context 1" (i.e. the target word was preceded by a Tone-1 word in the carrier sentence) for the fixed factor tonal context.

	Coefficient	Std. Error	t	Sig.	95% Confidence Interval	
					Lower	Upper
Intercept	2.810	0.583	4.819	.000	1.667	3.954
Anger	-1.815	0.558	-3.254	.001	-2.909	-0.721
Tone 4	-2.038	0.635	-3.208	.001	-3.284	-0.792

Tone 3	-1.243	0.633	-1.965	.050	-2.483	-0.002
Tone 2	-1.947	0.628	-3.100	.002	-3.179	-0.716
Proficiency 2	2.670	0.996	2.681	.007	0.717	4.624
Emotion*Tone						
Anger*Tone 4	2.751	0.717	3.839	.000	1.346	4.156
Tone*Proficiency						
Tone 3* Proficiency 2	-1.728	0.823	-2.099	.036	-3.343	-0.113
Tone 2*Proficiency 3	2.454	0.919	2.671	.008	0.652	4.256
Context*Proficiency						
Context 4*Proficiency 2	-1.130	0.509	-2.222	.026	-2.128	-0.133

6. Discussion

This study set out to address four research questions:

Question 1: Does tonal perception differ in different emotions in L2 Mandarin, compared to L1 Mandarin?

Question 2: Does tonal perception in emotional speech differ in different tones in L2 Mandarin, compared to L1 Mandarin?

Question 3: Does tonal context influence tonal perception in emotional speech in L2 Mandarin?

Question 4: Does proficiency in Mandarin affect tonal perception in emotional speech in L2 Mandarin?

Regarding the first research question, we hypothesized that tonal perception can vary in different emotions for L2 learners and may be the least accurate in the angry

emotion, compared to the neutral emotion. Regarding the second research question, we hypothesized that different tones would vary in emotional speech. According to the order of the acquisition difficulty of Mandarin tones, we predicted that the perception of Tones 2 and 3 will be more difficult than the perception of Tone 1 and 4 in both L1 and L2. Our data show that both the L2 Mandarin learners and native controls were significantly more accurate in tonal perception in the neutral emotion than in the angry emotion but not in the happy and sad emotions. Nevertheless, the participants' perception of Tone 4 was more accurate in the anger emotion, compared to the perception of Tone 1, different from what we expected from the tones.

Regarding the effects of tonal context on the perception of Mandarin tones, we hypothesized that the perception of tones in tonal context 3 would be less accurate than that in other tonal contexts, compared to the perception of tones in tonal context 1. Our data show that the effect of tonal context differed for different proficiency levels. The perception of tones was less accurate in tonal context 4 by advanced Dutch learners of Mandarin, compared to the perception of tones in tonal context 1. For the last research question, we hypothesized that native Mandarin speakers would perform significantly better than Dutch learners of Mandarin. However, our data show that the perception of tones did not differ in these three proficiency groups. When ran the analysis again with beginning Dutch learners of Mandarin as the reference category, we found that the advanced learners of Mandarin performed much better than the beginning learners of Mandarin. This language proficiency advantage also made the perception of Tone 3 more accurate by advanced Dutch learners of Mandarin, compared to that of beginning learners.

7. Conclusion

In the current study, we examined the influence of emotions on tonal perception. Additionally, we also investigated the effect of lexical tones, tonal context, and

language proficiency on tonal perception. Different from previous research which mainly focused on the perception of tones in neutral emotion, this study has shed light on the perception of lexical tones in emotional speech.

In line with previous studies suggesting emotions influence tonal perception, especially those that would cause great changes to the f0, the current study found that the perception of tones was less accurate in tones spoken with various emotions, especially with an angry emotion, compared to neutral emotion. This study indicated that the f0 was varied in emotional speech, which resulted in difficulty for listeners, especially for L2 learners. Besides, our study also found that tonal perception was affected by lexical tones. Previous studies found that non-native listeners do not perceive the four Mandarin tones equally well, and Tone 2 and Tone 3 are the most difficult to perceive (Shen, 1989; Gottfried & Suiter, 1997). Contrary to previous findings, the current study found that Tone 4 was the most difficult one to perceive, compared to Tone 1. Regarding the tonal contexts and language proficiency, the current study found that different proficiency groups were affected to different degrees, with advanced learners being influenced the most compared to beginners. For language proficiency, previous studies found that language proficiency affects tonal perception, which was proved in the current study.

In this study, there was only one female speaker who recorded the stimuli. However, there may be gender-related differences in prosodic expression of emotion (Klatt, D., & Klatt, L., 1990). In the previous studies, the tone patterns of “anger” were different in terms of amplitude between the male and female speakers. In addition, previous research suggests that female and male voices are processed differently (Yang et al., 2013). Female voices might impede language processing which is attributed to the high acoustic salience and complexity of female voices (Magnuson & Nusbaum, 2007) . Besides, female voices are different from male voices in terms of acoustic

dimensions including mean pitch, formant frequencies and breathiness. In the future, it might be interesting to examine the relation between the gender effect and the perception of tones.

References

Broselow, E., Hurtig, R. R., & Ringen, C. (1987). The perception of second language prosody. In G. Ioup & S. H. Weinberger (Eds.), Inter-language Phonology, *The Acquisition of Second Language Sound System* (pp. 350-361). Cambridge: Newbury House Publishers.

Blicher, D. L., Diehl, R. L., & Cohen, L. B. (1990). Effects of syllable duration on the perception of the mandarin tone 2/ tone 3 distinction: Evidence of auditory enhancement. *Journal of Phonetics*, 18(1), 37–49.

Birch, S., & Clifton, C. (2002). Effects of varying focus and accenting of adjuncts on the comprehension of utterances. *Journal of Memory and Language*, 47, 571–588.

Braun, B., & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide attention to pitch. *Journal of Phonetics*, 39, 585–594.

Boersma, P. & Weenink, D. (2014). Praat: Doing phonetics by computer [computer program]. Version 5.3.62. http://www.praat.org/

Burnham, D., Kasisopa, B., Reid, A., Luksaneeyanawin, S., Lacerda, F., & Attina, V. (2014). Universality and language-specific experience in the perception of lexical tone and pitch. *Applied Psycholinguistics*, 36(6), 1-33.

Chao, Y. R. (1933). Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology*, 4, 121- 134.

Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: California University Press.

Chuang, C. K., Hiki, S., Sone, T., & Nimura, T. (1972). The acoustical features and perceptual cues of the four tones of Standard Colloquial Chinese. *Proceedings of the Seventh International Congress on Acoustics* (Adadémial Kiado, Budapest), 297-300.

Chen, A., Gussenhoven, C., & Rietveld, A. (2004). Language-specificity in perception of paralinguistic intonational meaning. *Language and Speech*, 47(4), 311-350.

Chen, A. (2005). *Universal and language-specific perception of paralinguistic intonational meaning*. Utrecht: LOT.

Content, A., & Noémi, P. (2011). Categorical perception of tones in Vietnamese. *Paper presented at the16th International Congress of Phonetic Sciences*, August 6–10, 2007. Saarbrücken: Saarland University.

Cabrera, L., Tsao, F. M., Gnansia, D., Bertoncini, J., & Lorenzi, C. (2014). The role of spectro-temporal fine structure cues in lexical-tone discrimination for French and mandarin listeners. *The Journal of the Acoustical Society of America*, 136(2), 877– 882.

Chen, A., Chen, A., Kager, R., & Wong, P. (2014). Rises and falls in Dutch and Mandarin Chinese. *International Symposium on Tonal Aspects of Languages*, 83-86.

Darwin, C. (1872). *The expression of emotion in man and animals*. New York: Oxford University Press.

Duanmu, S. (2007). *The Phonology of Standard Chinese.* Oxford: Oxford University Press.

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion*, 6, 169– 200.

Ellgring, H., & Scherer, K. R. (1996). Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior*, 20, 83–110.

Farinas, J., & Pellegrino, F. (2001). Automatic rhythm modeling for language identification. Eurospeech 2001 Scandinavia, *European Conference on Speech Communication and Technology, INTERSPEECH Event, Aalborg, Denmark, September* (pp.2539-2542). DBLP.

Francis, A. L., Ciocca, V. C., & Ng, B. K. C. (2003). On the (non)categorical perception of lexical tones. *Perception & Psychophysics*, 65, 1029–1044.

Francis, A. L., Ciocca, V., Ma, L., & Fenn, K. (2008). Perceptual learning of Cantonese lexical tones by tone and non-tone language speakers. *Journal of Phonetics, 36*(2), 268-294.

Gandour, J. T. (1983). Tone perception in far Eastern languages. *Journal of Phonetics*, 11, 149–175.

Gussenhoven, C. (1993). The Dutch foot and the chanted call. *Journal of Linguistics*, 29, 37–63.

Gottfried, T. L., & Suiter, T. L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics*, 25, 207-231.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones.* Cambridge: Cambridge University Press.

Halle, P., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics*, 32(3), 395–421.

Hao, Y. C. (2012). Second language acquisition of Mandarin Chinese tones by tonal and non-tonal language speakers. *Journal of Phonetics*, 40(2), 269–279.

Klatt, D., & Klatt, L. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, *87*(2), 820–857.

Krishnan, A., Xu, Y., Gandour, J., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research, 25*(1), 161-168.

Lin, M. C. (1965). The pitch indicator and the pitch characteristics of tones in Standard Chinese. *Acta Acoustica (China)*, 2, 8-15.

Li, N.C. & Thompson, S. A. (1981). *Mandarin Chinese*. Berkeley and Los Angeles: University of California Press.

Leather, J. (1987). F0 pattern inference in the perceptual acquisition of second language tone. In A. James & J. Leather (Eds.), *Sound Patterns in Second Language Acquisition* (pp. 59-81). Dordrecht: Foris Publications,.

Leather, J. (1990). Perceptual and productive learning of Chinese lexical tone by Dutch and English speakers. In J. Leather & A. James (Eds.), *New Sounds 90: Proceedings of the 1990 Amsterdam symposium on the acquisition of second language speech* (pp.72-97). University of Amsterdam.

Ladd, D. R. (1992). An introduction to intonational phonology. In G. J. Docherty & D. R. Ladd (Eds.), *Papers in laboratory phonology II: Gesture, segment prosody* (pp. 321–334). Cambridge: Cambridge University Press.

Ladd, D. R. (1996). *Intonational Phonology*. Cambridge: Cambridge University Press.

Lee, Y. S., Vakoch, D. A., & Wurm, L. H. (1996). Tone perception in Cantonese and Mandarin: A cross-linguistic comparison. *Journal of Psycholinguistic Research*, 25(5), 527-542.

Liu, S., & Samuel, A. G. (2004). Perception of mandarin lexical tones when f0 information is neutralized. *Language & Speech*, *47*(2), 109-138.

Li, A., Fang, Q., & Dang, J. (2011). Emotional tone in a tone language: experimental evidence from Chinese. *Icphs XVII*, 17-21.

Murray, I. R. & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *Journal of the Acoustical Society of America*, 93(2), 1097-1108.

Mozziconacci, S. (1995). Pitch variations and emotions in speech. *Proceedings of the XIIIth International Congress of Phonetic Sciences*, Stockholm, 1, 178– 181.

Moore, C. B., & Jongman, A. (1997). Speaker normalization in the perception of Mandarin Chinese tones. *Journal of the Acoustical Society of America*, 102, 1864 –1877.

Nordenhake, M., & Svantesson, J. O. (1983). Duration of Standard Chinese word tones in different sentence environments. *Lund University Working papers in Linguistics*, 25, 105-111.

Pierrehumbert, J. B., & Steele, S. A. (1989). Categories of tonal alignment in English. *Phonetica*, 46, 181 – 196.

Shen, X. S. (1989). Toward a register approach in teaching Mandarin tones. *Journal of Chinese Language Teachers Association*, 24, 27-47.

Shen, X. S., & Lin, M. C. (1991). A perceptual study of Mandarin tones 2 and 3. *Language and Speech*, 34, 145–156.

Shih, C. (1997). Mandarin Third Tone sandhi and prosodic structure. In J. Wang, & N. Smith (Eds.), *Studies in Chinese phonology* (pp. 81–123). Mouton de Gruyter.

't Hart, J., Collier, R., & Cohen, A. (1990). *A perceptual study of intonation*. Cambridge University Press, UK.

Vance, T. J. (1976). An Experimental investigation of tone and intonation in Cantonese. *Phonetica*, 33, 368-392.

Van Heuven, V. J., & Haan, J. (2002). Temporal development of interrogativity cues in Dutch. In C. Gussenhoven & N. Warner (Eds.), *Papers in Laboratory Phonology 7* (pp. 61-86). Berlin: Mouton de Gruyter.

van Heuven, V. J. & Kirsner, R. S. (2004). Phonetic or phonological contrasts in Dutch boundary tones? In L. Cornips and J. Doetjes (Eds.), *Linguistics in the Netherlands*, 102-113.

Veenker, T. J. G. (2017). The Zep Experiment Control Application (Version 1.10). Beexy Behavioral Experiment Software. Retrieved from http://www.beexy.org/zep/

White, C. M. (1981). Tonal pronunciation errors and interference from English intonation. *Journal of the Chinese Language Association Teachers*, 16 (2), 27-56.

Wang, Y. (1995). Ye tan meiguoren xuexi hanyu shengdiao. *Yuyan Jaoxue Yu Yanjiu*, 2, 126-140.

Wang, Y., Spence, M. M., Jongman, A., & Sereno, J. A. (1999). Training American listeners to perceive Mandarin tones. *Journal of the Acoustical Society of America*, 106(6), 3649-3658.

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *Journal of the Acoustical Society of America*, 113, 1033-1044.

Welby, P. (2003). Effects of pitch accent position, type, and status on focus projection. *Language and Speech*, 46(1), 53–81.

Wayland, R. P., & Guion, S. G. (2004). Training native English and native Chinese speakers to perceive Thai tones. *Language Learning*, 54, 681-712.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25(1), 61–83.

Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, 33, 319-337.

Xu, Y, & Xu, C. X. (2005). Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 33(2), 159–197.

Xu, Y., Gandour, J. T., & Francis, A. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *Journal of the Acoustical Society of America*, 120(2), 1063–1074.

Xu, Y., Kelly, A., & Smillie, C. (2013). Emotional expressions as communicative signals. In S. Hancil & D. Hirst (Eds.), *Prosody and Iconicity* (pp. 33–60). Amsterdam: John Benjamins Publishing Company.

Yue-Hashimoto, A. O. (1986). Shengdiao jiaofa de Shangque. Diyijie Guoji Hanyu Jiaoxue *Taolunhui Lunwenxuan*, 229-235.

Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.

Yang, H., Yang, S., & Park, G. (2013). Her voice lingers on and her memory is strategic: effects of gender on directed forgetting. *Plos One*, 8(5), e64030.

Zhang, X., Samuel, A. G., & Liu, S. (2012). The perception and representation of segmental and prosodic Mandarin contrasts in native speakers of Cantonese. *Journal of Memory and Language*, 66, 438-457.

Appendices

Appendix A

Table 2. Experimental Stimuli

Order	Sentences	Pinyin	Dutch Translation
1	指出_mong_这个字。	zhi3 chu1 _mong_ zhe4 ge4 zi4.	Wijs naar het woord _mong_.
2	我会读_mong_这个字。	wo3 hui4 du2 _mong_ zhe4 ge4 zi4.	Ik kan het woord _mong_ lezen.
3	我会写_mong_这个字。	wo3 hui4 xie3 _mong_ zhe4 ge4 zi4.	Ik kan het woord _mong_ schrijven.
4	我想练_mong_这个字。	wo3 xiang3 lian4 _mong_ zhe4 ge4 zi4.	Ik wil het woord _mong_ oefenen.
5	指出_ging_这个字。	zhi3 chu1 _ging_ zhe4 ge4 zi4.	Wijs naar het woord _ging_.
6	我会读_ging_这个字。	wo3 hui4 du2 _ging_ zhe4 ge4 zi4.	Ik kan het woord _ging_ lezen.
7	我会写_ging_这个字。	wo3 hui4 xie3 _ging_ zhe4 ge4 zi4.	Ik kan het woord _ging_ schrijven.
8	我想练_ging_这个字。	wo3 xiang3 lian4 _ging_ zhe4 ge4 zi4.	Ik wil het woord _ging_ oefenen.
9	指出_ra_这个字。	zhi3 chu1 _ra_ zhe4 ge4 zi4.	Wijs naar het woord _ra_.
10	我会读_ra_这个字。	wo3 hui4 du2 _ra_ zhe4 ge4 zi4.	Ik kan het woord _ra_ lezen.
11	我会写_ra_这个字。	wo3 hui4 xie3 _ra_ zhe4 ge4 zi4.	Ik kan het woord _ra_ schrijven.
12	我想练_ra_这个字。	wo3 xiang3 lian4 _ra_ zhe4 ge4 zi4.	Ik wil het woord _ra_ oefenen.
13	指出_bü_这个字。	zhi3 chu1 _bü_ zhe4 ge4 zi4.	Wijs naar het woord _bü_.
14	我会读_bü_这个字。	wo3 hui4 du2 _bü_ zhe4 ge4 zi4.	Ik kan het woord _bü_ lezen.
15	我会写_bü_这个字。	wo3 hui4 xie3 _bü_ zhe4 ge4 zi4.	Ik kan het woord _bü_ schrijven.
16	我想练_bü_这个字。	wo3 xiang3 lian4 _bü_ zhe4 ge4 zi4.	Ik wil het woord _bü_ oefenen.

Appendix B

Introduction (Dutch version)

Bedankt voor je deelname aan deze studie.

In dit experiment zal je een aantal Chinese zinnen te horen krijgen.Het is mogelijk dat je dezelfde zin te horen krijgt met verschillende emoties. Jouw taak is om aandachtig te luisteren naar de opnames. Na elke zin zal je de toon van het onderstreepte woord (Bijvoorbeeld:指出_hing_这个字。) moeten aangeven door op een van de vier knoppen de drukken. Zoals je weet heeft het Chinees vier tonen: toon 1 (hoge toon), toon 2 (stijgend), toon 3 (dalend-stijgend), toon 4 (dalend). Denk b.v. de woorden 'ma-1' (moeder), 'ma-2' (hennep), 'ma-3' (paard) en 'ma-4' (schelden).

Gebruik de koptelefoon om naar de opnames te luisteren. Je zal elke opname maar één keer horen.We beginnen met een oefenronde. Druk op de ´Ga door´ knop wanneer je klaar bent om te beginnen.

Introduction (Chinese version)

指引

感谢您参加这项实验。

在这个实验中您会听到一些中文句子，有时同一个中文句子会用不同的感情表达。请认真聆听，在每个句子结束时，您必须通过按下按钮来选出带下划线的汉字的声调（例如：_____）。

请使用耳机收听录音，每次您只会听到一次录音。

在实验开始前，会有一个简短的练习部分。

当您准备好时请按下任意键。

English Translation

Thank you for participating in this study.

In this experiment you will hear some Chinese sentences. It is possible that you hear the same sentence with different emotions. Your job is to listen carefully to the recordings. After each sentence you will have to indicate the tone of the underlined word (For example: 指出 _hing_ 这个 字.) By pressing one of the four buttons. As you know, the Chinese has four tones: tone 1 (high tone), tone 2 (rising), tone 3 (falling-rising), tone 4 (falling). Think e.g. the words 'ma-1' (mother), 'ma-2' (hemp), 'ma-3' (horse) and 'ma-4' (scolding).

Use the headphones to listen to the recordings. You will hear every shot only once. We start with a practice round. Press the 'Go through' button when you are ready to start.

Appendix C The picture of button box

Appendix D Questionnaire

Vragenlijst

Testdatum: Testplaats: Proefpersoonnummer:

Naam: Geboortedatum:

Geslacht: Leeftijd:

Geboorteplaats: Studerend / werkend

1. Op welke leeftijd ben je Mandarijn Chinees gaan leren?

2. Wordt er bij jou thuis Mandarijn Chinees gesproken?

3. Is Nederlands jouw moedertaal?

4. Wordt er thuis een Limburgs dialect gesproken? Zo ja, welke?

Bedankt voor het invullen.