Effect of language context in non-native perception of intonation

Insights from Dutch listeners' perception of Mandarin

Eric Shek

Thesis submitted in partial
fulfillment for the degree of
Master of Arts (Linguistics)

Leiden University Centre for Linguistics
Faculty of Humanities
Leiden University

June 2018

Supervisor: Dr. Yiya Chen
Second reader: Dr. Willemijn Heeren

Universiteit Leiden
The Netherlands

For my parents

Contents

Acknowledgements

I came to Leiden with the original intent of documenting endangered languages. It was almost as an afterthought that I went to check out a course on the sounds of language (Experimental Phonetics). But the lecturer, Yiya Chen, drew me into the world of phonetics with such contagious enthusiasm that I didn't leave. Completing this thesis under her supervision has been an enriching experience. She inspired me to stretch my intellectual boundaries, with her vision that we were part of a broader research community striving not merely to affirm, but also to extend our knowledge on the workings of the world.

I wouldn't have been equipped with the technical skills to explore the world of phonetics if not for Jos Pacilly. He taught me all the ins and outs of various speech processing software, and freely gave up his time to assist me with any issues I had with the software, no matter how trivial.

Running a speech experiment like I did for this thesis requires the assistance of many individuals. I'm grateful for the goodwill of many people who enabled my experiment to come alive. Ans de Rooij has been an enthusiastic source of advice about my experiment ever since I introduced it to her. Amongst other things, she helped me design my experiment so that it would be accessible to students of Mandarin. Min Liu's help has also been instrumental in realizing my experiment. She generously shared her knowledge from her doctoral research, and kindly provided the E-Prime program file that formed the basis for my own experiment.

Naomi Nota, Yara Sleebos and Marijn van 't Veer offered valuable feedback on my initial experiment proposal. My discussions with Yuan Zhang enabled me to refine details on my experiment design. Jiang Wu and Siyun Wu reviewed the experiment stimuli and reviewed my Chinese language advertising flyer. Wei-Wei Lee reviewed my Dutch language advertising flyer. Ying-ting Wang, Zhaole Yang and Zhujun Liu, in their capacities as Mandarin teachers, helped me get in touch with participants for my experiment. Indeed, I thank the participants for their interest and enthusiasm in my experiment. Special acknowledgements go to Yuping Liu and Mengying Che, who recorded the stimuli, as well as Hang Cheng, Nick Lowe, Joren Pronk, Ans de Rooij, Siyun Wu and Arthur Zhang, who provided feedback on my pilot experiments. Thank you to Robert Frolla and Naomi Nota, who lent their sharp proofing skills in an advanced version of my thesis.

It wasn't just the world of phonetics that was new for me. The Netherlands was also new. A *hartelijk bedankt* to Wei-Wei Lee, Naomi Nota and Yara Sleebos who made me feel at home with their companionship and our frivolous times together eating, drinking, bouldering, exploring rustic Dutch villages, dancing the night away – you will be missed.

Although life has taken me to worlds away from home for the past decade, my parents have always been there for me. Mum and Dad, thank you for unconditionally supporting me emotionally and financially throughout my time in the Netherlands. The least I could do to honor your support is to do my best in all that I do. This thesis is dedicated to you.

Eric Shek
Leiden, June 2018

Abstract

Previous studies have indicated that native and non-native listeners' attention to differences in segments and lexical tones is heightened when language context is removed. Do they also display greater sensitivity to intonational differences in the absence of language context? To examine this question, this thesis tests the ability of Dutch and Mandarin listeners to identify Mandarin questions and statements that differ only in intonation in three different levels of language context: no language context, a neutral language context, and a constraining language context. All listeners were found to identify questions and statements better with each increasing level of language context. This suggests that the presence of a meaningful semantic context facilitates the perception of intonational meaning. Moreover, Mandarin listeners were better at identifying questions and statements than non-native listeners in sentences with language context. But the difference between Mandarin and Dutch listeners' abilities was minimal in sentences without language context. This result suggests that the effect of language experience on intonation perception is diminished at the lower auditory processing level.

List of Figures

Introduction

The effect of language experience on the perception of segments and lexical tone has been well documented. We are more attuned to variations in segments and tones[1] that are linguistically meaningful in our native language than those that are not. Nevertheless, a more nuanced picture of speech perception emerges when we consider the effect of language context. This effect has been demonstrated by a body of research that has tapped into the analogous effect of speech processing levels. We process speech at a low auditory level when we perceive speech without language context, and conversely, we process speech at a higher linguistic level when we perceive speech with language context. The body of research demonstrates that the effect of language experience is weaker at lower levels of processing segments and tones (e.g., Miyawaki et al., 1975; X. Luo & Ashmore, 2014). That is, the ability to perceive variations in segments and tones without a language context is comparable amongst native and non-native listeners. Moreover, it has also emerged from this body of research that both native and non-native listeners have a better ability to perceive segmental and tonal variation at the auditory processing level than at the higher linguistic processing level. This is explained by the idea that the absence of language context enables listeners to direct their attention to acoustic variations in speech.

However, there has been less research on the effects of language experience and the degree of language context in the perception of intonation (c.f. e.g., Ortega-Llebaria & Colantoni, 2013). Thus we have little insight as to whether listeners are more sensitive to intonation in the same way we are to segments and tones where there is no language context, compared with where there is language context. To illuminate this issue, we test the ability of Dutch listeners to identify intonational cues of statements and questions in Mandarin, in three types of sentences that increase in the degree of language context: filtered sentences with no language context, sentences with a neutral context, and sentences with a constraining context. This experiment

[1]The use of the word "tone" throughout this thesis will refer to lexical tone. It is important to note that our conception of tone here does not refer to its other use in the field of intonation: pitch units in the course of a phonological phrase. This latter conception of tone is explicated in works such as Gussenhoven (2004, Chapter 7).

thus advances our knowledge on how native and non-native listeners' perception of intonation varies according to the degree of language context.

The structure of the current study is as follows. Chapter 1 provides a review of the research on the effect of language context on speech perception. The chapter firstly presents a more detailed examination of the findings that listeners' perception of segments and tones is stronger at the low auditory processing level than at the higher linguistic level. Secondly, the chapter outlines our existing knowledge on the perception of intonation, with particular emphasis on the findings that have emerged from the limited research conducted on the effect of the degree of language context on non-native perception of intonation. The chapter then demonstrates how our understanding in this area can be extended through an experiment that examines Dutch listeners' perception of Mandarin questions and statements with varying degrees of context. Chapter 2 will provide a detailed explanation of the methodology of the experiment. Chapter 3 presents the results, and Chapter 4 discusses what these results reveal about the effects of language context on intonation perception.

Chapter 1

Literature review

The motivation for the current study is the dearth in our understanding on the effect of language context on the perception of intonation. This chapter firstly reviews the considerable number of studies on the effect of processing levels – analogous to the effect of language context – on the perception of segments and tone. These studies generally show that our perception of segments and tone is more refined where there is no language context. The second part of the chapter provides an overview on existing research on the perception of intonation. This research has found that intonation is perceived better in sentences without language context compared to sentences with language context. However, it remains to be seen how the ability to perceive intonation varies across three levels of language context: none, neutral and constraining. To investigate this issue, we set out our main research question in the final part of the chapter: how does the ability of non-native listeners to perceive intonation vary according to different levels of language context?

1.1 Effect of language context on the perception of segments and tone

1.1.1 Perception of segments

Experience with our native language has shaped our perception of speech from infancy (Werker & Tees, 1984). Extensive research has affirmed that this language experience influences the perception of segments, such that we are attuned to segmental contrasts in our native language, but have greater difficulty in distinguishing non-native contrasts (see e.g., Pisoni, Lively, & Logan, 1994; Best & Tyler, 2007). However, a body of evidence has accumulated to support the view that the perception of segments in the absence of language context is not affected by language experience. Studies have found that non-native and native listeners alike are equally attentive to segmental contrasts

without a language context, where listeners engage their auditory processing mechanisms (e.g., Miyawaki et al., 1975; Werker & Logan, 1985). Miyawaki et al. (1975)'s landmark study, for example, found that native Japanese listeners had difficulty in discriminating between the English syllables /ra/ and /la/, which involve a non-native contrast between /r/ and /l/ for Japanese listeners. Conversely, the Japanese listeners displayed a high degree of accuracy which was identical to that of English listeners in distinguishing between these syllables with an isolated third formant – where the acoustic difference between /r/ and /l/ lies. This evidence suggests that at the level of auditory processing, not only is the effect of language experience absent[1], but the ability to discriminate segmental contrasts is also heightened compared to the linguistic processing level.

1.1.2 Perception of tone

The effect of language experience has also been shown to influence the perception of tone – the use of pitch to signal lexical meaning (Yip, 2002). Native listeners of tonal languages have been shown to be generally better at distinguishing tones in their language than non-native listeners (Gottfried & Suiter, 1997; Lee, Tao, & Bond, 2009; c.f. Huang & Johnson, 2010). Several studies have investigated whether the influence of language background extends to the perception of tones at the auditory processing level. On the one hand, some studies have found no influence of language background in the perception of non-speech-like tones (Burns & Sampat, 1980; Burnham et al., 1996; Qin & Mok, 2012). For example, Burnham et al. (1996) found that native English listeners' ability to discriminate Thai tones were not as good as native Thai listeners. However, when these tones were presented as low-pass filtered or music, English listeners' ability to discriminate these tones improved to a commensurate level with native listeners.

On the other hand, an effect of language background in the perception of non-speech tones has been found in other studies (e.g., Y. Xu, Gandour, & Francis, 2006; X. Luo & Ashmore, 2014). X. Luo and Ashmore (2014), for example, exposed Mandarin and English listeners to stimuli based on a continuum of speech and non-speech tones ranging from rising to level. English listeners identified more tones in the continuum as rising than level, when compared to Mandarin listeners. This result was explained on the basis that Mandarin listeners' representation of the high-level tone includes fluctuations in fundamental frequency.

Nevertheless, a consensus from these studies is that listeners, regardless of their language background, show greater sensitivity to pitch differences in non-speech stimuli than speech stimuli (Y. Xu et al., 2006; X. Luo & Ashmore, 2014). Both the Mandarin

[1]This is notwithstanding evidence, as outlined in Sebastián-Gallés (2005, p.549), suggesting that language experience effects listeners' neural activity in the perception of speech without language context.

and English listeners in X. Luo and Ashmore (2014) could identify marginally rising tones as rising when listening to non-speech, while their ability to do so when listening to speech was reduced. A general enhanced ability to perceive pitch in less complex stimuli can explain this. Speech stimuli contain both high-order unresolved harmonics and low-order resolved harmonics, while non-speech stimuli usually only contain the later resolved harmonics. Resolved harmonics have been shown to contribute more to pitch perception than unresolved ones (e.g., Shackleton & Carlyon, 1994, cited in X. Luo & Ashmore, 2014, p.3591). Thus, to a large extent, our ability to perceive tones is enhanced in non-speech stimuli which lack language context, in the same way that the absence of language context enhances the perception of segments.

1.2 Effect of language context on the perception of intonation

Our understanding of the effects of language experience and language context on the perception of intonation is not as refined. Intonation refers to variations in the pitch contour that signal a sentence-level linguistic meaning other than a lexical meaning. This type of meaning is often referred to as "post-lexical" or "intonational meaning" (Wennerstrom, 2001; Braun & Johnson, 2011). Such intonational meanings include not only questions (e.g., Bolinger, 1978; Haan, 2002), but also the expression of focus to give prominence to certain elements in an utterance (e.g., Eady & Cooper, 1986) and the resolution of syntactical ambiguity (e.g., Carlson, 2009).

This section firstly explains how language-specific and universal factors influence the perception of intonation (e.g., Hadding-Koch & Studdert-Kennedy, 1964; Gussenhoven & Chen, 2000; Makarova, 2001). We then review the extant studies examining the effect of language context on the perception of intonation. These studies indicate that intonation is better perceived where there is no language context. However, there are no studies examining how the perception of intonation varies with three levels of language context: none, neutral and constraining. We also examine the effect of lexical tone on the perception of intonation in speech without language context (or a "non-speech context"). Finally, we raise the issue of whether the effect of language background extends to a non-speech context.

1.2.1 Universal and language-specific effects in the perception of intonation

Most studies on the perception of intonation have examined whether it is shaped by language experience or universal factors. These studies have been motivated by an observation that across the world's languages there is a consistent correlation between

the use of pitch and emotion, sometimes known as "paralinguistic meaning": high pitch conveys a heightened sense of emotion (A. Chen, 2005, p.2; Gussenhoven, 2004, p.51). This correlation is also evident between pitch and linguistic meaning, notably the signalling of the distinction between questions and statements, which we describe as the signalling of "interrogativity" in this thesis. Several surveys have found that in the majority of the world's languages, questions are signalled by some high-pitch element (Hermann, 1942; Bolinger, 1978). An inclining pitch on the final syllable, for example, signals questions in languages such as Dutch and English (e.g., Haan, 2002, p.39).

This use of intonation can be explained by the Frequency Code. Under this Code coined by Ohala (1983), higher pitch expresses meanings of deference or submission, while lower pitch expresses meanings of confidence or domination. Gussenhoven (2002) expanded on Ohala's account of Frequency Code to explain why questions are frequently expressed throughout many languages with some high-pitched element. In brief, questions can be regarded as a form of submission, appealing to the goodwill of another person to supply information.[2]

Several studies have demonstrated that the Frequency Code also extends to speech perception, showing that listeners can perceive intonational cues for interrogativity across different languages, regardless of language background (Hadding-Koch & Studdert-Kennedy, 1964; Gussenhoven & Chen, 2000; Makarova, 2001). For example, Gussenhoven and Chen (2000) exposed Hungarian, Mandarin and Dutch listeners to trisyllabic nonsense words which were synthesized for differing degrees of pitch accent height, pitch accent alignment and terminal incline. Although these listeners' languages use different markers of interrogativity, the study found that all listener groups associated higher and later pitch accents, and greater terminal incline with questions.

Nevertheless, an effect of language experience has also been found in the perception of intonation. While listeners can distinguish questions based on cues not used in their native language, they have been found to be more sensitive to cues of questions used in their own native languages. Gussenhoven and Chen (2000) found that the Hungarian listeners were the most sensitive to differences in pitch accent height and alignment, as the higher and later pitch accents that are used to mark questions in Hungarian are not salient markers in Mandarin or Dutch. Mandarin listeners were found to be the least sensitive, as Mandarin marks questions with a higher pitch register rather than changes in tonal contours (e.g., Yuan, 2006).

[2]Note that some languages use low pitch to express questions and a relatively higher pitch to express statements (Gussenhoven, 2002). This is notably the case for several African language families (Rialland, 2009).

1.2.2 Effect of language context on the ability to perceive intonation

Mixed results have emerged from the few studies that have examined the effect of language context on intonation perception (Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014; M. Liu, Chen, & Schiller, 2016a). On the one hand, the results of Ortega-Llebaria and Colantoni (2013) and B. R. Xu and Mok (2012a, 2012b, 2014) suggest that non-native perception of intonation is enhanced in the absence of language context, compared to its presence. On the other hand, M. Liu et al. (2016a)'s finding that a constraining language context promotes intonation perception over a neutral one raises the possibility that the ability to perceive intonation increases with increasing levels of context, from none to constraining. To explore this tension, we examine the methodology and results of these studies in detail in this section.

The aim of Ortega-Llebaria and Colantoni (2013) was to examine if higher levels of processing increased the effects of language experience in the non-native perception and production of focus in English. To examine perception at the linguistic level, English, Mandarin and Spanish listeners listened to a story in English. They were then presented with questions relating to the story. For each question, listeners listened to three possible answers differing only in focus intonation and selected the most appropriate one to the question. This was designated the "access to meaning (+AM)" condition. To examine perception at the lower auditory levels, the same listeners were exposed to utterances spoken with different focus intonation positions. After each of these unfiltered utterances were played, listeners were then exposed to three non-speech sentences differing in focus intonation, generated by removing higher frequencies in normal sentences ("low-pass filtered sentences"). Listeners were then asked to select the one that matched the original utterance. This was designated the "no access to meaning (−AM)" condition.

It was found that all listeners, including the non-native ones, were able to perceive utterances with the correct focus more frequently in the −AM condition than the +AM condition. The authors concluded that increased access to meaning led to a greater effect of language experience in non-native listeners. This aligns with the explanation for the enhanced ability to detect tones in non-speech (analogous to the −AM condition) compared to speech contexts (analogous to the +AM condition).

In the other studies examining the effect of context in the non-native perception of intonation, B. R. Xu and Mok conducted a series of experiments (B. R. Xu & Mok, 2012a, 2012b, 2014) to investigate the perception of Mandarin interrogativity in speech and non-speech contexts by Mandarin and Cantonese listeners. Before discussing their experiments, it is pertinent to provide some background on the nature of Mandarin as a tonal language. Mandarin has four lexical tones: a high-level tone (Tone1), a mid-rising tone (Tone2), a low-dipping tone (Tone3) and a high-falling tone (Tone4) (Y.-H. Lin,

2007, p.4). While these tones are realized at the syllabic level, the intonational cue of interrogativity is realized at the sentence level. The global pitch contour has been shown to be the marker of interrogativity in Mandarin, with a higher pitch contour throughout the sentence expressing a question (e.g., Shen, 1990, p.38; Yuan, Shih, & Kochanski, 2002; F. Liu & Xu, 2005; Yuan, 2006). Regardless of the interrogativity expressed, the canonical shape of the tone in the final syllable generally remains unchanged (e.g., Shen, 1990, p.129; M. Lin, 2004; c.f. Yuan & Shih, 2004, p.4).

In B. R. Xu and Mok (2012a), Mandarin questions and statements were presented to Mandarin and Cantonese listeners. All utterances also varied in their final tone. Listeners were to identify whether each of the utterances was a statement or question. B. R. Xu and Mok (2012b) repeated this experiment with low-pass filtered sentences. B. R. Xu and Mok (2014) conducted the same experiments in what appears to be a replication of B. R. Xu and Mok (2012a, 2012b). The results of the Cantonese listeners in the first experiments under B. R. Xu and Mok (2012a, 2012b) are displayed in Figure 1.1. It shows their accuracy in the perception of Mandarin interrogativity in unfiltered and filtered sentences across all four final tones. From this figure, it can be seen that that the rate of perception accuracy of statements is at a ceiling level across all tones in both unfiltered and filtered speech, except Tone1 statements which are perceived worse in the filtered condition. Notably, the graphs show that perception accuracy of questions is markedly higher across all tones in filtered speech as compared to unfiltered speech.

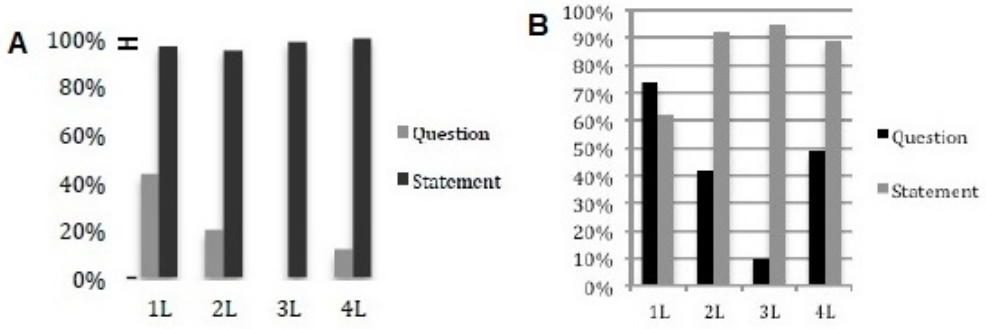

FIGURE 1.1: Cantonese listeners' perception accuracy of Mandarin interrogativity in B. R. Xu and Mok (2012a, 2012b). (A) shows their perception accuracy in unfiltered sentences (taken from B. R. Xu and Mok (2012a), Figure 6). (B) shows their perception accuracy in filtered sentences (B. R. Xu and Mok (2012b), Figure 6).

This trend's consistency is arguably grounds for a broader generalization that Cantonese listeners' ability to perceive differences in the global pitch contour – the salient cue of Mandarin interrogativity – is better in the absence of language context, compared to its presence. More importantly, the graphical data of this trend in Figure 1.1 is evidence that the authors' own claim – that intonation is better perceived in filtered speech compared to unfiltered speech (B. R. Xu & Mok, 2012b, p.5) – is not supported by their results.

The results from Ortega-Llebaria and Colantoni (2013) and B. R. Xu and Mok (2012a, 2012b) as a whole suggest that filtered speech promotes non-native listeners' ability to perceive differences in intonation. However, there is a need to be cautious about the strength of this evidence. Firstly, in B. R. Xu and Mok (2012a, 2012b), no statistical analysis was provided for the difference in the results between filtered and unfiltered speech. Moreover, the replicating experiment of B. R. Xu and Mok (2014) reveals a different trend in Cantonese listeners' perception of Mandarin interrogativity. As shown in Figure 1.2, statements across all final tones appear to be marginally better perceived in unfiltered speech compared to filtered speech. Questions ending in Tones1 and 2 attract ostensibly similar levels of perception accuracy across the two speech modes. But questions ending in Tone3 appear to attract a higher accuracy in filtered compared to unfiltered speech, while those ending in Tone4 attract a seemingly higher accuracy in unfiltered speech, compared to filtered speech. Given these results, it could be said that language context generally had a minimal effect on the perception of Mandarin intonation by Cantonese listeners in B. R. Xu and Mok (2014).

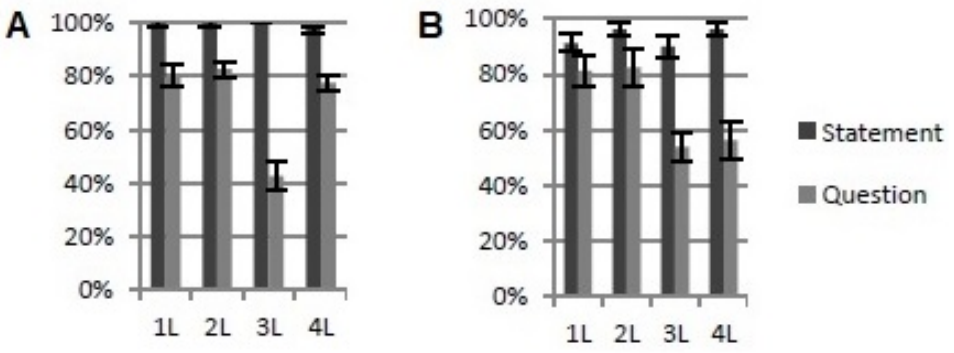

FIGURE 1.2: Cantonese listeners' perception accuracy of Mandarin interrogativity in B. R. Xu and Mok (2014). (A) shows their perception accuracy in unfiltered sentences. (B) shows their perception accuracy in filtered sentences (taken from B. R. Xu and Mok (2014), Figures 2 and 4 respectively).

Casting further doubt on the facilitative effect of the absence of language context is the study of M. Liu et al. (2016a). In an experiment with a similar design to the experiments of B. R. Xu and Mok, they tested Mandarin listeners' ability to identify Mandarin interrogativity in neutral and constraining language contexts. The experiment showed that their ability was generally better in a constraining language context compared to a neutral one. Extending B. R. Xu and Mok (2012b)'s claim that language context facilitates the perception of interrogativity, M. Liu et al. (2016a) suggested that Mandarin listeners' identification of interrogativity increased over three levels of increasing language context. Their ability was proposed to be weakest where there was no context, better in a neutral context, and strongest in a constraining language context.

M. Liu et al. (2016a)'s findings prompt several issues. Firstly, it reinforces the need to affirm whether the perception of intonation is better or worse in a non-speech context compared to a speech context. Secondly, would their finding of a facilitative effect of a constraining context over a neutral one also apply to non-native listeners? The studies that have examined the effect of language context on non-native perception of intonation (Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014) compared the dichotomous effect of the presence and absence of language context, without making further levels of distinction within sentences with language context. Lastly, M. Liu et al. (2016a)'s suggestion that the accuracy of native perception of Mandarin interrogativity could be gradated across three levels of language context – namely none (through filtered speech), neutral and constraining – has not been considered in relation to either native or non-native perception of intonation. To make a foray in this area, this thesis examines how Dutch and Mandarin listeners' ability to perceive interrogativity in Mandarin correlates with these three levels of language context.

1.2.3 Language-specific effects on the perception of intonation without language context

Another emerging issue in the perception of intonation is if the language-specific effects found in the perception of intonation, as raised in subsection 1.2.1, are only present where there is a language context. Only a handful of studies have addressed this issue. In the earliest study on this subject, Grabe, Rosner, García-Albea, and Zhou (2003) exposed British English, Spanish and Mandarin listeners to 11 general intonational contours of Southern British English, in both a speech version and a version using frequency-modulated sine waves. Listeners were asked to rate the similarity of pairs of contours. The study found that all listeners distinguished between the non-speech contours in the same way, with differences between listener groups emerging in the perception of normal speech contours. This result suggests that listeners' auditory processing of relatively slow-changing movements in pitch is based on a universal auditory mechanism that is not affected by language experience.

Conversely, Ortega-Llebaria and Colantoni (2013) and B. R. Xu and Mok (2012a, 2012b, 2014) found that the influence of language experience in the perception of intonation extends to a non-speech context. Ortega-Llebaria and Colantoni (2013) found that native English listeners, as well as non-native Spanish and Mandarin listeners, had a weaker perception of English focus in unfiltered sentences, compared to low-pass filtered sentences. The effect of language experience was discernible in both types of sentences. Compared to English listeners, Spanish listeners were less able to perceive focus in non-final positions in both unfiltered and filtered sentences. This was

attributed to a "negative" first-language transfer effect, as word order is a more salient marker of focus than intonation in Spanish. Mandarin listeners' ability to perceive focus in all positions in both sentence types was lower than that of English listeners.

In the two sets of experiments conducted by B. R. Xu and Mok (2012a, 2012b, 2014), Mandarin listeners' accuracy rates in identifying Mandarin interrogativity were generally higher than Cantonese listeners in both filtered and unfiltered speech. This result can be attributed to different markers of questions in the two languages. Cantonese marks questions with a terminal incline. Unlike Mandarin, Cantonese statements and questions do not differ in their global fundamental frequency contour (Ma, Ciocca, & Whitehill, 2011). The results of B. R. Xu and Mok suggest that the respective linguistic representations of interrogativity in Cantonese and Mandarin listeners carry over into their auditory processing mechanisms.

Evidently, there is no agreement on whether the effect of language experience extends to the perception of intonation in the absence of language context. It is reasonable to speculate that the emergence of this effect may be dependent on the perception task at hand (Bent, Bradlow, & Wright, 2006). Under the current study, we conduct a perception experiment that closely matches that of the studies of B. R. Xu and Mok. Thus we will be able to examine if the language-specific effects in the perception of Mandarin interrogativity in non-speech contexts can be replicated with this type of task.

1.2.4 Effect of tone on the perception of intonation without language context

Given that the studies of B. R. Xu and Mok are based on the tonal language of Mandarin, the question also arises as to whether the final tone influences non-native listeners' ability to perceive interrogativity in Mandarin. Several studies have investigated this question in relation to a speech context (Yang & Chan, 2010; S. Luo & Lin, 2015). These studies have been motivated by the fact that the final pitch direction is the most salient cue of interrogativity in languages including English and Dutch, with a final incline signalling a question (see e.g., Haan, van Heuven, Pacilly, & van Bezooijen, 1997; van Heuven & Haan, 2000; Heuven & Haan, 2002). This raises the possibility that non-native listeners may use the final tone movement rather than the global pitch contour as the cue for perceiving interrogativity in Mandarin. This was found to be the case in Yang and Chan (2010) and S. Luo and Lin (2015), who found that English learners of Mandarin tended to identify sentences ending with the rising Tone2 and the falling Tone4 as questions and statements respectively.

The studies of B. R. Xu and Mok found that this transfer of native intonational structures by non-native listeners also extended to the auditory processing level. That

is, they found that non-native listeners' ability to perceive intonation in non-speech contexts was also influenced by the final tone. In reviewing Figures 1.1 and 1.2, it can be seen that the Cantonese listeners in the experiments of B. R. Xu and Mok generally identify questions ending with the high-level Tone1 in filtered speech relatively well, while those ending with the low-dipping Tone3 are identified the worst. This pattern can be attributed to Cantonese listeners' transfer of their native intonational cues of interrogativity; in Cantonese, the salient marker of questions is a terminal incline that modifies that shape of the lexical tone on the final syllable, as explained previously (Ma et al., 2011). Our current study examines whether the effect of the final tone can be replicated in Dutch listeners' perception of Mandarin interrogativity in different levels of language context.

1.3 Research questions and hypotheses

To illuminate the above issues, our main research question is: how does the degree of language context affect the identification of Mandarin intonation by Dutch listeners? We thus examine how well they can perceive intonation as a cue of Mandarin interrogativity across three degrees of language context: no language context, neutral context and constraining context.

We hypothesize that Dutch listeners' ability to identify intonation will be best where there is no language context. This is given that the behavioural results from previous studies (Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b) are weighted towards the idea that non-native listeners perceive intonational meaning more accurately in a non-speech context than in a speech context. Secondly, M. Liu et al. (2016a)'s finding on the facilitative effect of a constraining context leads us to hypothesize that Dutch listeners would perceive intonation as a cue of Mandarin interrogativity better in a constraining context than in a neutral context.

We also examine whether the identification of Mandarin intonation by Dutch listeners in different contexts is influenced by two specific factors. Firstly, we assess if language experience (i.e. Mandarin vs Dutch listeners) affects the identification of Mandarin intonation in different contexts. We hypothesize that Mandarin listeners would perceive interrogativity better than non-native Dutch listeners, given the language-specific effects in intonation perception. Moreover, the native listeners' superior ability is hypothesized to extend to non-speech contexts. This is given the similarity of the task in our current study (which we outline in detail in the next chapter) with the task in B. R. Xu and Mok, who found that the language-specific effects of intonation extended to non-speech contexts.

Secondly, we assess whether the sentence-final lexical tonal identity affects Dutch listeners' identification of intonation. As mentioned in subsection 1.2.4, studies have found that the final tone influences non-native listeners' ability to perceive intonation as a cue of Mandarin interrogativity, in both speech and non-speech contexts. We thus hypothesize that Dutch listeners' identification of interrogativity will be similarly affected by the final tone in the three levels of language context that we will examine.

Chapter 2

Method

To examine how language context affects non-native listeners' ability to perceive interrogativity in Mandarin, we conducted a perception experiment with Mandarin listeners and Dutch listeners who understand Mandarin. They were tested on how accurately and quickly they perceived statements and questions at three levels of language context: none, neutral and constraining. This section provides a detailed explanation of the methodology behind the experiment. It firstly details the creation, recording and preparation of stimuli, followed by the conduct of the perception experiment. All stages of the research complied with the Ethics Code for linguistics research set out by the Leiden University Centre for Linguistics.

2.1 Creation of stimuli

Our stimuli were created on the basis of 80 base sentences for each of the three levels of language context. To tap into the effect of tone, the 80 base sentences comprised four groups of 20 sentences, with each group defined by different tones on the final syllable (representing the four standard tones in Mandarin). Sentences ending with the low-dipping Tone3 were designated to be the filler stimuli in our perceptual experiment. This was to avoid the complexities associated with analysing Tone3. It is the lexical tone in Mandarin with the most phonetic variation, as it can be realized with or without a final rising incline (Yuan, 2006, p.28; Duanmu, 2007, pp.238-9).

The primary issue in the creation of sentences was constructing suitable stimuli sentences for neutral and constraining context. Below, we outline the process by which we constructed these sentences.

2.1.1 Neutral language context

Sentences of neutral language context are defined as those in which the final syllable cannot be determined from the sentence's immediate preceding portion. We constructed

these sentences based on a group of 80 monosyllabic words representing the four tones in Mandarin. Each of these words was embedded as the final syllable in the sentence "ta1 gang1 gang1 shuo1 X" ("he just said X"[1]), the same sentence used in M. Liu et al. (2016a) to denote a neutral semantic context.

The main criterion for selecting the monosyllabic words was that they be comprehensible to non-native Mandarin learners. We thus compiled a list of the monosyllabic words required to be learnt for the first five levels of the Hanyu Shuiping Kaoshi ("HSK"), an international standardized exam that tests non-native listeners' Mandarin proficiency.[2]

We further narrowed down this list of suitable monosyllabic words based on several criteria. Firstly, we limited words to be those that can be generally used and understood as a single word. For example, the syllable "che1" ("car") can be understood as a standalone word. Such words are in contrast to monosyllabic words which are generally not used as standalone words but as part of a longer word. For example, the constituent syllables of "xi3 huan1" ("to like") are generally not used as standalone words. Secondly, we limited the word class of the monosyllabic words to nouns, verbs and adjectives. Ideally, the words should belong to the same word class, but we broadened the scope of word class to ensure that there would be sufficient stimuli in our experiment.

Thirdly, we further reduced the list based on phonological neighbourhood density, which refers to the number of words with the same pronunciation in segments and tone (i.e., homophones). This has been shown to have an effect on lexical processing in Mandarin (H.-C. Chen, Vaid, & Wu, 2009), and could potentially also affect the processing of intonation. The phonological neighbourhood density for each word was calculated from a list of Chinese characters drawn from a corpus of Modern Chinese e-texts, which were compiled by Da (2004).[3] Words with four or more homophones were then excluded from the list. Fourthly, we only included words with a high word frequency. Following M. Liu et al. (2016a), we define this as words occurring greater than 4,500 times in Da (2004)'s corpus. Finally, of the remaining words, we selected 80 words representing all four Mandarin tones, such that they were balanced across segments. A list of these words can be found in Appendix A.

[1] The number after each syllable denotes its lexical tone.

[2] Words lists for the HSK can be found on the website of *Hanban*, the administering organization of the HSK, at http://www.chinesetest.cn/userfiles/file/HSK/HSK-2012.xls (valid as of 26 April 2018).

[3] A link to the list can be found at http://lingua.mtsu.edu/chinese-computing/statistics/char/CharFreq-Modern.xls (valid as of 26 April 2018).

2.1.2 Constraining language context

Sentences of neutral language context are defined as those in which the final syllable, and thereby the final tone, can be determined from the pre-final portion of context. These sentences were constructed from 80 disyllabic words.

The mandatory vocabulary lists for the HSK were again used as the source of words. We firstly compiled all disyllabic words from the mandatory vocabulary lists for the first four levels of the HSK into an initial list. We then created sentences which ended with each of these disyllabic words. The creation of the sentences was underpinned by the following principles:

- *Length*: Sentences were to range from six to eight syllables, to ensure comparable word length.
- *Difficulty*: Sentences were created with vocabulary that were as simple as possible for third-year undergraduate students of Mandarin, who represented the minimal level of Mandarin required of non-native participants in our experiment, to understand. An experienced teacher of Mandarin verified that the sentences were comprehensible for students of this level of Mandarin.
- *Constraining context*: Sentences were constructed such that the context limited the potential disyllabic words that could end the sentence. For example, various disyllabic words could conclude the sentence "Fang2 jian1 li3 you3 yi1 zhang1 *X X*" ("Inside the room there is a *X X*"). In some sentences, the context was such that the final dissyllabic word could be predicted with near certainty. Such sentences, including "Zhong1 guo2 de shou3 du1 shi4 *Bei3 Jing1*" ("The capital of China is *Beijing*"), were discarded to ensure a commensurate level of constraint in the context across the sentences.
- *Subject pronouns*: Most sentences were constructed with the third-person subject pronoun – "ta1" ("he", "she" or "it") – as this has been contended to have no biasing effect on listeners' identification of interrogativity. In contrast, sentences with the first-person subject pronoun are prone to be interpreted as statements, and those with the second-person subject pronoun as questions (e.g., Beun, 1990). As such, use of the first or second person pronoun ("wo3" and "ni3" respectively) was avoided.

Finally, of the remaining words, we selected 80 sentences with their final tone representing all four Mandarin tones, such that they were balanced across segments. A list of these sentences can be found in Appendix B.

2.2 Recording of stimuli

One native female Mandarin speaker (26 years old), born and raised in Northern China, served as the speaker of the stimuli. She was recorded for two versions of each of the 160 sentences we had created: one version in statement intonation, and the other in question intonation. The speaker was recorded in the sound-attenuated phonetics booth at Leiden University, at 16-bit resolution and at a sampling rate of 44.1 kilohertz. The sentences were presented to the speaker in random order, one by one in Chinese characters on a computer screen.

The speaker was specifically instructed to produce sentences ending with a question mark as questions, and those ending with a period as statements. Moreover, she was instructed to say the sentences naturally without emphasis on any particular words or exaggerated emotional prosody. This was to ensure the control of focus, as focus has been shown to affect the production and perception of interrogativity in Mandarin (e.g., F. Liu & Xu, 2005).

Additionally, it was important to ensure that our recordings of sentences and questions aligned with the prototypical fundamental frequency ("F0") pattern of Mandarin interrogativity described in past studies. That is, the global F0 contour in the recordings of our questions should be higher than that of the statements, with the difference increasing exponentially over time such that the greatest difference is found in the final syllable (e.g., Yuan, 2006). To ensure our recordings reflected this acoustic pattern, we had originally recorded two native female Mandarin speakers from Northern China. An acoustic analysis of both speakers through the speech processing software Praat (Boersma & Weenink, 2017) revealed that the production of one of the speakers was more aligned with the prototypical F0 pattern of interrogativity.

Wilcoxon matched-paired tests confirmed that in the chosen speaker's recordings, the F0 of questions was significantly higher than that of statements in the majority of all syllables, in both sentences of neutral and constraining context (all ps < 0.05).[4] The difference in F0 peaked at the final syllable in all sentences. Figures 2.1 and 2.2 depict this quasi-exponential increase in the F0 difference between statements and questions over time in a sentence of neutral and constraining context.

In relation to the durational properties, each syllable before the final one in statements was generally longer than those in questions. These differences were mostly significant in sentences of neutral context, and in six- and seven-syllable sentences of constraining context (all ps < 0.05), but insignificant in eight-syllable sentences of constraining context (all ps > 0.05). The final syllable in all sentences was longer than in questions, with the difference being significant in sentences of neutral context and

[4]The statistical analyses relating to the sentences of constraining context were separated by the length of the sentences (i.e., six, seven and eight syllables).

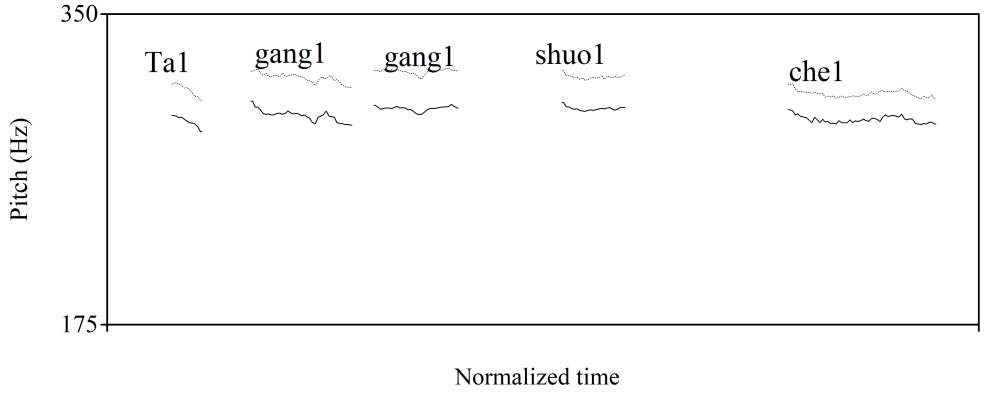

FIGURE 2.1: F0 contours of a sentence of neutral language context: "Ta1 gang1 gang1 shuo1 che1" ("He just said 'car' "). The darker lines depict the F0 contour of the version produced as a statement; the grey lines correspond to that produced as a question.

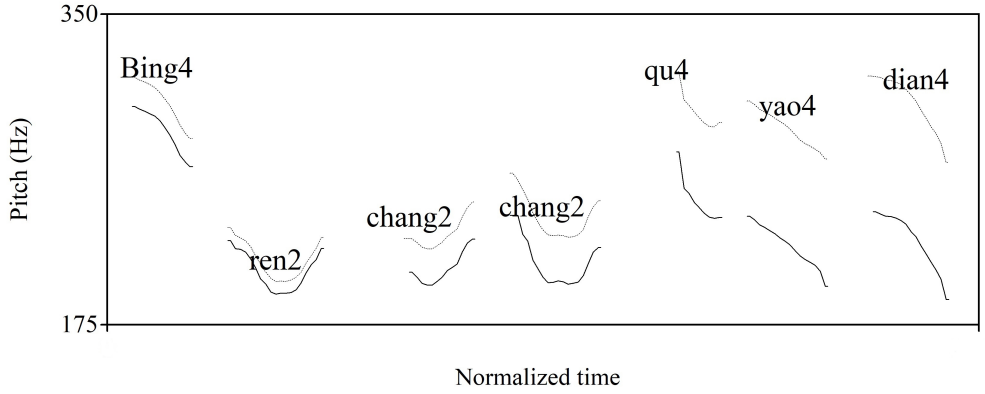

FIGURE 2.2: F0 contours of a sentence of constraining language context: "Bing4 ren2 chang2 chang2 qu4 yao4 dian4" ("Sick people often go to the pharmacy"). The darker lines depict production as a statement; the grey lines as a question.

seven-syllable sentences of constraining context (both ps < 0.01), but insignificant in six- and eight-syllable sentences of constraining context (both ps > 0.05). These durational patterns are consistent with the durational trends found by Yuan (2006). Nevertheless, we will assume that the primary perceptual cue of interrogativity is pitch. This is given not only the identification of F0 as the primary acoustic marker of intonation in Mandarin (e.g., Y. Xu & Wang, 2001; Y. Xu, 2004), but also because of the centrality of F0 in the production and perception of intonational meaning (e.g., Vaissière, 2005).

From the recording, we obtained 320 sentences (80 base sentences x 2 intonation types (i.e. sentence or question) x 2 contexts). The amplitude of these sentences was normalized to 75 decibels (dB). To generate sentences of no language context, the 160 sentences of neutral context were low-pass filtered at 400 hertz (Hz) with 100 Hz

bandsmoothing. The threshold of 400 Hz was based on acoustic analyses of the maxima F0 of the final syllables in the sentences of neutral context. The low-pass filtered stimuli were normalized to an amplitude of 82 dB to ensure that the stimuli had a commensurate perceived intensity with the unfiltered stimuli.

2.3 Subjects

Most subjects were recruited from student and alumni groups of Leiden University. Twenty-two subjects (17 females, 5 males), aged between 20 and 30 years old ($M \pm SD$: 23.6 $\pm$ 2.6), participated as Dutch listeners. They were all raised in the Netherlands and had completed or were completing a Mandarin course at a third-year undergraduate level. All Dutch listeners had lived in China or Taiwan from three months to three years (M = 11 months).

Twenty-two subjects (19 females, 3 males), aged between 19 and 30 years old ($M \pm SD$: 25.2 $\pm$ 2.7), participated as Mandarin listeners. They were all born and raised in China and had limited experience living outside China, ranging from one month to 2.5 years (M = 9 months). To ensure a degree of homogeneity in the language background of the Mandarin listeners, we recruited Mandarin listeners whose native dialect was a Mandarin dialect. Those who used non-Mandarin dialects such as Cantonese or Shanghainese were excluded.[5]

All subjects could speak English as a second language. None of the subjects had any reported speech or hearing problems. All subjects gave informed consent before the experiment and were reimbursed for their participation.

2.4 Perception experiment

Participants completed an intonation identification task. Instructions were given to participants orally by the experimenter in English before the experiment, as well as on a monitor during the experiment. Participants were tested individually in the sound-proof phonetics booth at Leiden University. Sentences were played through headphones using the experimental software E-Prime 2.0 (*E-Prime*, n.d.) at a comfortable listening level. Throughout the experiment, participants were to fix their gaze on a cross on the monitor to help maintain their focus.

The presentation of each sentence stimulus followed the method used in M. Liu et al. (2016a). The presentation began with a 100 millisecond (ms) warning beep, followed by a 300 ms pause. The sentence was then played. Participants had two seconds from the offset of the sentence to indicate whether they heard a question or

[5] See e.g., Norman (1988, p.191) for a detailed discussion of the categories of Chinese dialects.

statement. They indicated this on a keyboard with either the "f" or "j" key, with the left and right index fingers respectively. The coding of the keys differed for every other participant, such that one participant identified questions with his or her preferred hand, and the next participant did so with his or her non-preferred hand (J. Pacilly, personal communication). Once the participant pressed a key, a 500 ms pause was activated before the presentation of the next sentence stimulus. If no response was given within two seconds, the program automatically presented the next stimulus, with a preceding 500 ms pause.

Each participant listened to 480 sentences altogether (160 sentences x 3 levels of context). The experiment was completed in two parts. In the first part, participants listened to the 160 filtered sentences in randomized order. Playing the non-speech stimuli first was deliberately aimed at tapping into participants' auditory processing level and at minimizing the possibility that they would process the stimuli as speech (e.g., Bent et al., 2006; Huang & Johnson, 2010). Listeners completed one practice block of eight trials. Following this, they listened to the randomized filtered sentences in two blocks of 80 trials, with a short break between the blocks.

In the second part, the 320 sentences of neutral and constraining context were randomly mixed in the same blocks to minimize potential learning affects from repeated listening to sentences of the same level of language context (M. Liu, personal communication). Just before the second part of the experiment, participants were told that none of the filtered sentences had any lexical markers of questions. This was to minimize the possibility that their identification would be based on the presence or absence of their markers. Listeners completed another practice block of eight trials relating to the unfiltered stimuli. They then listened to the randomized unfiltered sentences in four blocks of 80 trials, with short breaks between all blocks. The experiment lasted 30 minutes on average.

2.5 Analysis of results

The E-Prime software provided two pieces of information for each sentence and each listener: (1) a response as to whether the interrogativity identification was correct or incorrect; and (2) the reaction time (RT). Responses enabled us to obtain identification rates ("ID rates"), which are defined as the percentage of statements or questions identified correctly. RTs are also analyzed because they reflect how easy a perceptual decision is, as shown by studies investigating the identification of interrogativity in German (Schneider, Dogil, & Möbius, 2011) and Mandarin (M. Liu et al., 2016a). Following these studies, we define RT as the time from the onset of the final syllable for correct responses.

Before analyzing results, the data was cleaned so that responses given before the onset of the final syllable were excluded. Data points were also removed where no response was given within two seconds of the onset of the final syllable. Finally, responses that had a RT of three standard deviations beyond the mean were excluded for each listener, following Baayen (2008, p.244).

Analyses of results were conducted in the statistical processing software R version 3.4.3 (R Core Team, 2013) using the lme4 package (Bates, Maechler, Bolker, & Walker, 2013) that enables mixed-effects regression models to be generated. We built two models corresponding to the ID rate and RT for the overall data set. For analyzing ID rate, a mixed-effects binomial logistic regression model was constructed with the following main fixed effects: degree of language context ("context": 0 (none) vs. 1 (neutral) vs. 2 (constraining)); interrogativity of the sentence ("interrogativity": sentence vs question); final tone of the sentence ("tone": Tone1 vs. Tone2 vs. Tone4); and language background ("background": Dutch vs. Mandarin), as well as their interactions. Firstly, models with each of the individual fixed effects were compared with a null model with only the random effects of subject (44 listeners) and item (120 different sentences). Log-likelihood ratios were used to evaluate the significance of each of the fixed effects. Only effects that were found to be significant were added to the model.

Secondly, this process was repeated with the two-way interactive effects. That is, the model (with significant individual fixed effects) was compared to models with each of the two-way interactive fixed effects. Once again, only effects that were found to be significant were added to the model. This process of evaluating the significance of each interactive effect through comparison with a model containing significant lower-order effects was repeated in relation to the three-way and four-way interactive fixed effects. After adding the four significant main fixed effects and their interactions, trial-by-trial dependency ("trial") – i.e. the order in which sentences were presented within each level of language context – was then added as another fixed effect, as it was found to have a significant, albeit small effect.

In the final stage of building the model, random slopes were added for the by-subject effect of tone, by-subject effect of intonation, and by-item effect of intonation. The by-item effect of tone was found to be insignificant, and was thus excluded in the final model.

For the analysis of RT, a mixed-effects linear regression model was employed, with the same fixed effects, random effects and random slopes added and evaluated in the same way. RT was log-transformed beforehand for better normalcy.

The fit of each of the models was evaluated using marginal and conditional R^2 values computed with the MuMIn package (Bartoń, 2018). The marginal R^2 value measures the variance accounted for by fixed effects, and the conditional R^2 value

represents the variance accounted for by fixed effects, and random effects and slopes.

To assess the significance of differences between levels of in a fixed effect, post-hoc pairwise tests were conducted, using the lme4 package again. As the lme4 package does not provide p-values for pairwise tests based on mixed-effects linear models, p-values for pairwise tests in relation to RT were obtained through a supplementary lmerTest package, which computes the p-values based on Satterthwaite's approximations (Kuznetsova, Brockhoff, & Christensen, 2017). Effect sizes of differences are mentioned where relevant. For ID rate, the relative risk (RR) – the ratio of the two ID rates being compared (Davies, Crombie, & Tavakoli, 1998) – was used as the measure of effect size of the probability of a correct response over an incorrect one, as the more conventional measure of the odds ratio produced inflated results.[6] For RT, the r-value according to Rosenthal (1991, p.19 (equation 2.16)) was used as a measure of effect size.

[6]This is in line with evidence demonstrating that the odds ratio overstates the effect size where the prevalence of an outcome is high in the two groups that are being compared (see Altman, Deeks, & Sackett, 1998; Davies et al., 1998).

Chapter 3

Results

3.1 Overall results

The presentation of the results begins with a summary of the statistical results for the mixed-effects models corresponding to the ID rate and RT for the whole data set. The summary is presented in Table 3.1. The χ^2 values, df and p-values for each of the fixed effects, random effects[1] and random slopes were calculated from loglikelihood-tests.

Most of the four main fixed effects (i.e., context, interrogativity, tone and background) and their interactions were significant in both models. To explore the results, we firstly examine the main fixed effect of context in more detail. This is because the effect of context is relevant to our main research question as to how listeners' ability to identify interrogativity differs with different levels of context. Although there are higher-order significant interactive effects, we will demonstrate that the effect of context is consistent in both the whole data set, and in further data subsets.

Secondly, we will examine the highest-order interactive effects on the whole data set in relation to ID rate and RT respectively. Thus we will examine the significant four-way interaction of context x interrogativity x tone x background on ID rate. We will then examine the three significant three-way interactive effects on RT, namely: interrogativity x tone x background; context x interrogativity x background; and context x interrogativity x tone. These interactions shed light on our ancillary research questions as to whether language experience and the sentence-final lexical tonal identity affects listeners' identification of Mandarin intonation in different contexts.

[1]The χ^2 values in relation to the random effects (by-subject and by-item) represent the difference between the full model (with all fixed and random effects) and the model without the random effects.

	ID rate			RT		
Fixed effects	χ^2	df	p	χ^2	df	p
context	290.75	2	$<.001$	214.23	2	$<.001$
interrogativity	536.42	1	$<.001$	83.45	1	$<.001$
tone	1.25	2	0.53	6.57	2	0.04
background	14.91	1	$<.001$	11.41	1	$<.001$
context x interrogativity	342.14	2	$<.001$	67.84	2	$<.001$
context x tone	6.09	6	0.41	5.26	4	0.26
context x background	36.48	2	$<.001$	362.41	2	$<.001$
interrogativity x tone	971.99	4	$<.001$	130.77	2	$<.001$
interrogativity x background	64.43	1	$<.001$	33.16	1	$<.001$
tone x background	4.24	4	0.37	7.34	2	0.03
interrogativity x tone x background	27.28	4	$<.001$	12.22	2	0.002
context x tone x background	53.17	10	$<.001$	11.71	8	0.16
context x interrogativity x background	17.25	2	$<.001$	22.27	2	$<.001$
context x interrogativity x tone	107.48	8	$<.001$	28.68	8	$<.001$
context x interrogativity x tone x background	21.98	4	$<.001$	14.88	8	0.06
trial	21.69	1	$<.001$	76.32	1	$<.001$
Random effects and slopes						
1|subject	206.85	1	$<.001$	3,211.40	1	$<.001$
1|item	499.11	1	$<.001$	152.78	1	$<.001$
1+tone|subject	65.00	5	$<.001$	22.89	5	$<.001$
1+interrogativity|subject	436.24	2	$<.001$	162.58	2	$<.001$
1+interrogativity|item	316.47	2	$<.001$	41.57	2	$<.001$
Marginal R^2		0.70			0.24	
Conditional R^2		0.78			0.48	

TABLE 3.1: Summary of mixed effects models of all listeners' ID rates and RTs.

3.2 Effect of context

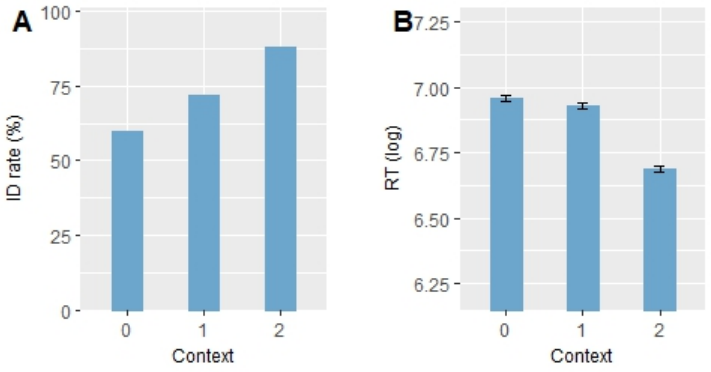

FIGURE 3.1: (A) ID Rates and (B) RTs in each context. Contexts 0, 1 and 2 refer to a non-speech, neutral and constraining language context respectively.

The effect of context on response and RT is illustrated in Figure 3.1. It is evident that the ID rate increases with each increasing level of context. ID rates are significantly

higher in context 1 compared to context 0 ($p < .001$), and in context 2 compared to context 1 ($p < .001$). In relation to RT, responses appear to be quicker with each increasing level of context. RTs are significantly quicker in context 1 than context 0 ($p < .001$, $r = 0.03$), although the effect size is small. Context 2 attracts a quicker RT than context 1 ($p < .001$, $r = 0.88$).

To ascertain whether this effect of context holds across all conditions of tone and interrogativity, within both Dutch and Mandarin listeners, we break down the results further by tone, interrogativity and background. Figure 3.2 displays Dutch listeners' ID rates at each level of context in each condition of tone, with the two plots showing these results in relation to statements and questions respectively. Figure 3.3 shows the same ID rates in relation to Mandarin listeners. The effect of context is significant in most of the six different sentence conditions of tone and interrogativity (i.e., Tone1 statements, Tone2 statements, Tone4 statements, Tone1 questions, Tone2 questions and Tone4 questions), for both listener groups (all ps $< .001$). Across most of these six conditions, ID rates generally increase with each increasing level of context, such that ID rates in context 1 are significantly higher than in context 0 (all ps < 0.01), and those in context 2 are significantly higher than in context 1 (all ps < 0.05).

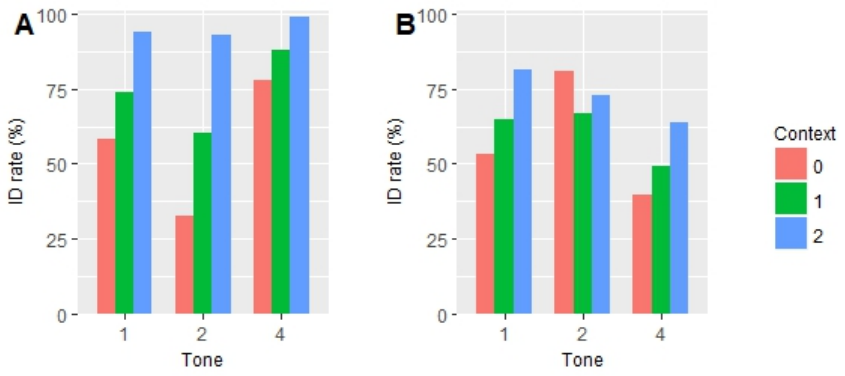

FIGURE 3.2: Dutch listeners' ID rates in each context in each tonal condition in (A) statements and (B) questions.

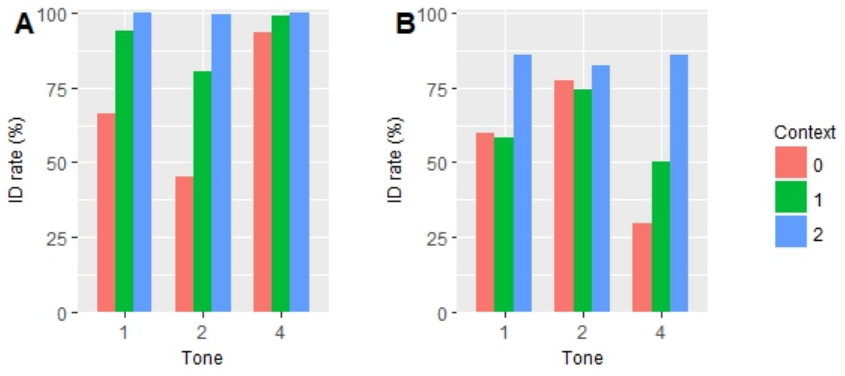

FIGURE 3.3: Mandarin listeners' ID rates in each context in each tonal condition in (A) statements and (B) questions.

Notably, this effect of context is not evident in Tone2 questions. For Dutch listeners, Tone2 questions are most accurately perceived where there is no language context (80.87%). Their ID rate is significantly lower in context 1 (66.97%) ($p < .001$). Moreover, their ID rate in context 2 (73.04%) is not significantly higher than in context 1 ($p = 0.15$). For Mandarin listeners, the effect of context in their ID rate of Tone2 questions is insignificant ($p = 0.08$). There are no significant differences in their ID rates between contexts 1 and 0 (74.25% vs. 77.31%: $p = 0.22$), nor between contexts 2 and 1 (82.15% vs. 77.31%: $p = 0.06$). Furthermore, there are no significant differences in Mandarin listeners' ID rates between contexts 1 and 2 in Tone4 statements (which have reached a ceiling level at 99.09% and 100.00% respectively: $p = 0.99$), and between contexts 0 and 1 in Tone1 questions ($p = 0.63$).

The effect of context on Dutch listeners' RTs in all tonal conditions, in both statements and questions, is shown in Figure 3.4. Figure 3.5 shows the same results for Mandarin listeners. For both listener groups, the effect of context is significant in both sentences and questions, across all tones (all ps < 0.01). Context 2 consistently attracts a quicker RT than context 1 in all sentence conditions for both listener groups (all ps $< .001$). The differences between contexts 1 and 0 do not appear to be as large. For Dutch listeners, RTs are not significantly different between contexts 1 and 0 in most sentence conditions (all ps > 0.30). This is not the case for Tone1 statements, which attract a significantly quicker RT in context 0 compared context 1 ($p = 0.04$, $r = 0.07$). For Mandarin listeners, in contrast, responses are significantly quicker in context 1 than context 0 in most sentence conditions (all ps < 0.05). This is not the case for Tone1 and Tone2 questions, where there are no significant differences in RTs between contexts 0 and 1 (both ps > 0.10). Nevertheless, the effect size of the difference in Mandarin listeners' RTs between contexts 1 and 0 has a r-value of at most 0.16 (in Tone1 statements). This contrasts with their larger differences between contexts 2 and 1: the minimum effect size of these differences has a r-value of 0.81 (in Tone1 questions).

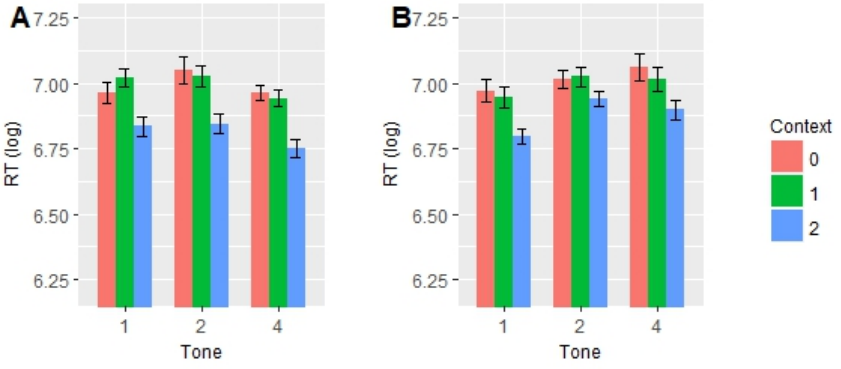

FIGURE 3.4: Dutch listeners' RTs in each context in each tonal condition in (A) statements and (B) questions.

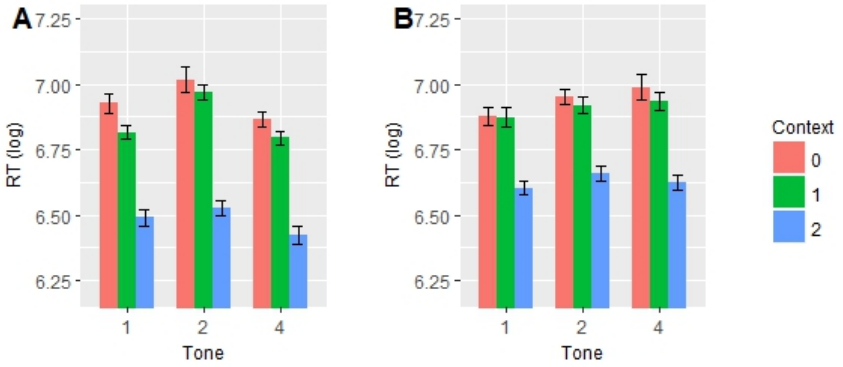

FIGURE 3.5: Mandarin listeners' RTs in each context in each tonal condition in (A) statements and (B) questions.

Overall, these results indicate that the identification of Mandarin interrogativity generally improves as language context increases. This trend is especially borne out in the ID rates, which increase with each increase context. The facilitative effect of context is less salient in the RTs. Although a constraining language context assists listeners in identifying interrogativity quicker than a neutral context, the facilitative effect of a neutral context over a non-speech context is markedly smaller.

3.3 Significant interactive effects on ID rate

The four-way interactive effect of context x interrogativity x tone x background on ID rate is displayed in Figures 3.6 and 3.7 below. Figure 3.6 shows the ID rates in sentences and questions in each tonal condition, with the three plots displaying this data in each level of context for Dutch listeners. Figure 3.7 shows the same data in relation to Mandarin listeners.

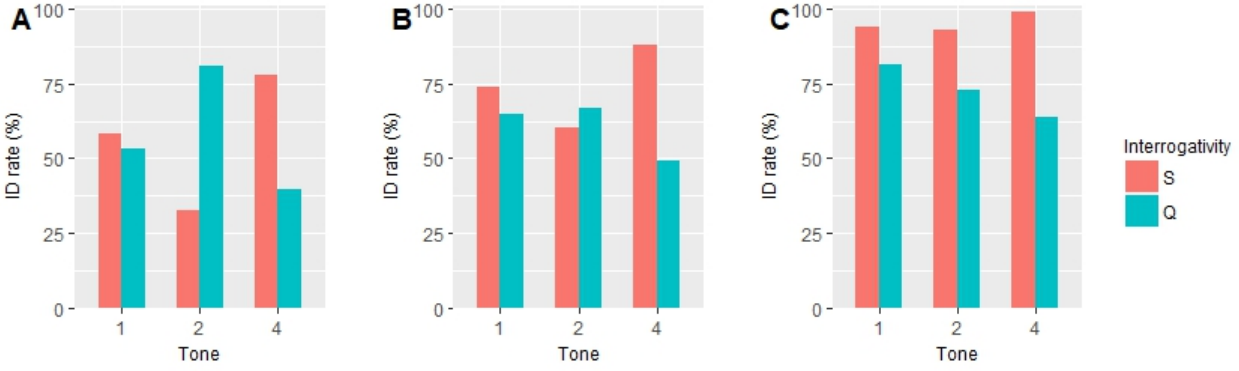

FIGURE 3.6: Dutch listeners' ID rates in (A) context 0, (B) context 1 and (C) context 2. "S" and "Q" refer to statements and questions respectively.

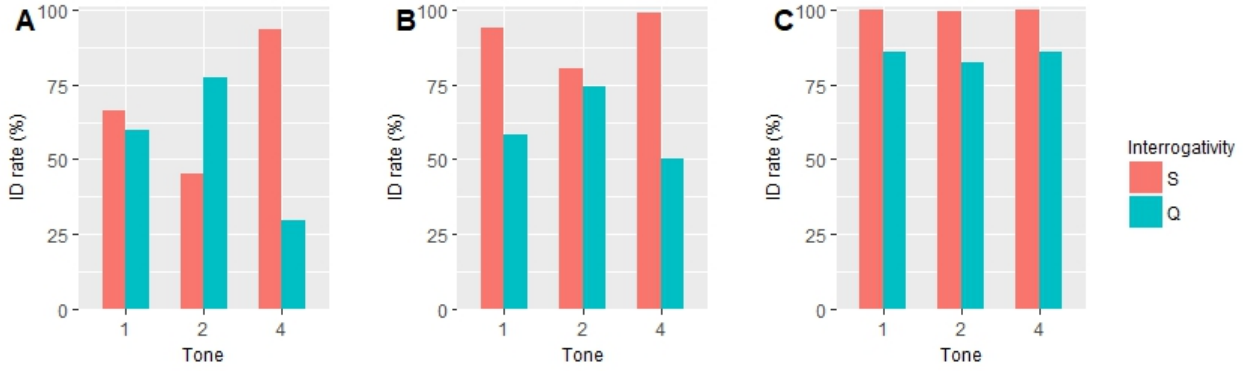

FIGURE 3.7: Mandarin listeners' ID rates in (A) context 0, (B) context 1 and (C) context 2.

In relation to Dutch listeners' ID rate, the effect of context x tone x interrogativity is significant ($p < .001$). Moreover, the effect of tone x interrogativity is significant at each level of context (all ps $< .001$). Nevertheless, there appears to be a difference in the influence of tone in their identification of sentences and questions in contexts 0 and 1 on the other hand, and in context 2 on the other. In contexts 0 and 1, statements are generally identified with a significantly higher ID rate than questions, where sentences end in Tone1 or Tone4 (all ps < 0.01), with an insignificant difference between statements and questions in Tone1 sentences in context 0 ($p = 0.09$). In Tone2 sentences in contexts 0 and 1, questions attract a significantly higher ID rate than statements (both ps < 0.05). In context 2, on the other hand, statements are perceived significantly more accurately than questions (all ps $< .001$), regardless of tone.

In relation to Mandarin listeners' ID rate, the effect of context x tone x interrogativity is significant ($p < .001$). This can be explained by analyses revealing that the effect of interrogativity x tone is significant in contexts 0 and 1 (both ps $< .001$), but insignificant in context 2 ($p = 0.18$). Pairwise analyses nevertheless indicate that a difference can be drawn between the effect of tone on the identification of statements and questions in context 0 on the one hand, and contexts 1 and 2 on the other. In context 0, statements are perceived with a significantly higher ID rate than questions in Tone1 and Tone4 (both ps < 0.05), while questions are perceived with a significantly higher ID rate than statements in Tone2 ($p < .001$). In contexts 1 and 2, in contrast, statements generally attract a significantly higher ID rates than questions across all tones (all ps < 0.05). An insignificant difference can be found between the ID rates of questions and statements in Tone4 sentences in context 2 ($p = 0.99$).

Overall, these results show that the influence of the final tone on both Dutch and Mandarin listeners' identification of interrogativity varies according to the level of context. Where there is no language context, questions are identified better than statements where there is a final rising Tone2, while statements are generally identified

better than questions where there is a final falling Tone4 or high-level Tone1. This suggests that listeners tend to associate rising and falling tones with questions and statements respectively where there is no language context. Conversely, where there is a constraining language context, sentences are identified consistently better than questions, regardless of the final tone.

Where there is a neutral language context, there is a difference between Dutch and Mandarin listeners' ID rates. Here, Dutch listeners still associate rising and falling tones with questions and statements respectively. Conversely, Mandarin listeners identify statements better than questions regardless of the final tone. It is this difference that can account for the significant interactive effect of context x interrogativity x tone x background.

3.4 Significant interactive effects on RT

In this section, we provide a detailed examination of the results as they relate to the three highest-order interactive effects on RT that were significant. We note in passing that the highest-order interaction of context x interrogativity x tone x background on RT was not significant ($p = 0.06$). However, its near significant effect suggests a trend whereby RT is influenced by this four-way interaction.

3.4.1 Effect of interrogativity x tone x background

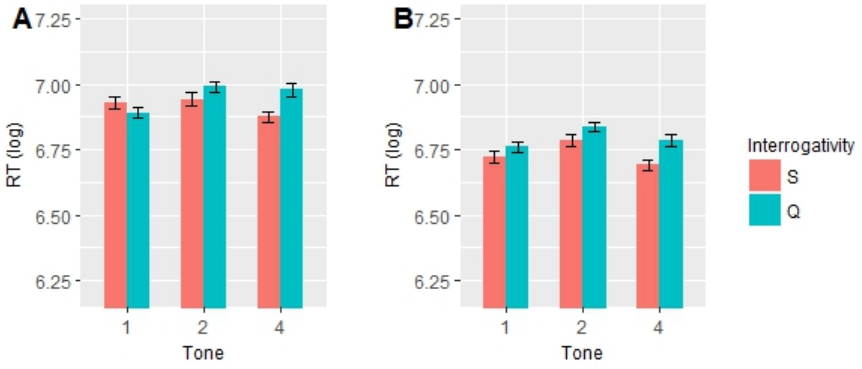

FIGURE 3.8: RTs of (A) Dutch and (B) Mandarin listeners in listeners in statements and questions in each tonal condition. Error bars represent the 95% confidence interval from the mean RT.

The significant effect of interrogativity x tone x background on RT is displayed in Figure 3.8. This shows the RTs of statements and questions in each tonal condition, with the two plots corresponding to the results of Dutch and Mandarin listeners respectively. For Dutch listeners, the effect of tone x interrogativity was significant ($p < .001$). For these listeners, questions attracted a quicker response than statements in Tone1 ($p <$

.001), while statements attracted a quicker response than questions in Tone4 ($p < .001$). There were no significant differences in their RTs between statements and questions in Tone2 ($p = 0.34$). For Mandarin listeners, the effect of tone x interrogativity was also significant ($p < .001$). Statements attracted quicker RTs than questions in Tone1 ($p < .001$, $r = 0.10$), Tone2 ($p = 0.02$, $r = 0.05$) and Tone4 ($p < .001$, $r = 0.27$). The significant effect of tone x interrogativity in Mandarin listeners' RTs may be attributed to differences in effect size.

The effect of interrogativity x tone x background can primarily be pinned down to the differences in the identification of Tone1 sentences between Dutch and Mandarin listeners. In this tonal condition, questions attract marginally quicker responses than statements for Dutch listeners. Conversely, statements attract slightly quicker responses than questions for Mandarin listeners.

3.4.2 Effect of context x interrogativity x background

The significant interactive effect of context x interrogativity x background on RT is displayed in Figure 3.9. It shows the RTs of Dutch and Mandarin listeners at each level of context, with the two plots displaying these RTs in relation to statements and questions respectively. In relation to statements, the effect of background x context was significant ($p < .001$). Although Mandarin listeners' RTs for statements are quicker than those of Dutch listeners in all contexts, they were significantly quicker only in contexts 1 and 2 (both ps < 0.05). The differences in the RTs between Dutch and Mandarin listeners in context 0 was insignificant ($p = 0.16$). In relation to questions, the effect of background x context was also significant ($p < .001$). Again, Mandarin listeners' RTs for questions were quicker than those of Dutch listeners across all contexts. However, the difference was only significant in context 2 ($p < .001$). Mandarin listeners' RTs were not significant quicker than those of Dutch listeners in contexts 0 and 1 (both ps > 0.05).

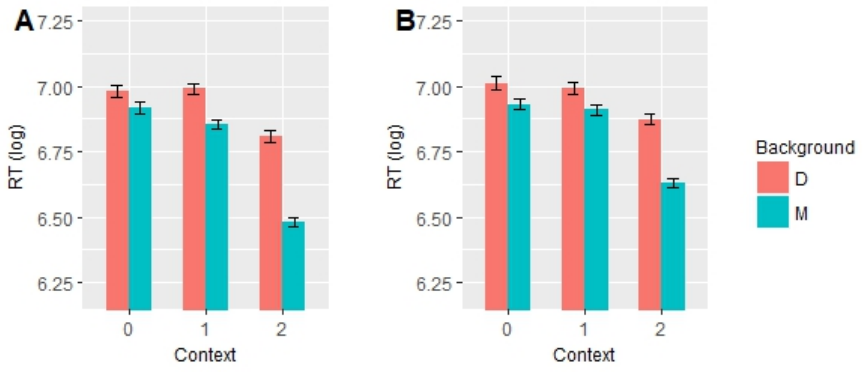

FIGURE 3.9: Dutch and Mandarin listeners' RTs in each context in (A) statements and (B) questions. "D" and "M" refer to Dutch and Mandarin listeners respectively.

The results in relation to RT suggest that Mandarin listeners' responses are significantly quicker than Dutch listeners' responses in a constraining context, but not in a non-speech context. This holds true in both statements and questions. The significant effect of interrogativity x context x background is attributable to differences between statements and questions in a neutral context. Here, Mandarin listeners' responses are significantly quicker than Dutch listeners in statements, but not significantly so in questions.

3.4.3 Effect of context x interrogativity x tone

The significant effect of context x interrogativity x tone is displayed in Figure 3.10. It shows the RTs for statements and questions in each tonal condition, with the three plots displaying these RTs at each level of context. At each level of context, the effect of interrogativity x tone was significant (all *p*s $< .001$). In context 0, RTs do not significantly differ between statements and questions in Tone1 ($p > 0.70$). In Tone2, questions attract a significantly quicker RT than statements ($p < 0.01$), while in Tone4, statements attract a significantly quicker RT than questions ($p < .001$). The same trends hold in context 1. In context 2, statements attract a significantly quicker RT than questions for Tone1 ($p = 0.02$, $r = 0.06$), Tone4 ($p < .001$, $r = 0.21$), and Tone4 ($p < .001$, $r = 0.30$). The significance of the interactive effect of interrogativity x tone in context 2 may be attributable to differences in effect size.

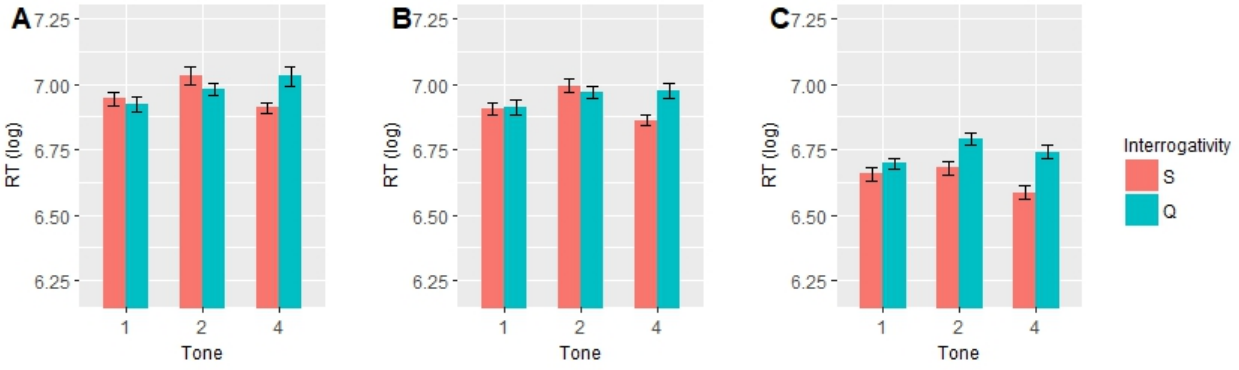

FIGURE 3.10: RTs in statements and questions in each tonal condition in (A) context 0, (B) context 1 and (C) context 2.

The results in relation to Tone2 and Tone4 show a trend whereby listeners' identification of statements and questions is influenced by the final tone at lower levels of language context. Where there is no or neutral language context, their processing speed is such that they tend to associate a final rising tone with questions, and a final falling tone with statements. However, the influence of the final tone on RT is reduced where there is a constraining language context. Statements are identified quicker than questions, regardless of the final tone.

Chapter 4

Discussion

In this thesis, we sought to illuminate the effect of language context on the perception of intonation. To this end, we conducted an experiment that investigated how Dutch and Mandarin listeners' identification of intonation as a cue of Mandarin interrogativity was influenced by the level of language context, as well as by language background and the final lexical tone identity. In this section, we firstly summarize our findings and relate them to previous research. Secondly, we attempt to account for aspects of our findings which were inconsistent with our hypotheses and previous research. Thirdly, we identify a limitation of our study relating to the construction of our stimuli.

4.1 General findings

Our results revealed that both Dutch and Mandarin listeners' identification of intonation as a cue of Mandarin interrogativity is generally enhanced by language context. This was illustrated by the significant effect of context on all listeners' ID rates and RTs. As displayed in Figure 3.1 (Part A), listeners' ID rates increase with each increasing level of context. Similarly, their RTs also become quicker with each increasing level of context. This is notwithstanding that that facilitative effect of a neutral context over a non-speech context on RT is minimal, compared to the larger facilitative effect of a constraining context over a neutral context (as seen in Figure 3.1, Part B). These facilitative effects of context on ID rates and RT were also evident in most sentence conditions of tone and interrogativity (i.e., Tone1 statements, Tone2 statements, Tone4 statements, Tone1 questions and Tone4 questions), for both Dutch and Mandarin listeners (see Figures 3.2 to 3.5).

The significant interactive effects on our listeners' ID rates and RTs revealed how the effects of language background and the final lexical tone identity influenced listeners' identification of interrogativity across different contexts. The significant interactive effect of context x interrogativity x tone x background on ID rates showed

that for both Dutch and Mandarin listeners, the influence of the final tone decreases with increasing levels of context. Where there is no language context, Dutch and Mandarin listeners identify questions better than statements were there is a final rising tone, and statements better than questions where there is a final falling tone (see Figures 3.6 and 3.7). This indicates that listeners associate rising and falling tones with questions and statements respectively where there is no language context. Conversely, listeners identified statements better than questions regardless of the final tone, where there was a constraining context. The differences between Dutch and Mandarin listeners lay in a neutral context: Dutch listeners associated rising and falling tones with questions and statements respectively, while Mandarin listeners identified statements better than questions regardless of the final tone.

In relation to RT, the significant interactive effects of context x interrogativity x background and context x interrogativity x tone are illuminating. An exploration of the interactive effect of context x interrogativity x background revealed that Mandarin listeners' responses were consistently quicker than Dutch listeners, with the difference between the two listener groups diminishing with lower levels of context (see Figure 3.9). Specifically, Mandarin listeners' responses were significantly quicker than those of Dutch listeners in a constraining context. In a neutral context, Mandarin listeners' responses were only significantly quicker in relation to questions. In a non-speech context, Mandarin listeners' responses were not significantly quicker than those of Dutch listeners in relation to either questions or statements. This result suggests that the effect of language experience in intonation perception at the auditory processing level is minimal compared to its effect at the linguistic processing level. Furthermore, the significant effect of context x interrogativity x tone on RT also revealed that listeners' processing speeds were influenced by the final tone at lower levels of context. As displayed in Figure 3.10, listeners were quicker to associate a rising tone with questions than statements, and a falling tone with statements than questions in a non-speech or neutral context. But in a constraining context, listeners identified statements quicker than questions regardless of the final tone.

Various aspects of these findings in the current study resonate with previous research. Our finding that the identification of intonation improves with increasing language context corroborates the findings of M. Liu et al. (2016a). In their experiment, they showed that Mandarin native listeners' ability to identify intonation was better in a constraining context than a neutral one. Building upon this result, our experiment revealed that the facilitative effect of language context extends over three levels (i.e., none, neutral and constraining), with the ability to identify intonation improving with each higher level of context. We also found that the facilitative effect of language context on intonation identification extended to non-native listeners.

Moreover, it can be argued that the reduced influence of the final lexical tonal identity at higher levels of context is evidence of the facilitative effect of a constraining context on the identification of intonation. In a constraining context, listeners were predisposed towards the identification of statements over questions. This result aligns with previous studies (Makarova, 2001; Yuan, 2011), confirming that the identification of questions requires more mental effort (M. Liu, Chen, & Schiller, 2016b, pp.314-5). Nevertheless, both statements and questions generally attracted higher ID rates and quicker responses in a constraining language context than in lower levels of context. This, coupled with the reduced effect of the final tone, supports the contention that listeners can perceive differences in the global pitch contour – the salient marker of Mandarin interrogativity – better in a constraining language context compared to lower levels of context.

Our findings resonate with a longer line of studies which have found that a constraining semantic context facilitates speech comprehension. It has been found, for example, that a sentential and discourse context facilitates lexical access (e.g., Marslen-Wilson & Welsh, 1978, p.61; Tabossi, 1988). An aspect of speech comprehension of immediate relevance to our study is the perception of tone. Ye and Connine (1999) found that tonal recognition by native Mandarin listeners is enhanced when perceived as the final syllable of an idiomatic expression, compared to their perception in isolation. Similarly, M. Liu et al. (2016a) found that Mandarin listeners had significantly quicker RTs when identifying tones in a constraining context compared to a neutral context. The results in these behavioural studies have also been supported by neurophysiological evidence from Kung, Chwilla, and Schriefers (2014). In their study of the identification of Cantonese tones by native listeners, they found not only higher accuracy rates in a constraining context compared to a neutral context, but also the disappearance of the P600 effect, an event-related potential reflecting the ease of language processing.

A comparison between Dutch and Mandarin listeners' results confirms the universal and language-specific effects in intonation perception as found in previous studies (e.g., Gussenhoven & Chen, 2000). The universal effects are reflected by the trend that both listener groups associated rising tones and falling tones with questions and statements respectively at lower levels of context. Furthermore, the high ID rates of both listener groups in a constraining context (see Figures 3.2 and 3.3) show that they can identify high-pitched global pitch contours as questions, low-pitched ones as statements. In a constraining context, the minimum ID rate was 63.64% (for Dutch listeners' perception of Tone4 questions). Thus both Dutch and Mandarin listeners apply the Frequency Code in identifying interrogativity across all contexts, although the acoustic cues they use differ across contexts: the final tone is used as the cue in

lower levels of context,[1] and the global pitch contour at higher levels of context.

The language-specific effect in intonation perception was reflected through Mandarin listeners' quicker RTs compared to Dutch listeners. Of particular interest is our finding that that this effect of language background extends to the auditory processing level, albeit to a minimal degree. This finding has common ground with previous studies that have also found the effect of language experience in intonation perception at the auditory processing level. A close examination of the graphical data of Ortega-Llebaria and Colantoni (2013, p.344) shows that the difference between the Mandarin and English listeners' ability to perceive focus intonation in English is smaller in the non-speech context compared to the speech context. The diminished effects of language experience in a non-speech context is also evident in the Spanish listeners, albeit to a smaller extent. The difference in English and Spanish listeners' ability to perceive focus in subject and object positions appear to be comparable in speech and non-speech contexts. However, where the focus was on the verb, the difference between the two listener groups' abilities was markedly smaller in a non-speech context, when compared to a speech context.

Moreover, the reduced effect of background in a non-speech context can be found in the experiment of B. R. Xu and Mok (2012a, 2012b). Their graphical data suggest that Mandarin and Cantonese listeners' accuracy rates in identifying Mandarin interrogativity are less divergent in a non-speech context than a speech context. In their replicating experiment (B. R. Xu & Mok, 2014), the difference between the two listener groups do appear to be similar across the two speech contexts. Nevertheless, the results from previous studies as well as our current study is weighted towards the view that the effect of background on intonation perception is diminished at the auditory processing level compared to the higher linguistic processing level.

4.2 Accounting for more complex findings

Nevertheless, some aspects of our results are not as straightforward to explain. Firstly, we discuss the finding that listeners' identification of intonation is worst where there is no language context, as this was contrary to our hypothesis. Secondly, we provide suggestions as to why some studies (including the current study) found an effect of language background on pitch perception at the auditory processing level, while other studies did not find such an effect. Thirdly, we offer an account as to why the perception

[1] However, it is hard to explain why the results in relation to Tone1 do not follow the Frequency Code in a non-speech or neutral context. Recall that listeners' ID rates are higher for statements than questions in sentences with a final high-level Tone1 (see Figures 3.6 and 3.7). Moreover, listeners' RTs do not differ significantly across statements and questions ending with this tone (see Figure 3.10).

of questions with a final rising Tone2 was seemingly immune from the effect of language context. There are grounds to contend that this is not a trivial exception.

4.2.1 Comparing the presence and absence of language context

Our results further the debate as to whether listeners would find it easier to perceive intonation in the absence or presence of language context. We had hypothesized that they would find it easier to perceive intonation in the absence of language context. This was given the established finding that segments and tone are better perceived in the absence of language context: the absence of higher frequencies is said to enable listeners to be more perceptive to differences in the fundamental frequency (e.g., Miyawaki et al., 1975; X. Luo & Ashmore, 2014). Additionally, within the limited studies examining how the absence of language context effects intonation (i.e., Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014), we had argued that the weight of evidence tended towards the view that listeners perceive intonation better without language context. However, our results were opposite to what we had hypothesized. This was especially evident in the ID rates of our listeners, and to a lesser extent, in their RTs.

The difference in the perception of intonation and tone could be explained as follows. Tone is generally restricted to the domain of a syllable, while intonation can extend beyond the domain of a syllable, being "a feature at the sentential level" (M. Liu et al., 2016b, p.309). Consequently, studies examining the effect of a non-speech context on the perception of tone have used single-syllable stimuli (e.g. X. Luo & Ashmore, 2014). Conversely, the studies on intonation have based their stimuli on multi-syllabic sentences (Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014). This means that the degree of language context in the single-syllable speech stimuli used in the tonal studies is not the same as that used in the multi-syllabic speech stimuli in the intonational studies. While the former stimuli do include segmental information, they do not contain sufficient semantic content. This is given that a single-syllabic speech syllable in tonal languages such as Mandarin could refer to more than one lexeme. Moreover, non-native listeners without knowledge of tonal languages, as employed in the studies examining the effect of speech and non-speech on tonal perception (e.g., X. Luo & Ashmore, 2014), would be unable to assign a lexical meaning to these words. Conversely, the multi-syllabic speech stimuli used in the intonational studies are meaningful sentences and thus have sufficient semantic content.

As such, it may be the case that the ability to perceive variation in pitch is hindered by segmental content alone, but facilitated by a broader semantic context. This would explain the deprecatory effects of speech context on the perception of tone on the

one hand, and its facilitative effects in the perception of intonation on the other hand. In the case of tonal perception, the presence of segmental information detracts from listeners' ability to perceive differences in F0. Conversely, the perception of intonation is enhanced by the presence of language context. This could be because the semantic context enables listeners to interpret differences in intonation in a meaningful way.

Our result that language context facilitates the perception of intonation for both native and non-native listeners resonates with other research findings investigating factors besides language background. Saindon, Trehub, Schellenberg, and van Lieshout (2016) found that both children and adults could perceive intonation as a cue of English interrogativity better in a speech context compared to in a non-speech context. In a study examining the perceived voice quality of filtered speech compared to non-filtered Dutch speech, van Bezooijen and Boves (1986) found that filtered speech attracted higher ratings than non-filtered speech on the scales of "pitch variability" and "varied pitch patterns". These studies propose that the enhanced sensitivity towards intonational differences in speech compared to non-speech can be explained by the perceived sound quality of filtered speech. van Bezooijen and Boves (1986, p.413) describe filtered speech as having a "smooth and murmur like overall quality", which can explain why it attracts lower ratings of pitch variability than unfiltered speech. Similarly, Saindon et al. (2016, p.1188) suggest that filtered speech can have an "unusual sound quality" which may detract from listeners' ability to perceive intonational differences.

Saindon et al. (2016, p.1188) raise the possibility that listeners could adapt to this unusual sound quality, citing past findings that listeners can adapt to distorted speech (including low-pass filtered speech) (e.g., Hervais-Adelman, Davis, Johnsrude, Taylor, & Carlyon, 2011). However, amongst the listeners in our current study, it was found that trial did not have a significant effect on ID rate in context 0 ($p = 0.34$).[2] RT did have a significant effect such that responses became significantly quicker over time ($p < .001$, $r = 0.11$), but the small effect size suggests that listeners did not adapt to a high degree.[3] As such, it can be said that the unusual sound quality could detract from listeners' sensitivity to intonational differences, and provide an additional reason accounting for the reduced ability to perceive intonation in a non-speech compared to a speech context.

If the absence of a semantic context and a presence of an unusual sound can account for the listeners' reduced ability to attend to intonation in a non-speech context, then the obvious question arises as to why results in previous studies revealed instead

[2]The significant effect of trial on ID rate (which we found earlier) appears to be limited to contexts 1 and 2 (both ps < 0.01).

[3]This does not discount the possibility that listeners' ability to perceive intonation without a speech context would be better after a training session, as raised by Saindon et al. (2016, p.1188).

an enhanced ability to perceive intonation in non-speech over speech contexts. These differing results could be explained by differences in task. The listeners in the study of Ortega-Llebaria and Colantoni (2013) were tasked with matching one of three filtered stimuli differing in intonation with an unfiltered sentence they had heard immediately prior. Their inclusion of an initial unfiltered sentence may have induced listeners to process the subsequent filtered stimuli at a linguistic level. Moreover, in B. R. Xu and Mok (2012a), listeners could listen to the non-speech stimuli repeatedly. This could be a factor explaining Cantonese listeners' superior ability in a non-speech context over a speech context.

4.2.2 Effect of background in a non-speech context

Our study, as well as other studies (Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014), have found an effect of language background in intonation perception at the auditory processing level. The question then remains as to why other studies have found an absence of an effect of language experience in tone perception (e.g., Burnham et al., 1996) and intonation perception (Grabe et al., 2003) at the auditory processing level.

Bent et al. (2006, p.103) propose that the absence or presence of language experience in non-speech tone perception is task-dependent in some way. Similarly, we contend that the studies which found an effect of language background in intonation perception at the acoustic processing level adopted an experimental method that activated listeners' linguistic processing mechanisms. Ortega-Llebaria and Colantoni (2013) adopted such a method, which we detailed in subsection 4.2.1. In B. R. Xu and Mok (2014), listeners were exposed to speech and non-speech stimuli mixed in the same blocks. This inclusion of speech stimuli as part of these tasks may have encouraged users to think of the non-speech stimuli as speech. Moreover, in our own study, the task of identifying filtered sentences as statements or questions, which are linguistic phenomena in themselves, may also have tapped into listeners' linguistic processing mechanisms. In contrast, Grabe et al. (2003) examined the perception of intonation in non-speech stimuli by having subjects listen to two non-speech stimuli and assign a rating of the degree of similarity. The researchers' adoption of a pitch discrimination task, as well as their recruitment of a separate group of subjects for the speech stimuli, would have tapped more into listeners' auditory processing mechanisms compared to our intonation identification task.

Additionally, Bent et al. (2006, p.98) suggest that linguistic processing mechanisms may be activated by "nonspeech stimuli whose spectral... properties bear some resemblance to speech stimuli". In this respect, the non-speech stimuli used in Ortega-Llebaria and Colantoni (2013); B. R. Xu and Mok (2012b, 2014), and our own

study were naturally produced and low-pass filtered speech. Conversely, Grabe et al. (2003) used frequency-modulated sine waves that were artificially generated. Arguably, the acoustic properties of the stimuli in Grabe et al. (2003) have less resemblance to speech stimuli, compared to those of the stimuli used in the former studies. This could also explain the absence of the effect of language background in intonation perception at the auditory level in Grabe et al. (2003). Thus the stimuli type, as well as the task type, could account for the differential effects of language background across different studies.

4.2.3 Tone2 questions: an exceptional exception?

A notable exception to our finding that the perception of interrogativity in Mandarin improves with increasing levels of context was found in the ID rate of Tone2 questions. In this sentence condition, Dutch listeners' ID rate in a neutral context was significantly lower than in a non-speech context. Their ID rate in a constraining context were not significantly different from that in either of the two lower levels of context (see Figure 3.2, Part B). For Mandarin listeners, context did not have a significant effect (see 3.3, Part B). The lack of a facilitative effect of context in the ID rate of Tone2 questions was also found by M. Liu et al. (2016a), whose Mandarin listeners identified Tone2 questions worse in a constraining context than in a neutral context.

A possible explanation for this exception derives from the idea that the phonological identity of the final tone influences the identification of Mandarin interrogativity, as proposed by Yuan (2011). Simply put, this means that it is easier to discern intonation in sentences with certain final tones, and harder to do so in sentences ending with other tones. In this regard, several neurophysiological studies support the idea that intonation as a cue of interrogativity is better perceived with a final Tone4 than a final Tone2. In Ren, Yang, and Li (2009), Mandarin listeners displayed a mismatch-negativity (MMN) effect (an indicator that an acoustic change was detected) when exposed to single-syllable questions and statements with Tone4, at the pre-attentive stage (that is, their attention was diverted to another stimulus). However, the MMN effect was not found when they were exposed to single-syllable questions and statements ending with Tone2 in a subsequent study (Ren, Tang, Li, & Sui, 2013).

A similar finding emerged in M. Liu et al. (2016b), who found a P300 effect (an indicator of the ease of language processing) in Mandarin listeners attentively listening to multi-syllabic statements and questions of a neutral context ending in Tone4, but not in relation to those ending in Tone2. Granted, these neurophysiological results are not congruent with the behavioral data of our current study. Dutch and Mandarin listeners showed difficulties in identifying interrogativity in sentences ending with both

Tone2 and Tone4, especially at lower levels of language context (see Figures 3.2 and 3.3). For example, Dutch listeners' ID rates of Tone2 statements and Tone4 questions in a non-speech context were at a below-chance level, being 32.49% and 39.58% respectively. Regardless, the neurophysiological data do lend support to the idea that differences in the perception of Mandarin intonation can be attributed to the differences in the phonological identity of Tone2 and Tone4. In the case of our current study, we can argue that the facilitative effect of the context on the perception of interrogativity is tempered by the phonological identity of the final tone. Specifically, increasing levels of context facilitate the identification of questions ending with a final Tone1 or Tone4, but do not facilitate the identification of questions ending with Tone2 to the same degree.

4.3 Limitation in stimuli construction

One limitation to our current study relates to the way in which our stimuli of neutral and constraining language context were constructed. Our sentences of neutral context all consist of four initial high tones, while the sentences of constraining context consist of differing tones. Shih (1988, pp.90-1) had found that Mandarin statements consisting of only level tones were less susceptible than those consisting of varied tones to downstep: the gradual lowering of F0 throughout the course of a statement. As such, the difference in the global F0 contour between statements and questions would be expected to be smaller in our sentences of neutral context, than in those of constraining context. This is confirmed by an impressionistic acoustic analysis. Figure 4.1 below is a boxplot graph showing the difference in F0 (in Hz) between statements and questions in each of the last five syllables. The difference in F0 is generally greater in sentences of constraining context compared to those of a neutral context, with the difference between the two contexts appearing to be greatest in the final syllable.

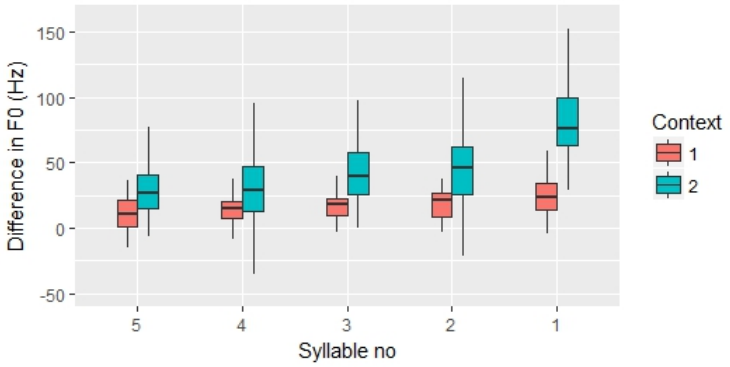

FIGURE 4.1: Boxplots showing the difference in F0 between statements and questions for each syllable no, in both sentences of context 1 and 2. "Syllable no" is defined such that "1" denotes the first syllable from the end of the sentence.

Consequently, the enhanced ability of listeners to perceive intonation in the presence of a constraining context could be explained not only by the higher degree of semantic constraint, but also the larger acoustic differences between statements and questions. Moreover, the comparatively smaller acoustic difference in a neutral context could explain why listeners based their judgment of interrogativity on the final pitch movement in a neutral context more so than a constraining context. In hindsight, to minimize the differences in F0 between the neutral and constraining contexts, it may have been better to construct our sentences of constraining context using the same four initial high-tone syllables as the neutral context, followed by the same disyllabic words we had used as the basis for our constraining context (per Kung et al., 2014). As such, there is a need to be cautious in overstating the effect of a constraining context in our current study. This is especially given our finding that the facilitative effect of a constraining context over a neutral context on listeners' RTs is greater than the minimal facilitative effect of a neutral context over a non-speech context. Had we used Kung et al. (2014)'s method of constructing stimuli, the facilitative effects of a constraining context and a neutral context on listeners' RTs may have possibly been more comparable.

Conclusion

Our motivation for the current study stemmed from emerging tensions in the literature on how language context affects listeners' ability to perceive intonation. Past studies have proposed different views on whether intonation is better perceived in a non-speech context compared to a speech context. On the one hand, past studies on this question have lent weight to the view that intonation is better perceived without language context (Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014). This view follows a previous line of studies demonstrating the perception of segments and tones is more refined at the auditory processing level than the linguistic processing level (e.g., Miyawaki et al., 1975; X. Luo & Ashmore, 2014). On the other hand, M. Liu et al. (2016a) raised the possibility that the perception of intonation improves with increasing language context, with the ability to perceive intonation lowest in filtered speech, higher in a neutral language context, and highest in a constraining language context. This was following their finding that a constraining context facilitated intonation perception over a neutral context. Moreover, different views have also emerged as to whether the effects of language background in intonation perception extend to a non-speech context (Grabe et al., 2003; Ortega-Llebaria & Colantoni, 2013; B. R. Xu & Mok, 2012a, 2012b, 2014).

To elucidate these tensions and extend our understanding on how language context effects the perception of intonation in both native and non-native listeners, we tested their ability to perceive intonation in Mandarin at three levels of language context: none, neutral and constraining. Our main research question was how Dutch listeners' identification of intonational cues of interrogativity in Mandarin differ according to different levels of language context. We also examined the effect of language background and the final lexical tonal identity on their identification patterns.

Our results, in terms of ID rates and RTs, revealed that both Mandarin and Dutch listeners generally identified interrogativity better with each increasing level of context. An examination of the influence of the final tone revealed a tendency for both Dutch and Mandarin listeners to associate final rising and falling tones with questions and statements respectively at lower levels of language context. In a constraining context, they based their judgment more so on differences in the global pitch contour –

the salient marker of Mandarin interrogativity, associating high and low pitch contours with questions and statements respectively. As such, both listener groups relied on the universal Frequency Code to identify intonation, although the cue that they used differed across language contexts. The effect of language background on intonation perception was revealed through Mandarin listeners' quicker RTs compared to those of Dutch listeners. However, the differences between the two groups' RTs decreased with lower levels of language context, being minimal in a non-speech context.

Some aspects of our results confirm findings in past research. Our finding of a facilitative effect of a constraining context in intonation identification resonates with past findings on the facilitative effect of a constraining context on pitch processing (e.g., Ye & Connine, 1999). Moreover, our results demonstrate that both universal and language-specific factors influence the perception of intonation, as established in previous studies (e.g., Gussenhoven & Chen, 2000). Additionally, the reduced effect of language background in intonation perception at the auditory processing level is in line with the findings of Ortega-Llebaria and Colantoni (2013) and B. R. Xu and Mok (2012a, 2012b, 2014).

Importantly, our findings make a step in elucidating conflicting views in the existing research on intonation perception. Our study illuminated the differential effects of the presence and absence of language context on the perception of intonation. The identification of intonation was found to be stronger where there is language context. This could be explained by the idea that the perception of acoustic features in speech is enhanced by the presence of a meaningful semantic context. Additionally, our study suggests that the effect of language background in intonation perception decreases with decreasing language context. Nevertheless, how do we reconcile our findings with previous findings that intonation perception is stronger in the absence of language context (e.g., Ortega-Llebaria & Colantoni, 2013), and not influenced by language background at the auditory processing level (Grabe et al., 2003)? The answer may lie in differences in task and stimuli. The interplay of the effect of these additional factors with the effect of context could be scrutinized in future research that seeks to refine our knowledge on the perception of intonation.

Appendix A

Sentences of neutral language context

The words which form the basis for the sentences of neutral language context are tabulated in the following page, grouped by tone. Each word is listed as a Chinese character, followed by its pinyin (the standard romanization of Chinese characters). The number after each pinyin syllable represents its tone. Each of these words was embedded as the final word of the sentence "他刚刚说 X" ("ta1 gang1 gang1 shuo1 X" – "He just said X").

Tone1			**Tone2**			**Tone4**		
帮	bang1	help	白	bai2	white	菜	cai4	vegetable
包	bao1	package	鼻	bi2	nose	错	cuo4	wrong
车	che1	car	船	chuan2	ship	大	da4	big
吃	chi1	eat	床	chuang2	bed	放	fang4	release
吹	chui1	blow	肥	fei2	fat	怪	guai4	strange
春	chun1	spring	国	guo2	country	贵	gui4	expensive
刀	dao1	knife	来	lai2	come	坏	huai4	bad
光	guang1	light	龙	long2	dragon	看	kan4	look
黑	hei1	black	楼	lou2	floor	妹	mei4	sister
花	hua1	flower	门	men2	door	梦	meng4	dream
空	kong1	air	年	nian2	year	面	mian4	surface
哭	ku1	cry	爬	pa2	climb	怕	pa4	afraid
拉	la1	pull	皮	pi2	skin	热	re4	hot
猫	mao1	cat	婆	po2	grandma	日	ri4	day
秋	qiu1	autumn	球	qiu2	ball	上	shang4	on
刷	shua1	brush	人	ren2	person	跳	tiao4	jump
酸	suan1	sour	头	tou2	head	外	wai4	outside
天	tian1	day	学	xue2	learn	问	wen4	ask
听	ting1	listen	牙	ya2	tooth	信	xin4	letter
租	zu1	rent	爷	ye2	grandpa	用	yong4	use

Appendix B

Sentences of constraining language context

The sentences of constraining language context used in the experiment are listed here in order of tone. The sentences are listed in Chinese characters and pinyin, along with an English translation. The disyllabic words, on the basis of which the sentences were created, are bolded.

Tone1

他 要 吃 一 片 **面包** Ta1 yao4 chi1 yi1 pian4 **mian4 bao1**	He wants to eat a piece of **bread**
中 国 人 不 吃 **西餐** Zhong1 guo2 ren2 bu4 chi1 **xi1 can1**	Chinese people do not eat **Western food**
他 在 这 一 站 **下车** Ta1 zai4 zhe4 yi2 zhan4 **xia4 che1**	He will **get off** at this station
他 想 要 看 **菜单** Ta1 xiang3 yao4 kan4 **cai4 dan1**	He wants to see **the menu**
他 买 了 一 包 **饼干** Ta1 mai3 le yi1 bao1 **bing3 gan1**	He bought a bag of **cookies**
他 想 听 我 **唱歌** Ta1 xiang3 ting1 wo3 **chang4 ge1**	He wants to hear me **sing**
他 们 刚 刚 **结婚** Ta1 men gang1 gang1 **jie2 hun1**	They just **got married**
他 很 害 怕 坐 **飞机** Ta1 hen3 hai4 pa4 zuo4 **fei1 ji1**	He is afraid of going on a **plane**
他 吃 了 一 个 **香蕉** Ta1 chi1 le yi1 ge4 **xiang1 jiao1**	He ate a **banana**
他 到 乡 下 **爬山** Ta1 dao4 xiang1 xia4 **pa2 shan1**	He is heading to the countryside **to hike**
他 今 天 有 点 **发烧** Ta1 jin1 tian1 you3 dian3 **fa1 shao1**	He has a bit of **a fever** today

他	是	最	聪	明	的	**学生**	He is the smartest **student**
Ta1	shi4	zui4	cong1	ming2	de	**xue2 sheng1**	

他	喜	欢	看	**小说**		He likes to read **novels**
Ta1	xi3	huan1	kan4	**xiao3 shuo1**		

老	人	应	该	坐	**电梯**	Elderly people should take **the elevator**
Lao3	ren2	ying1	gai1	zuo4	**dian4 ti1**	

他	跟	朋	友	**聊天**	He is **chatting** with some friends
Ta1	gen1	peng2	you3	**liao2 tian1**	

这	里	有	中	国	**餐厅**	There is a Chinese **restaurant** here
Zhe4	li	you3	zhong4	guo2	**can1 ting1**	

他	喝	了	一	杯	**果汁**	He drank a glass of **juice**
Ta1	he1	le	yi1	bei1	**guo3 zhi1**	

他	穿	着	黑	色	**西装**	He is wearing a black **suit**
Ta1	chuan1	zhuo2	hei1	se4	**xi1 zhuang1**	

房	间	里	有	一	张	**书桌**	The room has **a desk**
Fang2	jian1	li	you3	yi1	zhang1	**shu1 zhuo1**	

他	没	有	钱	付	**房租**	He has no money to pay **rent**
Ta1	mei2	you3	qian2	fu4	**fang2 zu1**	

Tone2

他	杀	了	两	个	**警察**		He killed two **policemen**
Ta1	sha1	le	liang3	ge4	**jing3 cha2**		

他	通	常	很	早	**起床**	He usually **gets up** very early
Ta1	tong1	chang2	hen3	zao3	**qi3 chuang2**	

他	用	字	典	查	**生词**	He uses a dictionary to search for **new words**
Ta1	yong4	zi4	dian3	cha2	**sheng1 ci2**	

饭	后	必	须	收	拾	**厨房**	After meals you must clean **the kitchen**
Fan4	hou4	bi4	xu1	shou1	shi2	**chu2 fang2**	

他	做	运	动	来	**减肥**	He does exercise **to lose weight**
Ta1	zuo4	yun4	dong4	lai2	**jian3 fei2**	

他	们	去	过	**中国**	They have been to **China** before
Ta1	men	qu4	guo4	**zhong1 guo2**	

他	去	附	近	的	**银行**	He is going to a nearby **bank**
Ta1	qu4	fu4	jin4	de	**yin2 hang2**	

他	过	着	开	心	的	**生活**	He is leading a happy **life**
Ta1	guo4	zhe	kai1	xin1	de	**sheng1 huo2**	

问	题	还	没	有	**解决**	The problem still has not been **solved**
Wen4	ti2	hai2	mei2	you3	**jie3 jue2**	

他	在	音	乐	会	弹	**钢琴**	He will play **the piano** at the concert
Ta1	zai4	yin1	yue4	hui4	tan2	**gang1 qin2**	

他	对	学	习	很	有	**热情**	He is very **enthusiastic** about studying
Ta1	dui4	xue2	xi2	hen3	you3	**re4 qing2**	

他	们	在	踢	**足球**	They are playing **football**
Ta1	men	zai4	ti1	**zu2 qiu2**	

他 不 想 麻 烦 **别人** He does not want to trouble **others**
Ta1 bu4 xiang3 ma2 fan **bie2 ren2**

他 让 妈 妈 **头疼** He gave mom a **headache**
Ta1 rang4 ma1 ma **tou2 teng2**

他 煮 了 一 碗 **面条** He cooked a bowl of **noodles**
Ta1 zhu3 le yi1 wan3 **mian4 tiao2**

游 客 在 看 **地图** The tourists are looking at **a map**
You2 ke4 zai4 kan4 **di4 tu2**

音 乐 会 明 天 **举行** The concert will be **held** tomorrow
Yin1 yue4 hui4 ming2 tian1 **ju3 xing2**

妈 妈 教 我 **数学** Mom teaches me **math**
Ma1 ma jiao4 wo3 **shu4 xue2**

他 睡 觉 前 要 **刷牙** He needs **to brush his teeth** before going to bed
Ta1 shui4 jiao4 qian2 yao4 **shua1 ya2**

他 们 开 车 去 **公园** They drove to the **park**
Ta1 men kai1 che1 qu4 **gong1 yuan2**

Tone4

他 在 公 园 里 **跑步** He is **jogging** in the park
Ta1 zai4 gong1 yuan2 li3 **pao3 bu4**

他 早 餐 吃 了 **鸡蛋** He ate **eggs** for breakfast
Ta1 zao3 can1 chi1 le **ji1 dan4**

病 人 常 常 去 **药店** Patients often go to **the pharmacy**
Bing4 ren2 chang2 chang2 qu4 **yao4 dian4**

他 不 同 意 我 的 **决定** He disagreed with my **decision**
Ta1 bu4 tong2 yi4 wo3 de **jue2 ding4**

他 每 餐 吃 一 碗 **米饭** He eats a bowl of **rice** each meal
Ta1 mei3 can1 chi1 yi1 wan3 **mi3 fan4**

他 不 同 意 我 的 **计划** He does not agree with my **plan**
Ta1 bu4 tong2 yi4 wo3 de **ji4 hua4**

他 能 改 变 **世界** He can change **the world**
Ta1 neng2 gai3 bian4 **shi4 jie4**

房 子 里 有 很 多 **家具** There is a lot of **furniture** in the house
Fang2 zi li you3 hen3 duo1 **jia1 ju4**

他 很 晚 才 **睡觉** He **sleeps** only when it is very late
Ta1 hen3 wan3 cai2 **shui4 jiao4**

他 要 喝 一 瓶 **可乐** He wants to drink a bottle of **coca-cola**
Ta1 yao4 he4 yi1 ping2 **ke3 le4**

他 让 爸 爸 很 **生气** He made dad **angry**
Ta1 rang4 ba4 ba hen3 **sheng1 qi4**

他 参 加 足 球 **比赛** He participated in a football **match**
Ta1 can1 jia1 zu2 qiu2 **bi3 sai4**

附 近 有 一 家 **超市** There is a **supermarket** nearby
Fu4 jin4 you3 yi1 jia1 **chao1 shi4**

表 演 还 没 有 **结束**	The show is not **over** yet
Biao3 yan3 hai2 mei2 you3 **jie2 shu4**	
他 想 去 商 场 **购物**	He wants to go **shopping** at the mall
Ta1 xiang3 qu4 shang1 chang3 **gou4 wu4**	
爸 爸 送 你 到 **学校**	Dad sent you to **school**
Ba4 ba song4 ni3 dao4 **xue2 xiao4**	
学 生 有 太 多 **作业**	The students have too many **assignments**
Xue2 sheng1 you3 tai4 duo1 **zuo4 ye4**	
他 带 病 人 去 **医院**	He took the patient to **the hospital**
Ta1 dai4 bing4 ren2 qu4 **yi1 yuan4**	
他 喜 欢 听 **音乐**	He likes to listen to **music**
Ta1 xi3 huan1 ting1 **yin1 yue4**	
他 有 中 国 **护照**	He has a Chinese **passport**
Ta1 you3 zhong4 guo2 **hu4 zhao4**	

References

Altman, D. G., Deeks, J. J., & Sackett, D. L. (1998). Odds ratios should be avoided when events are common. *British Medical Journal*, *317*(7170), 1456-1456.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge, U.K.: Cambridge University Press.

Bartoń, K. (2018). MuMIn: Multi-model inference. R package version 1.40.4 [Computer software manual]. Retrieved 2018-03-15, from https://cran.r-project.org/web/packages/MuMIn/MuMIn.pdf

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2013). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–9 [Computer software manual]. Retrieved 2017-11-20, from https://cran.r-project.org/web/packages/lme4/lme4.pdf

Bent, T., Bradlow, A. R., & Wright, B. A. (2006). The influence of linguistic experience on the cognitive processing of pitch in speech and nonspeech sounds. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 97–103. doi: 10.1037/0096-1523.32.1.97

Best, C. T., & Tyler, M. (2007). Nonnative and second-language speech perception: Commonalities and complementaries. In M. J. Munro & O.-S. Bohn (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 13–34). Amsterdam: John Benjamins.

Beun, R.-J. (1990). The recognition of Dutch declarative questions. *Journal of Pragmatics*, *14*(1), 39–56. doi: 10.1016/0378-2166(90)90063-J

Boersma, P., & Weenink, D. (2017). *Praat: Doing phonetics by computer [computer program]. version 6.0.34*. Retrieved 2018-01-08, from http://www.praat.org/

Bolinger, D. (1978). Intonation across languages. In J. Greenberg (Ed.), *Universals of human language* (Vol. 2: Phonology, pp. 471–524). Palo Alto, U.S.A.: Stanford University Press.

Braun, B., & Johnson, E. K. (2011). Question or tone 2? How language experience and linguistic function guide pitch processing. *Journal of Phonetics*, *39*(4), 585–594. doi: 10.1016/j.wocn.2011.06.002

Burnham, D., Francis, E., Webster, D., Luksaneeyanawin, S., Attapaiboon, C., Lacerda,

F., & Keller, P. (1996). Perception of lexical tone across languages: Evidence for a linguistic mode of processing. In *Proceedings of the Fourth International Conference on Spoken Language Processing* (Vol. 4, pp. 2514–2517). doi: 10.1109/ICSLP.1996.607325

Burns, E. M., & Sampat, K. (1980). A note on possible culture-bound effects in frequency discrimination. *The Journal of the Acoustical Society of America*, *68*(6), 1886–1888. doi: 10.1121/1.385179

Carlson, K. (2009). How prosody influences sentence comprehension. *Language and Linguistics Compass*, *3*(5), 1188–1200.

Chen, A. (2005). *Universal and language-specific perception of paralinguistic intonational meaning* (Doctoral Dissertation). Radboud University.

Chen, H.-C., Vaid, J., & Wu, J.-T. (2009). Homophone density and phonological frequency in Chinese word recognition. *Language and Cognitive Processes*, *24*(7-8), 967–982. doi: 10.1080/01690960902804515

Da, J. (2004). A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. In *Proceedings of the Fourth International Conference on New Technologies in Teaching and Learning Chinese* (pp. 501–511). Beijing: Tsinghua University Press. Retrieved from http://lingua.mtsu.edu/academic/dajun-4thtech-A4.pdf

Davies, H. T. O., Crombie, I. K., & Tavakoli, M. (1998). When can odds ratios mislead? *British Medical Journal*, *316*(7136), 989–991.

Duanmu, S. (2007). *The phonology of Standard Chinese* (2nd ed.). Oxford: Oxford University Press.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *The Journal of the Acoustical Society of America*, *80*(2), 402–415.

E-prime (Computer software. Version 2.0.8.22 ed.). (n.d.). Pittsburgh, U.S.A.: Psychology Software Tools.

Gottfried, T. L., & Suiter, T. L. (1997). Effect of linguistic experience on the identification of Mandarin Chinese vowels and tones. *Journal of Phonetics*, *25*(2), 207–231. doi: 10.1006/jpho.1997.0042

Grabe, E., Rosner, B. S., García-Albea, J. E., & Zhou, X. (2003). Perception of English intonation by English, Spanish, and Chinese listeners. *Language and Speech*, *46*(4), 375–401. doi: 10.1177/00238309030460040201

Gussenhoven, C. (2002). Intonation and interpretation: Phonetics and phonology. In *Speech Prosody 2002* (pp. 47–57). Retrieved from https://www.isca-speech.org/archive_open/sp2002/sp02_047.pdf

Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge, U.K.: Cambridge University Press.

Gussenhoven, C., & Chen, A. (2000). Universal and language-specific effects in the perception of question intonation. In *Sixth International Conference on Spoken Language Processing* (Vol. 2, pp. 91–94). Retrieved from https://www.isca-speech.org/archive/archive_papers/icslp_2000/i00_2091.pdf

Haan, J. (2002). *Speaking of questions: An exploration of Dutch question intonation* (Doctoral Dissertation). Radboud University.

Haan, J., van Heuven, V. J., Pacilly, J. J. A., & van Bezooijen, R. (1997). On the anatomy of Dutch question intonation. In H. de Hoop & J. Coerts (Eds.), *Linguistics in the Netherlands 1997* (pp. 99–110). Amsterdam: John Benjamins.

Hadding-Koch, K., & Studdert-Kennedy, M. (1964). An experimental study of some intonation contours. *Phonetica*, *11*(3-4), 175–185. doi: 10.1159/000258338

Hermann, E. (1942). *Probleme der Frage. [Problems of the question].* Göttingen: Vandenhoeck Ruprecht.

Hervais-Adelman, A. G., Davis, M. H., Johnsrude, I. S., Taylor, K. J., & Carlyon, R. P. (2011). Generalization of perceptual learning of vocoded speech. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(1), 283. doi: 10.1037/a0020772

Heuven, V. J., & Haan, J. (2002). Temporal distribution of interrogativity markers in Dutch: A perceptual study. In C. Gussenhoven & N. Warner (Eds.), *Laboratory phonology* (Vol. 7, p. 61-86). Berlin: De Gruyter. doi: 10.1515/9783110197105.1.61

Huang, T., & Johnson, K. (2010). Language specificity in speech perception: Perception of Mandarin tones by native and nonnative listeners. *Phonetica*, *67*(4), 243–267. doi: 10.1159/000327392

Kung, C., Chwilla, D. J., & Schriefers, H. (2014). The interaction of lexical tone, intonation and semantic context in on-line spoken word recognition: An ERP study on Cantonese Chinese. *Neuropsychologia*, *53*, 293–309. doi: 10.1016/j.neuropsychologia.2013.11.020

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmertest: Tests in linear mixed effects models. R package version 2.0–34 [Computer software manual]. Retrieved 2017-11-20, from https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf

Lee, C.-Y., Tao, L., & Bond, Z. S. (2009). Speaker variability and context in the identification of fragmented Mandarin tones by native and non-native listeners. *Journal of Phonetics*, *37*(1), 1–15. doi: 10.1016/j.wocn.2008.08.001

Lin, M. (2004). On production and perception of boundary tone in Chinese intonation. In *International symposium on tonal aspects of languages: With emphasis on tone languages.* Retrieved from https://www.isca-speech.org/archive_open/tal2004/tal4_125.pdf

Lin, Y.-H. (2007). *The sounds of Chinese.* Cambridge, U.K.: Cambridge University Press.

Liu, F., & Xu, Y. (2005). Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, *62*(2-4), 70–87. doi: 10.1159/000090090

Liu, M., Chen, Y., & Schiller, N. O. (2016a). Context effects on tone and intonation processing in Mandarin. In *Speech Prosody 2016.* doi: 10.21437/SpeechProsody .2016-217

Liu, M., Chen, Y., & Schiller, N. O. (2016b). Online processing of tone and intonation in Mandarin: Evidence from ERPs. *Neuropsychologia*, *91*, 307–317. doi: 10.1016/ j.neuropsychologia.2016.08.025

Luo, S., & Lin, H. (2015). English learners' perception and production of Mandarin intonation. In *Proceedings of the Eighteenth International Congress of Phonetic Sciences.* Retrieved from https://www.internationalphoneticassociation .org/icphs-proceedings/ICPhS2015/Papers/ICPHS1035.pdf

Luo, X., & Ashmore, K. B. (2014). The effect of language experience on perceptual normalization of Mandarin tones and non-speech pitch contours. *The Journal of the Acoustical Society of America*, *135*(6), 3585–3593. doi: 10.1121/1.4874619

Ma, J. K.-Y., Ciocca, V., & Whitehill, T. L. (2011). The perception of intonation questions and statements in Cantonese. *The Journal of the Acoustical Society of America*, *129*(2), 1012–1023. doi: 10.1121/1.3531840

Makarova, V. (2001). Perceptual correlates of sentence-type intonation in Russian and Japanese. *Journal of Phonetics*, *29*(2), 137–154. doi: 10.1006/jpho.2001.0137

Marslen-Wilson, W. D., & Welsh, A. (1978). Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology*, *10*(1), 29–63. doi: 10.1016/0010-0285(78)90018-X

Miyawaki, K., Jenkins, J. J., Strange, W., Liberman, A. M., Verbrugge, R., & Fujimura, O. (1975). An effect of linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics*, *18*(5), 331–340. doi: 10.3758/BF03211209

Norman, J. (1988). *Chinese.* Cambridge, U.K.: Cambridge University Press.

Ohala, J. J. (1983). Cross-language use of pitch: An ethological view. *Phonetica*, *40*(1), 1–18. doi: 10.1159/000261678

Ortega-Llebaria, M., & Colantoni, L. (2013). The L2 acquisition of English intonation: Relations between form-meaning associations, access to meaning and L1 transfer. *Studies in Second Language Acquisition*, *36*(2), 331–353. doi: 10.1017/S0272263114000011

Pisoni, D. B., Lively, S. E., & Logan, J. S. (1994). Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception. In J. Goodman

& H. Nusbaum (Eds.), *The development of speech perception* (pp. 121–166). Cambridge, U.S.A.: MIT Press.

Qin, Z., & Mok, P. (2012). The perception of speech and non-speech tones by tone and non-tone language listeners. In *Speech Prosody 2012* (pp. 366–369). Retrieved from https://www.isca-speech.org/archive/sp2012/papers/sp12_366.pdf

R Core Team. (2013). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved 2017-11-20, from http://www.R-project.org/

Ren, G.-Q., Tang, Y.-Y., Li, X.-Q., & Sui, X. (2013). Pre-attentive processing of Mandarin tone and intonation: evidence from event-related potentials. In F. Signorelli (Ed.), *Functional brain mapping and the endeavor to understand the working brain* (pp. 95–108). Croatia: IntechOpen.

Ren, G.-Q., Yang, Y., & Li, X. (2009). Early cortical processing of linguistic pitch patterns as revealed by the mismatch negativity. *Neuroscience*, *162*(1), 87–95. doi: 10.1016/j.neuroscience.2009.04.021

Rialland, A. (2009). The African lax question prosody: Its realisation and geographical distribution. *Lingua*, *119*(6), 928–949. doi: 10.1016/j.lingua.2007.09.014

Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Newbury Park, U.S.A.: Sage Publicatons.

Saindon, M. R., Trehub, S. E., Schellenberg, E. G., & van Lieshout, P. (2016). Children's identification of questions from rising terminal pitch. *Journal of Child Language*, *43*(5), 1174–1191. doi: 10.1017/S0305000915000458

Schneider, K., Dogil, G., & Möbius, B. (2011). Reaction time and decision difficulty in the perception of intonation. In *Interspeech 2011* (pp. 2221–4). Retrieved from https://www.isca-speech.org/archive/archive_papers/interspeech_2011/i11_2221.pdf

Sebastián-Gallés, N. (2005). Cross-language speech perception. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 546–566). Oxford, U.K.: Blackwell Publishing.

Shackleton, T. M., & Carlyon, R. P. (1994). The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *The Journal of the Acoustical Society of America*, *95*(6), 3529–3540. doi: 10.1121/1.409970

Shen, X.-n. S. (1990). *The prosody of Mandarin Chinese*. Berkeley, U.S.A.: University of California Press.

Shih, C.-l. (1988). Tone and intonation in Mandarin. *Working Papers of the Cornell Phonetics Laboratory*, *3*, 83–109.

Tabossi, P. (1988). Accessing lexical ambiguity in different types of sentential contexts. *Journal of Memory and Language*, *27*(3), 324–340. doi: 10.1016/0749-596X(88)

90058-7

Vaissière, J. (2005). The perception of intonation. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 236–263). Oxford, U.K.: Blackwell Publishing.

van Bezooijen, R., & Boves, L. (1986). The effects of low-pass filtering and random splicing on the perception of speech. *Journal of Psycholinguistic Research*, *15*(5), 403–417. doi: 10.1007/BF01067722

van Heuven, V. J., & Haan, J. (2000). Phonetic correlates of statement versus question intonation in Dutch. In A. Botinis (Ed.), *Intonation: Analysis, modelling and technology* (pp. 119–143). Dordrecht: Springer Science.

Wennerstrom, A. (2001). *The music of everyday speech: Prosody and discourse analysis.* Oxford: Oxford University Press.

Werker, J. F., & Logan, J. S. (1985). Cross-language evidence for three factors in speech perception. *Perception & Psychophysics*, *37*(1), 35–44. doi: 10.3758/BF03207136

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *6*, 49–63. doi: 10.1016/S0163-6383(84)80022-3

Xu, B. R., & Mok, P. (2012a). Cross-linguistic perception of intonation by Mandarin and Cantonese listeners. In *Speech Prosody 2012.* Retrieved from http://www.isca-speech.org/archive/sp2012/papers/sp12_099.pdf

Xu, B. R., & Mok, P. (2012b). Intonation perception of low-pass filtered speech in Mandarin and Cantonese. In *Tonal Aspects of Languages - Third International Symposium.* Retrieved from http://www.isca-speech.org/archive/tal_2012/papers/tl12_02-02.pdf

Xu, B. R., & Mok, P. (2014). Cross-linguistic perception of Mandarin intonation. In *Speech Prosody 2014* (pp. 638–642). Retrieved from http://www.cuhk.edu.hk/lin/people/peggy/XuMok_SP2014.pdf

Xu, Y. (2004). Separation of functional components of tone and intonation from observed F0 patterns. In G. Fant, H. Fujisaki, J. Cao, & Y. Xu (Eds.), *From traditional phonology to modern speech processing: Festschrift for Professor Wu Zongji's 95th birthday* (p. 483—505). Beijing: Foreign Language Teaching and Research Press.

Xu, Y., Gandour, J. T., & Francis, A. L. (2006). Effects of language experience and stimulus complexity on the categorical perception of pitch direction. *The Journal of the Acoustical Society of America*, *120*(2), 1063–1074. doi: 10.1121/1.2213572

Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, *33*(4), 319–337. doi: 10.1016/S0167-6393(00)00063-7

Yang, C., & Chan, M. K. M. (2010). The perception of Mandarin Chinese tones and intonation by American learners. *Journal of Chinese Language Teachers*

Association, *45*(1), 7–36.

Ye, Y., & Connine, C. M. (1999). Processing spoken Chinese: The role of tone information. *Language and Cognitive Processes*, *14*(5-6), 609–630. doi: 10.1080/016909699386202

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

Yuan, J. (2006). Mechanisms of question intonation in Mandarin. In Q. Huo, B. Ma, E.-S. Chng, & H. Li (Eds.), *Proceedings of the Fifth International Symposium on Chinese Spoken Language Processing* (pp. 19–30). Berlin: Springer. doi: 10.1007/11939993

Yuan, J. (2011). Perception of intonation in Mandarin Chinese. *The Journal of the Acoustical Society of America*, *130*(6), 4063–4069. doi: 10.1121/1.3651818

Yuan, J., & Shih, C. (2004). Confusability of Chinese intonation. In *Speech Prosody 2004*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.227.8574&rep=rep1&type=pdf

Yuan, J., Shih, C., & Kochanski, G. P. (2002). Comparison of declarative and interrogative intonation in Chinese. In *Speech Prosody 2002*. Retrieved from https://www.isca-speech.org/archive_open/sp2002/sp02_711.pdf