國立臺灣大學文學院翻譯碩士學位學程

碩士論文

Graduate Program in Translation and Interpretation

National Taiwan University

Master Thesis

聲音語調對聽眾注意力及中文同步口譯之影響

The Effects of Intonation on Listeners' Attention and Mandarin Chinese Simultaneous Interpreting

莊英里

Cristina Chuang

指導教授：郭貞秀博士

Advisor: Chen-Hsiu Grace Kuo, Ph.D.

共同指導教授：吳敏嘉教授

Co-advisor: Min-Chia Michelle Wu

中華民國 108 年 11 月

November 2019

doi:10.6342/NTU202000533

國立臺灣大學碩士學位論文
口試委員會審定書

聲音語調對聽眾注意力及中文同步口譯之影響
The Effects of Intonation on Listeners' Attention and Mandarin Chinese Simultaneous Interpreting

本論文係莊英里君（學號 R05147012）在國立臺灣大學翻譯碩士學位學程完成之碩士學位論文，於民國 108 年 11 月 18 日承下列考試委員審查通過及口試及格，特此證明

口試委員：郭貞秀

（指導教授）

吳敏嘉

（共同指導教授）

張嘉倩

陳雅玲

doi:10.6342/NTU202000533

Acknowledgement

I would like to first like to express my deepest gratitude to my advisors, Grace and Michelle for all the help and support they have given me during this time. Thank you, Grace for your guidance and support despite your busy schedule. I have always enjoyed our meetings and you are truly a grace of God. Thank you, Michelle for giving me all the resources and help that I needed to finish this work. You changed my life in so many ways and I can never be grateful enough.

I would also like to sincerely thank Dr. Chia-chien Chang and Dr. Ya-lin Chen for their insightful comments and encouragement during the oral defense and throughout the writing process. Your passion for research is truly admirable and inspiring.

I truly appreciate the generosity that Prof. Danny Huang, Prof. Shan-shan Wang, and Amy Wu have shown me this year. Without your help and contribution, I would have never recruited so many participants.

Thank you, Jacky Siu and David Chuang for offering me your advice and experience. I am truly thankful to have had the helping hand of such brilliant thinkers.

Last but not least, I thank God for his abundant provision, for he gave me all the help and love I needed during what ended up being one of the toughest times of my life. Thank you for placing people who believed in me when I could not believe in myself. Thank you for giving me the wisdom and courage to strive forward when I only felt like giving up. Without you, I would not have been able to come this far.

doi:10.6342/NTU202000533

Abstract

During the simultaneous interpreting (SI) process, interpreters are often not visible to the audience and intonation is the most salient non-verbal element for interpreters to convey the speaker's message. Despite intonation's crucial role, existing research on intonation in the interpreting field have mostly focused on how intonation affects the audience's perception of the interpreter's ability. There is very little investigation about the role of intonation in speech communication and its effect on listeners' attention. The objectives of this research are to (1) uncover factors underlying intonation in SI and study whether intonation can serve as a *contextual cue* in Treisman's attenuation theory (1960) and increase listeners' attention; (2) investigate how intonation may benefit future interpreting training and practice. This paper includes a quantitative study (Part 1) and a qualitative study (Part 2). Part 1 is a dichotic test participated by 132 subjects from National Taiwan University to explore whether a lively intonation (wider F0 range) can enhance listening attention in Mandarin Chinese. Part 2 is a questionnaire investigating 29 professional interpreters' subjective evaluation of intonation's role in interpreting, the effects of monotony on their performance, and the strategies they adopt in order to overcome the challenges.

doi:10.6342/NTU202000533

The findings of this research prove that intonation can significantly enhance listeners' attention in Mandarin Chinese, especially for statement sentences. Results also reveal that the majority of interpreters acknowledge intonation's significance as a non-verbal cue and believe it is crucial for effective delivery.

This research offers insights into the practice and training of interpreting, establishing why and how intonation plays a decisive role in Mandarin Chinese SI. Interpreting trainers should factor in intonation when selecting material for classroom instruction in order to raise intonation awareness and explore the possibility of using intonation as a strategy while performing Mandarin Chinese SI.

Key words: Simultaneous interpreting, intonation, monotony, listening attention, contextual cue, Mandarin Chinese prosody

doi:10.6342/NTU202000533

摘要

口譯員的語調（intonation）是同步口譯過程中傳達講者信息的重要非語言交際。過去相關研究大多著重於口譯員的語調如何影響聽眾對於其口譯品質的評價，並未探討語調在此過程中所扮演的角色以及其如何影響聽眾注意力。本研究指在了解語調是否可作為崔斯曼在減弱理論(1960)中所提出的的一種情境提示(contextual cue)提升聽眾注意力及語調如何影響口譯教學及實務工作。本研究第一部分為一個雙耳聽覺測試，共有 132 名受試者參與，實驗目的旨在了解語調變化(基本頻率範圍)是否可以提升中文聽眾之注意力。第二部分則是針對 29 名專業口譯員進行線上問卷調查，了解口譯員對於語調的重視程度、單調的講者是否會對其口譯表現造成影響及面對此類型講者所採取的口譯策略。研究結果指出變化較為豐富的語調可以顯著提升中文聽眾之注意力，尤其是增加其對於敘述句（statement sentences）的理解及大部分口譯員都同意聲音語調是一個重要的非語言交際，能夠協助其有效溝通。本研究透過證實聲音語調對於中文同步口譯之重要性為口譯教學及實務工作帶來新的觀點：口譯教學工作者在篩選練習材料可以將講者的聲音語調納入考量、提升學生對於聲音語調的意識，並將聲音語調作為一種口譯策略。

關鍵字: 同步口譯、語調、單調、情境提示、聽覺注意力、中文語調

doi:10.6342/NTU202000533

Table of Contents

Chapter One Introduction.....1

1.1 Intonation in Simultaneous Interpreting..... 1

1.2 Motivation..... 3

1.3 Objective of the Study..... 4

1.4 Organization..... 5

Chapter Two Literature Review.....7

2.1 Intonation..... 7

2.2 Intonation in Interpreting.....12

2.3 Intonation's effect in Mandarin Chinese.....14

2.4 Intonation and Attention..... 20

2.5 Summary..... 23

Chapter Three Part 1..... 25

3.1 Method.....25

3.1.1 Procedure..... 25

3.1.2 Material..... 27

3.1.2.1 Stimuli Design..... 27

doi:10.6342/NTU202000533

3.2 Procedure..... 30

3.3 Data Analysis..... 31

3.4 Subjects..... 31

3.5 Hypothesis..... 32

3.6 Results..... 33

3.6.1 Group scores for dichotic test..... 34

3.6.2 Group Scores for Emotional Texts and Non-emotional Texts..... 36

3.6.3 Observed common group mistakes..... 43

3.6.4 Part 1 questionnaire results..... 45

Chapter Four Part 2..... 49

4.1 Method..... 49

4.1.1 Procedure..... 49

4.1.2 Material..... 49

4.1.3 Subjects..... 50

4.2 Results..... 52

4.2.1 Interpreters' View on Intonation's Role in Mandarin Chinese SI..... 52

doi:10.6342/NTU202000533

4.2.2 Monotony's Effect on Interpreters During SI..... 55

4.2.3 Strategies Applied when Interpreting Monotonous Speakers..... 56

4.2.4 Should Interpreting trainers focus more on voice training?..... 57

Chapter Five General Discussion..... 59

5.1 Traffic Jam Theory: How Intonation Affects Listeners' Attention.. 59

5.2 How Intonation Affects Listeners' Attention in SI..... 65

5.3 How Interpreters Perceive the Role of Intonation in Mandarin Chinese SI..... 69

5.4 Implications of the Study 71

Chapter Six Conclusion..... 76

6.1 Answering the research questions 76

6.2 Contributions of the study..... 78

6.3 Suggestions for future research..... 78

References..... 80

Appendix 1 Rating Survey for Part 1..... 88

Appendix 2 Passages used for Part 2 90

Appendix 3 Part 1 Questionnaire..... 92

doi:10.6342/NTU202000533

Appendix 4 Part 2 Questionnaire..... 95

doi:10.6342/NTU202000533

List of Figures

Figure 1. F0 of the final Tone2 (S: Statement; Q: Question) (Yuan, 2006, pg.26) 17

Figure 2. PRAAT graph of a lively version of "森林裡住著一隻狐狸" 28

Figure 3. PRAAT graph of a monotonous version of "森林裡住著一隻獼猴" 29

Figure 4. Average scores for Group LM, LL, ML, and MM 35

Figure 5. Average scores for Group LM and Group LL 35

Figure 6. Average score of Group ML and Group MM 36

Figure 7. Average group score of emotional texts and non-emotional texts 37

Figure 8. Average group score of question sentences 38

Figure 9. Average group score of statement sentences. 39

Figure 10. Average group score of emotional texts. 40

Figure 11. Average group score of emotional texts with the emotion joy. 41

Figure 12. Average group score of emotional texts with the emotion sadness 41

Figure 13. Average group score for emotional texts with the emotion fear. 42

Figure 14. Average group score for emotional texts with the emotion anger 43

Figure 15. Subjects' responds regarding intonation and their will to listen. 45

Figure 16. Experienced interpreters' years of practice 51

doi:10.6342/NTU202000533

Figure 17. Novice interpreters' years of practice..... 51

Figure 18. Ratings of Essential Elements in SI..... 52

Figure 19. Ratings of Intonation's role in SI 53

Figure 20. Diagram illustrating how subjects in Group LM process signals..... 62

Figure 21. Diagram illustrating how subjects in Group LL process signals..... 63

Figure 22. Diagram illustrating how subjects in Group MM process signals. 65

Figure 23. Group scores for the different emotional texts. 68

doi:10.6342/NTU202000533

List of Tables

Table 1 The acoustic features of the four basic emotions obtained from Yuan (2002)'s studies. 18

Table 2 The F0 features of English and Mandarin Chinese speech in the four basic emotions (anger, fear, joy, sadness) obtained from Murray et al.'s and Yuan's studies. 18

Table 3 The different tests performed by the groups 26

Table 4 The F0 values of the lively vs. monotonous versions. 29

Table 5 The means and standard deviations of the monotonous and lively passages 30

doi:10.6342/NTU202000533

Chapter One Introduction

1.1 Intonation in Simultaneous Interpreting

When simultaneous interpreting (SI) first debuted in the Nuremberg trials in 1945, it was shrouded with mystery and the world was fascinated by how interpreters could transfer one language to another in an effortless manner almost instantaneously as the speaker spoke (Seeber, 2015). Since then, an increasing number of national and international organizations have relied on SI to operate and communicate on a daily basis. In a simultaneous interpreting setting, interpreters usually sit inside a booth and deliver rendered messages to the audience via a microphone while the audience listens through headsets. In most cases, the audience solely relies on the interpreter's verbal language and paralanguage in order to comprehend the speaker's speech. Since interpreters are often not visible to the audience, intonation, as one of the prosodic cues, is the most commonly used non-lexical feature to help interpreters deliver the emotional or attitudinal status of the speaker (Poyatos, 1997).

The terms "prosody" and "intonation" are often used interchangeably since prosody is generally regarded as a complex system of features including intonation (Cutler & Ladd, 1983). As Seeber (2015) succinctly defined, "intonation is the melody of a sentence that surpasses its grammatical form (p. 69)." Despite the lack of consensus on intonation's definition due to its complex relation with other prosodic features such as rhythm and speech rate, most interpreting research has defined intonation as *the pitch contour of an utterance* (Ahrens, 2004; Seeber, 2001). Prosody, on the other hand, involves changes in fundamental frequency (F0), duration and amplitude. All these components interact with

doi:10.6342/NTU202000533

one another to create a final product that surpasses the whole of its units (Beckman 1986; L. C. Nygaard & D. S. Herold, L. L, 2009). In line with most of the previous work, intonation will be defined as *the change in fundamental frequency* (*a.k.a.* F0 range; the difference between the maximum F0 and the minimum F0) which is the major acoustic measurement in this paper and is used to facilitate the design of the first experiment (Holub, 2010; Langlet, 2012).

Researchers from the fields of linguistics have long studied the diverse functions and the role of intonation in speech production. Linguistic research has shown that acoustic features in non-lexical speech have a greater impact on listeners' comprehension of speech attitude and emotion compared to verbal communication (Mehrabian & Wiener's, 1967). Intonation provides semantic information, conveying tense, emotion, and mood (Halliday, 1967; Shlesinger, 1994) and is perceived as a contextual cue, aiding participants to comprehend the relationship among different actors, the intentions and emotions involved, and the context of the conversation (Steling,1992). Moreover, intonation is regarded as an integral part of communication since it provides grammatical functions to help listeners differentiate the syntactic structure of the sentence. Despite intonation's crucial role in facilitating comprehension in speech communication, it has received surprisingly little focus in the field of interpreting (Holub 2010, Seeber 2015, Langlet 2013).

The existing literature studies on intonation in the interpreting field have mostly focused on how intonation affects audience's perception of the interpreter's ability (Moser, 1996). Little research has been done to investigate intonation's role in the communication between the interpreters and the audience. Among the few studies on intonation's role in SI, the empirical studies of Collados Aís (2001), Holub (2010), Holub and Rennet (2011) and

doi:10.6342/NTU202000533

Lenglet (2013) have contributed to the ways in which intonation may affect listeners' comprehension. Collados Aís (2001) investigated the impact of monotony, or the lack of intonation, on listeners' ability to recall information by creating variations of the same text through reciting it with different levels of liveliness. Holub (2010) utilized the software PRAAT to replicate a monotonous version of the original text by flattening the F0. The results from both studies suggested a tendency of monotonous text hindering listeners' comprehension. In addition, Holub and Rennet's (2011) follow-up experiment revealed that intonation had more impact on listeners' comprehension than fluency.

1.2 Motivation

In spite of the existing studies, the lack of consensus on how intonation and comprehensibility were defined in each study leaves the role of intonation in SI yet inconclusive. In addition, the lack of observation studying intonation's relation with listeners' comprehension show that almost none of the existing literatures have focused on explaining the mechanism behind such findings. Listening comprehension is a highly complex cognitive process; therefore, a more focused research investigating how, and to what extent, intonation may affect listeners' comprehension would be necessary to avoid overgeneralized claims.

It is also important to mention that most of these experiments were conducted using stress-timed languages such as English and German (Holub,2010; Seeber, 2015) and there seems to be no comprehensive literature investigating the relation of intonation and comprehension in Mandarin Chinese within the interpreting field. Mandarin Chinese is a tonal language through which tones carry lexical and intonational meanings, hence

doi:10.6342/NTU202000533

Mandarin Chinese demonstrates a greater average rate of F0 fluctuation (Eady, 1928). Since Mandarin Chinese is a syllable-timed language, intonation's effect on speech acts and discourse effects are expected to be different from those in stress-timed languages. This discrepancy between previous research shows a need for further investigation and data collection in order to understand intonation's role in listener comprehension in Mandarin Chinese SI.

Attention is a multifaceted concept and psychological research has long studied the process which the perceptual filter screens the stimulus (Broadbent, 1958; Cherry, 1953; Moray, 1959; Treisman, 1960). Broadbent's early perceptual theory (1958) suggests that physical traits like loudness and pitch are the prime criteria in selective listening. In follow-up experiments, Treisman (1960) added that psychological factors such as one's relevance to the message and *contextual cues* can also lower the threshold for signals to pass through. If Treisman's theory (1960) can also apply to a SI setting, then by exploring intonation's role as a contextual cue may offer significant insight to how intonation affects listeners' attention.

1.3 Objective of the Study

As previously mentioned, interpreters deliver messages to the audience using verbal language and prosodic cues. One of the most salient features of intonation is being a contextual cue to help listeners understand the speech (Halliday, 1967; Shlesinger, 1994). It is therefore crucial that interpreters understand how prosodic features in their interpretation may serve as a double-edged sword— aiding them to deliver the message as powerfully as the speaker, or hindering them and altering the meaning of the original text. In spite of the

doi:10.6342/NTU202000533

importance of intonation in speech communication, intonation's role in listening attention is not completely understood and little research has been done to investigate intonation's role in Mandarin Chinese in the interpreting field.

In order to fill in the gaps in the existing literature and offer insight for interpreting practice and pedagogical purposes, the current study attempts to address the following questions:

1. Can intonation (e.g., a wider F0 range), serve as a contextual cue as defined in Treisman's attenuation theory (1960) and help increase listeners' attention in Mandarin Chinese SI. If yes, then how?
2. How will this study benefit future interpreting training and practice?

1.4 Organization

The present study is divided into five chapters. The first chapter *Introduction* explains the background of the research and specifies the purpose and importance of the study. The second chapter *Literature Review* presents an overview of the relevant literature in order to contextualize the research of this paper. The researcher will review studies on intonation, its effect in Mandarin Chinese, and studies related to intonation within the interpreting field. Then, the chapter will briefly discuss psychological studies regarding listening attention to explore the notion of intonation as a contextual cue. The third chapter *Part 1* lays out the mechanism, purpose, design, and statistical results for the dichotic test while the fourth chapter *Part 2* presents the details and results of a questionnaire investigating interpreters' views on intonation. The fifth chapter *General Discussion* provides an analysis

doi:10.6342/NTU202000533

and explanation of the data acquired from the quantitative and qualitative research from Part 1 and Part 2 in order to answer the research questions and the sixth chapter *Conclusion* recapitulates the purpose and importance of the study and highlights how it can benefit future interpreting practice and training.

doi:10.6342/NTU202000533

Chapter Two Literature Review

The existing literature studies on intonation in the interpreting field have mostly focused on how intonation affects audience's perception of the interpreting quality (see section 2.2 for review). The role of intonation in speech perception and its effect on listeners' attention remains a relatively unexplored area. The aim of this chapter is to present an overview of the relevant literatures in order to contextualize intonation's effect on speech production and explore the notion of intonation as a contextual cue. For this purpose, this chapter will include: (1) the notion of intonation and its diverse functions in the process of communication (2) interpreting studies on intonation (3) intonation's role in Mandarin Chinese (4) the relation of intonation and attention. By building on these studies, this research aims to establish the importance of intonation serving as a contextual cue and explore its various effect on listeners' attention.

2.1 Intonation

What is Intonation?

Recent studies have focused on prosody and discussed its important role in speech. As Cutler stated (1997, p.141), "Prosody is an intrinsic determinant of the form of spoken language". The definition of prosody often varies depending on the purpose of the research. It is observed that the majority of the studies would define prosody as "the structure that organizes sounds" or "the synonym for pitch, tempo, loudness, and pause" (Cutler, 1997). The term "prosody" is often used as a synonym for "intonation" since they both involve the measurement of fundamental frequency (F0) acoustically (Holub, 2010).

doi:10.6342/NTU202000533

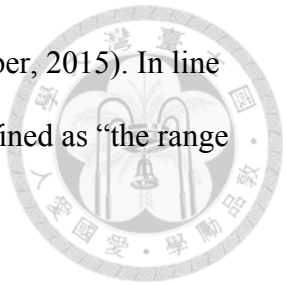
Although intonation has gained significant importance, it is crucial to note that there seems to be no single definition or shared consensus on how to describe intonation due to its varying features (Seeber, 2015). For instance, Beckman (1986) referred to intonation as "a combination of the acoustic patterns of fundamental frequency, duration, amplitude, segmentation, and articulation". Hargrove and McGarr (1994:16) viewed intonation as the "communicative use of pitch". Shlesinger (1994) characterized intonation with tones (pitch movement) and their features such as duration and speed. Similarly, Ahrens (2005) defined intonation as "the pitch contour of an utterance". In addition, Seeber (2015) proposes that "intonation is the melody of a sentence that surpasses its grammatical form". It is important to acknowledge that intonation is composed of changes in F0, duration, and amplitude; all these components interact with each other to create a final product that surpasses the whole of its units (Crystal,1975). As mentioned before, the terms "prosody" and "intonation" are often used interchangeably due to its common features. Therefore, the most comprehensive definition of intonation is perhaps the definition provided by L. C. Nygaard, D. S. Herold, L. L:

> "Prosody is the melody of speech, consists of intonation, rhythm, and relative loudness and timing of components of an utterance and is instantiated primarily in the acoustic correlates of fundamental frequency, amplitude, and relative duration." (L. C. Nygaard, D. S. Herold, L. L, 2009, p.128)

Despite the lack of consensus on intonation's definition due to its complex relation with other prosodic features, research on intonation within the interpreting field has mostly focused on "the changes of pitch" (Ahrens, 2004; Möbius et al.,1993; Seeber, 2001) or "the contour of fundamental frequency (F0)" in order to "establish correlation between

doi:10.6342/NTU202000533

instrumentally measured frequencies and perceptual judgements" (Seeber, 2015). In line with previous work (Holub, 2010; Ahrens, 2004) intonation will be defined as "the range and variation of F0" in order to facilitate the design of the experiment.



The Effects and Functions of Intonation

In the process of spoken communication, extra linguistic elements such as expressions, gestures, movements and proxemics provide rich communicative signals to listeners (Fantini, 1995). Studies on the effect of intonation have discovered two important functions of intonation: (1) intonation serves to clarify the structure of spoken language, and (2) intonation provides information on the psychological and attitudinal state of the speaker (Poyatos,1997).

Intonation contour helps the receiver disambiguate syllables, words and phrases. Observations have showed that children and language learners use syllable duration and intonation contour to identify the syntactic structure of sentences (L. C. Nygaard, D. S. Herold, L. L. Namy, 2009). For example, the end of a sentence is often marked by a lower pitch, which helps listeners distinguish whether the phrase is complete or not. Intonation also helps listeners differentiate questions from statements. In English, interrogative phrases that elicit either a "yes" or "no" response are often characterized by a rise in intonation in the end while complete statements show a falling pattern in intonation.

Intonation also helps provide information on the psychological state or attitude of the speaker. Prosodic cues help indicate speaker's intentions and emotion (Bachorowski, 1999; Scherer, 1994) and they also help listeners identify the speaker's emphasis or focus in a

doi:10.6342/NTU202000533

speech (Cohen, Douaire, & Elsabbagh, 2001; Liu, 1995). A wide range of studies have shown that emotional tone of voice may contribute to the semantic and linguistic interpretation of words. Among them, Nygaard & Launders' (2002) experiment discovered that homophones (bridal/bridle or die/dye) spoken with emotional prosody would lead the listeners to interpret words with an emotional meaning rather than a neutral meaning. Their experiment revealed that prosody contributes to the semantic comprehension of words, acting as a communicative signal, which facilitates one's process of interpreting novel words. Nygaard & Lauders' (2002) experiment concluded that pitch, duration, and phonation are correlated and paired with specific semantics. As a result, listeners utilize the relation of prosodic features and their associated meaning as semantic cues.

Though the association of intonation with particular meanings is unique to each community, some general rules can still be observed. For example, the pitch range tends to be wider at the beginning of a sentence while narrower and lower towards the end (Vaissière,1983; Hargrove & McGarr, 1994). A wider pitch range also indicates interest whereas a narrower pitch range implies familiarity with the topic, boredom, or depression (Ladd, 1992; Seeber, 2015). All these examples show how intonation can prove syntactic and semantic meaning in speech communication.

Intonation as a Contextual Cue

Activities which help participants interpret an utterance in a particular scenario are referred as contextualizing language (Auer, 1992). These activities may include speech, mood, modality, topic, as well as the social relationship between speakers. Contextual cue is therefore defined as the mean which helps speakers contextualize language (Auer, 1992;

doi:10.6342/NTU202000533

Tzeng et al.,2018). Though there is a wide range of studies related to contextualization, studies on non-lexical contextual cues like prosody and linguistic variations are most commonly explored and investigated.

It is observed that intonation acts as a formal cue and helps participants comprehend the relationship between different participants and also the intentions, emotions, and atmosphere of the conversation. These information help participants interpret the conversation in a more holistic manner and act as signals to highlight the contrast or sameness in speech; therefore, attracting participant's attention to the cohesive relation of the conversation (Levinson, 2003; Steling, 1992). Intonation, as a contextual cue, acts like "a knot in the handkerchief as a formal reminder of something" as Levinson (2003) succinctly defined.

Similarly, many early findings in psycholinguistics also highlight the role of prosodic features in speech recognition. Research has shown how a string of nonsense syllables is easier to recall if presented with sentence prosody than without (Cutler, 1997; Epstein, 1961). Furthermore, grammatical strings are only easier to shadow than ungrammatical strings if they possess sentence prosody (Cutler, 1997; Martin, 1968). Other research also found that nonsense utterances could be recalled and recognized more accurately if they were spoken with prosody, suggesting that prosodic features may "form part of the memory representation" which in turns helps listeners form the input (Speer et al.,1993; Culter, 1997). All the research has shown that prosody is crucial to speech communication and that prosody helps convey messages more efficiently (Cutler,1997).

doi:10.6342/NTU202000533

2.2 Intonation in Interpreting

Intonation is an extensively studied subject in linguistics and psychology yet it has received surprisingly little interest in the field of interpreting. The studies of intonation in the interpreting field have mostly focused on how intonation may affect audience's perception of the interpreter's professional ability (Collados Aís, 1998; Moser 1996, Seeber, 2015; Lenglet, 2012; Zwischenberger, 2010). Only few authors conducted research related to prosody, or intonation in SI (Collados Aís, 2001; Holub, 2010; Seeber, 2015). Moreover, the number of research examining the link between intonation and attention is even more limited.

Collados Aís (2001) designed a ten-minute source text in German and then made three versions of translation in Spanish: one with lively intonation and sense inconsistencies, one with lively intonation without sense inconsistencies, and one monotonous version without sense inconsistencies. Forty-two experts and fifteen interpreters were randomly divided into three groups to listen to one of the three versions and asked to rate the quality of the interpretation and complete a comprehension test related to the content they heard. Collados Aís discovered that the lively interpretations (with or without sense consistencies) received a significantly higher rating than the monotonous interpretation and this finding revealed that listeners tended to value intonation over accuracy. Though there was no significant difference between the comprehension score of the monotonous group and the two lively groups, Collados Aís claimed that the difference in scores between the groups suggested how monotonous intonation might negatively impact listeners' attention. However, the lack of acoustic measures and statistical data left this conclusion inconclusive.

doi:10.6342/NTU202000533

Similarly, Holub (2010) investigated the impact of intonation on listeners' comprehension and assessment of SI by simulating a real-life conference. She selected a specialized source text in English and recorded its SI in German. Unlike Collados Aís (2001), Holub used the software PRAAT to produce an artificial monotonous version of the content in order to control other speech parameters such as duration, pause, and amplitude. Forty-nine economics students were divided into two groups to complete a comprehension and assessment test after listening to separate versions of the translation. The results indicated a moderate correlation between intonation and interpreting assessment, but a "strong tendency that flattened fundamental frequency impedes comprehension." (Holub, 2010, p.124)

In a follow-up experiment, Holub & Rennert (2011) used the same material from the Holub's experiment (2010) to study how monotony and fluency would affect listeners' comprehension. In addition to the source text and the monotonous text, they created two other versions: a non-fluent version with disfluencies such as pauses, repetitions, false starts, and consonant lengthening; and a monotonous text with disfluencies. Holub & Rennert (2011) then studied listeners' subjective and objective understanding of their comprehensibility. Findings revealed that poor fluency alone did not affect listeners' objective comprehension. Moreover, monotony worsened listeners' objective comprehensibility even though it did not influence listeners' subjective comprehension negatively. Judging on the results, Holub & Rennert (2011) concluded that poor fluency may lead to a worse impression of an SI performance yet monotony had the greatest impact in listeners' comprehension.

doi:10.6342/NTU202000533

Lenglet (2013) tested whether the peculiar prosody of SI, mainly the sentence segmentations and rising tone in the end of sentences, impacted listeners' comprehension and perception of quality. Forty-nine economics students and thirty translation students were randomly divided into two groups to listen to two versions of interpretation. One was genuine SI while the other was a shadowing version that sounded like a prepared speech. Listeners then filled a comprehension test and a questionnaire to assess their subjective evaluation of their comprehension and their perception of quality. Though the experiment proved how listeners link their perception of fluency and accuracy with better interpreting quality, the experiment revealed no significant difference in comprehensibility between the two groups.

From the literatures presented in this section, we find that intonation may influence listeners' comprehension of the interpretation as well as their perception of the interpreting quality. The lack of the empirical studies on exploring the role of intonation in SI and on the effect of intonation on listener's comprehension are still not completely understood. It is also worth mentioning that the majority of the experiments presented in this section were conducted in stress-timed languages and so far there is no interpreting research investigating how intonation may affect Mandarin Chinese audience's perception and comprehension of the interpreting. The discrepancy between previous research shows a need for further investigation and data collection in order to address knowledge gaps in the literature.

2.3 Intonation's effect in Mandarin Chinese

doi:10.6342/NTU202000533

In order to examine how intonation may affect listeners' attention in Mandarin Chinese SI, it is important to understand the prosodic features of Mandarin Chinese and the role intonation plays in Mandarin Chinese speech production. This section will summarize the studies on Mandarin Chinese's prosodic features, the relation between intonation and speech act, and the F0 features of Mandarin Chinese's emotional speech.

Mandarin Chinese is a tone language where tones are lexically specified. Different F0 registers (or changes in F0 pitch) of a syllable have different meanings. There are four lexical tones in Mandarin Chinese which are often known as Tone 1 (high level tone), Tone 2 (mid-rising tone), Tone 3 (low-dipping one), and Tone 4 (high-falling tone). The F0 height and the F0 trajectory over a syllable are crucial components in tone production and perception. Aside from prosody at the lexical level, Mandarin Chinese also has a prosody at the utterance level. This second level of prosody is when F0 interacts with duration and amplitude over a phrase. Hence, the F0 in Mandarin Chinese reflects the interaction between tones and intonation (Yang, 2011).

In Chinese, tone and intonation are both presented in the F0 curves and therefore this is a great amount of questions regarding how Chinese intonation should be defined. Some researchers explore intonation at the sentence level while others refer intonation to the F0 changes of prosodic units. Despite of their different focuses, most researchers acknowledge the fact that Chinese intonation patterns may occur repeatedly at the sentence level (Pan, 2012). It is also observed that as a tonal language, speeches in Mandarin Chinese display a greater average rate of F0 change and fluctuation than in stress-timed languages such as English where F0 patterns are only prominent in certain lexical items (Eady, 1928).

doi:10.6342/NTU202000533

Different intonations for sentences such as statements, exclamation sentences or questions are realized as different F0 changes and different discourse effects (Shih, 2000; Thorsen, 1980; Garding, E. 1987). For example, *declination effect,* a gradual declination of F0 in a sentence which creates a downtrend tilt of intonation, is often regarded as an element of statement intonation (Shih, 2000; Pierrehumbert, 1980). In statement sentences, the pitch is typically higher in the initial position and lowered in the final position (Shih, 2000; Hirschberg & Pierrehumbert, 1986; Garding, E. 1987). In addition, a higher pitch followed by a downward trend of F0 can also signal the beginning of a new sentence (Shih, 1988; Liao, 1994). Declination effect is also prominent in post-focus sentences since focus in Mandarin Chinese is often marked by a sharp rise of F0 (Shih, 2000; Yuan, 2002).

Yuan (2006) investigated the characteristics of question intonation in Mandarin Chinese and how the question intonation differs from statement intonation. He observed that question intonation has a higher F0 contour than statement intonation. Question intonation also has greater strength of sentence final tones, resulting in a widening F0 intensity gap between the question intonation and the statement intonation toward the end of sentences. Generally speaking, question intonation shows higher F0 than statement intonation, especially for questions ending in Tone 2 (rising) syllables (See Figure 1). In addition, the F0 of the last syllable in Mandarin Chinese greatly affects listeners' performance on identifying the question intonation and the statement intonation. The tone of the last syllable tends to only affect question intonation identification partly because the last syllable in a question is usually a question particle, such as *ma* and *ne* (Yuan, 2006).

doi:10.6342/NTU202000533

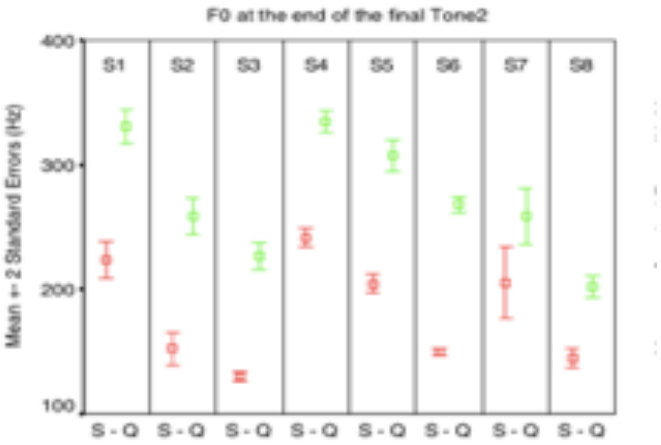


Figure 1. F0 of the final Tone2 (S: Statement; Q: Question) (Yuan, 2006, p.26) The study demonstrated that for all subject, question intonation tended to be higher than statement intonation.

Intonation and Emotions

As mentioned before, intonation plays an important role in conveying vocal emotions. Though the representation of different emotions with respective prosodic features has mostly focused on non-tone languages like English (Murray et. al., 1993), studies analyzing vocal emotions in tone languages like Mandarin Chinese are receiving increasing attention (Li, et al. 2011). In Chao's (1933) studies, he studied how intonation, or expressive intonation, conveys emotion in Mandarin Chinese. He proposed that expressive intonation is defined by the quality of the voice, the unusual degree of stress, and the general intensity and duration of the speech. Though emotional speech is the produce of different prosodic features, the researcher will only briefly present literatures which investigate emotional speech focusing on the variation of F0 contour for the purpose of the study.

doi:10.6342/NTU202000533

Yuan (2002) studied the acoustic realization of anger, fear, joy, and sadness in Mandarin Chinese by researching how speakers associate phonation, articulation and prosody with each emotion (See Table 1 and Table 2). Yuan concluded that joy is mainly realized on F0 while anger and fear were mainly realized on phonation. The F0 of joy was marked with a high intensity and great F0 fluctuation, while anger and fear were characterized by a creaky and breathy phonation with narrower pitch range. Sadness, on the other hand, is a combination of phonation and intonation, marked by a low F0 with small fluctuation and a breathy phonation. Similar research also showed that the F0 magnitude (= F0 range) of emotional speech in Mandarin Chinese were ranked as follow: anger > joy > neutral > fear > sadness, indicating how active emotions demonstrated a larger F0 fluctuation compared to passive emotions (Li et al., 2011; Gu & Lee, 2007).

Table 1. The acoustic features of the four basic emotions obtained from Yuan (2002)'s studies.

	Anger	Fear	Joy	Sadness
Phonation:	creaky	breathy	modal	breathy
F0 height	high	high	high	low
F0 fluctuation	big	small	small	small
Sentence length	anger< fear< joy< sadness			
Pause	joy < anger< fear < sadness			

Table 2. The F0 features of English and Mandarin Chinese speech in the four basic emotions (anger, fear, joy, sadness) obtained from Murray et. al.'s and Yuan's studies.

doi:10.6342/NTU202000533

	Murray et. al. (English)	Yuan et al. (Mandarin Chinese)
Anger	A high F0 average and abrupt on stressed	Widest F0 range and raises in a certain syllable to create F0 peak
Fear	A higher F0 average and a normal F0 inflection	Narrow F0 range with weaker declination of F0 compared to anger and joy. Fear is very similar to sadness, only showing a higher F0 contour
Joy	A relative higher F0 average and a smooth rising	Larger declination and wider F0 range compared to fear
Sadness	A slightly lower F0 average with a downward effect	Shows lowest F0 value and narrowest F0 range. F0 contours are significantly lower and flatter

From these observations, one can observe how intonation plays a crucial role in helping Chinese speakers communicate more effectively. The rising intonation at the end of a sentence shows question and doubt; a sentence with declination effect indicates assuredness or completeness. Intonation is especially important for listeners to distinguish questions and statements and to interpret whether the sentence is a new topic or a post-focused sentence (Grice and Bauman, 2007). In addition, different moods and meanings can also be better conveyed through the fluctuation of intonations. However, one of the greatest features of Mandarin Chinese is that most of the changes of intonation were found in the sentence-final position. This is different from other stress-timed languages, such as English, where the intonation and pitch movement would mostly occur in the middle of the sentence (Pan, 2012). Since the intonation patterns of Mandarin Chinese are most

doi:10.6342/NTU202000533

prominent in the tail of sentences, it is reasonable to suggest that speakers would pay most attention to the intonation of sentence-finals.

2.4 Intonation and Attention

Interpreters are communicators who interact between speakers, listeners, and other interpreters. Understanding the process of listening comprehension may enable interpreters to render messages more effectively and as a result, increase audiences' engagement in the conference. It would also offer suggestions to how interpreters' keep their listeners "attended" regardless of all the other stimuli in the conference room. As mentioned before, listening attention is an integral part of listening comprehension due to the fact that the level of attention a person devotes may significantly impact the degree of information he or she can encode (Houston, D. M., & Bergeson, T. R, 2013). It would therefore be reasonable to assume that attention is the cornerstone to listening comprehension. In this section, the researcher will briefly review two of the most prominent attention theories and relate these concepts to intonation's role in facilitating communication.

Attention is the action of processing incoming sensory information; it is selective, with limited capacity, and is reactive as well as deliberative (Posner, 1982; Driver, 2001; Musiek & Chermak, 2015). This process "can either be controlled voluntarily by the subject, or it can be captured by some external event" (Roda, 2006, p.560). Simply put, attention is the process that enables and guides the incoming perceptual information (Roda, 2006) and research on attention is primary concerned with the process of prioritizing relevant information while neglecting irrelevant or interfering information (Chun & Wolfe, 2001;Lavie & Tsal, 1994). Most attention studies focus on two aspects: the degree of

doi:10.6342/NTU202000533

control that the individual has over his or her attention and the external factors that may capture one's attention (Chun & Wolfe, 2001; Roda 2006).

Broadbent's early perceptual model (1958) is one of the most established theories and the foundation of many latter studies. He proposed two selection stages of perceptual stimuli: a pre-attentive stage and an attentive stage. In the pre-attentive stage, stimuli were assessed based on its relevance, loudness, and its frequency characteristics. In the attentive stage, stimuli were processed based on more abstract properties such as semantics. Unattended stimuli at the pre-attentive stage would be completely filtered out and only stimuli which reaches the attentive stage would receive higher level of processing (Broadbent, 1958; Roda, 2006). Nevertheless, latter studies showed that there were cases in which stimulus from the unattended ear could "break through" and interfere with the attended ear (Cherry, 1953). Treisman's (1960) "attenuation theory", a modification of Broadbent's model, suggested that the unattended stimuli received a greater level of processing based on properties such as physical traits, relevance, and contextual cues.

Contextual Cue in Treisman's Attenuation Theory

Treisman conducted a dichotic test in order to explore whether a contextual cue, or "expectancy based on transition probabilities between words" (1960, p.1) would be a factor which influences how stimuli pass through the selective attention filter. In her experiment, Treisman recorded four different passages, each fifty words long, and the passages were randomized and put into pairs. Subjects would listen to two passages at the same time, one from each ear, while "shadowing", or repeating what they heard on only one ear. In between the twentieth and the thirty-fifth word, the passages would be interchanged from one ear to the other. In other words, subjects would hear the first part from passage 1

doi:10.6342/NTU202000533

followed by the second part from passage 2 on one track, and the first part of passage 2 and the second part of passage 1 on the other track. Among the 18 subjects, 15 subjects occasionally repeated one or more words from the wrong ear during the switch, and in retrospect, only one subject was aware that the passage had been switched to the other ear. Treisman(1960) observed that the probability of this phenomena was highly linked with the transition probabilities between the passages. For example,

Passage in the left ear: "...I **SAW THE GIRL**/ song was wishing..."

Passage in the right ear: "…me that bird/ **JUMPING** in the street…"

(Words in capital letters were spoken by the subjects)

The experiment suggests that words with more contextual cues are more probable to pass the selective filter and be heard from the rejected ear, on the other hand, words with lower contextual cues will be "attenuated" rather than being absolutely filtered (Treisman, 1960). Though the transition probability lowered the thresholds of words, but the filter still operates in favor of the selected channel (the selected ear) and therefore subjects will return to the correct ear instead of shadowing to the end of the passage.

Treisman (1960) further explored the notion of a "dictionary" of known words in people's mind. Within the dictionary, there are certain group of words that "have permanently lower threshold for activation". For example, if the words "I sang a" were heard, the stored word "song" in the dictionary would have a considerably lower threshold, making it more likely to be activated, or heard. Other "important" words, such as one's name, words that signal danger such as "fire, watch out, look", are also contextually high and therefore more likely to be activated. This explanation provided a new perspective,

doi:10.6342/NTU202000533

suggesting (1) the selective filter (Broadbent, 1958) selectively raises thresholds for stimuli from the rejected ear rather than serving as an all-or-none barrier and (2) in addition to "physical" cues like pitch and loudness, the selective filter can also act according to the features of meaning.

In Treisman's theory, contextual cue was mostly defined as the semantic and syntactic relevance between words. The role of how prosodic features may influence the filtering process was never answered. As mentioned previously, intonation is also regarded as a contextual cue, proven to have semantic and syntactic functions in speech, facilitating communication. Based on these literatures, the researcher would like to explore the notion of intonation serving as a contextual cue and the how it may increase listeners' attention.

2.5 Summary

Despite the abundant amount of literature suggesting intonation's crucial role in speech recognition and communication, there is little research investigating intonation's influence on listeners within the interpreting field. Though past interpreting research has suggested that audience associate lively interpretation with professionalism and that a monotonous interpretation may have a negative impact on listener's comprehension (Holub, 2010) and satisfaction (Moser, 1996), there is still scant literature and statistics to explain intonation's role in SI and its effect on the audience. Moreover, a majority of the research studying intonation's effect on listeners in SI have been conducted in stress-timed languages and almost no studies were investigated in Mandarin Chinese.

In order to overcome intuitive claims, the researcher hopes to build on Treisman's (1960) attenuation theory to investigate whether intonation can serve as a contextual cue

doi:10.6342/NTU202000533

and facilitate listeners' listening process in SI. In other words, if intonation, as a contextual cue, can increase listeners' attention, then it may offer a possible explanation to why monotony impedes listeners' comprehension (Holub, 2010). Moreover, the experiments will be conducted in Mandarin Chinese in order to investigate the role of intonation in a tonal language SI. In conclusion, the researcher hopes to fill in the gaps in the literature by exploring how intonation may affect listeners' attention and listeners' subjective perception of the interpreting. The next section lays out the methodology for the two experiments conducted in order to answer these questions.

doi:10.6342/NTU202000533

Chapter Three Part 1

Part 1 was a quantitative research conducted to explore the notion of intonation and its effect on listeners' attention in Mandarin Chinese SI. The main purpose of the experiment was to investigate whether intonation, as a contextual cue, can lower the threshold of listeners' attention filter, and increase their attention when listening to interpreting.

The following section lays out the method and materials used for Part 1.

3.1 Method

3.1.1 Procedure

Part 1 utilized a dichotic listening test (Broadbent, 1958; Treisman, 1964) in which subjects were asked to shadow (i.e., repeat aloud) one of the two stimuli directed to both ears. By studying which ear (channel) subjects shadowed, one could distinguish which stimuli passed through the attention filter. There were four groups of subjects, and each performed a dichotic test listening to 16 pairs of passages (see Appendix 2). Group LL was exposed to two lively passages, Group MM to two monotonous passages, and both groups were asked to shadow a specific ear. On the other hand, Group ML and Group LM listened to a monotonous passage in one ear and a lively passage in another, and were asked to shadow the monotonous (ML) and the lively version (LM) respectively. Recordings were deliberately switched between ears and subjects were asked to shadow different ears in order to weaken the right-ear-advantage (REA).

doi:10.6342/NTU202000533

Table 3. The different tests performed by the groups.

	Group LL	Group MM	Group LM	Group 4 ML
Ear 1	Lively	Monotonous	Monotonous	Monotonous
Ear 2	Lively	Monotonous	Lively	Lively
Shadow	Lively	Monotonous	Lively	Monotonous

The experiment took place at NTU Foreign Language Teaching & Resource Center's multi-media teaching rooms. The multi-media teaching rooms accommodated 50 students and each featured computers and the Sanako Lab 100 hardware system. The Sanako Lab 100 hardware was used to play the recordings on subject's computers and also to record their answers. Subjects were given conference-used headsets to block outside sounds and disturbances, and to listen to the recordings. In addition, the microphone on the headset was used to record subject's performance of the dichotic test.

A music file was first played for sound check and subjects were asked to ensure they heard sounds from both ears. After the researcher explained the instructions and assigned subjects to shadow a specific ear, they were asked to perform a practice trial in order to familiarize with shadowing. The practice trial recording contained five pairs of sentences which did not appear in the formal experiment. A slide presenting the sentences that should have been shadowed was then shown to the subjects after the practice trial in order to verify their understanding of the instructions. Students were asked to complete an online questionnaire after finishing the dichotic test. The questionnaire contained questions (see Appendix 3) regarding their age, native language, hearing and psychological conditions, as well as their general reflection of the dichotic test. The entire procedure lasted within15

doi:10.6342/NTU202000533

minutes and all subjects received sweets worth 50 New Taiwanese Dollars after completing the experiment.

3.1.2 Material

3.1.2.1 Stimuli Design

Reading Material

There were 16 pairs of passages used as the reading material in the experiment. Among them, 8 pairs were coded as the "emotional text" because they contained emotional expressions such as *ren-xing* "willful". The other 8 "non-emotional text" pairs did not contain any of the emotional expressions but varied with their sentence type – half of them were statements, and half of them were questions. (see Appendix 2). The emotional texts included four kinds of emotions: joy, anger, sadness, and fear.

All sentences were generated using the free-association method and adjusted in order to ensure they sounded natural. Each pair of sentences shared the same word length and level of word frequency. To contrast the presented pair of stimuli, each pair of sentences contained different keywords of one or two characters. All the keywords shared identical lexical tones in order to avoid the interference of tones. For example:

"東京航空明年將結束營業" （Tokyo Airline will stop operating next year)

"京都航空明年將結束營業" (Kyoto Airline will stop operating next year)

The keywords for the emotional texts were placed at the end or near the end of sentences in order to give listeners time to distinguish the emotion of the text. For example:

doi:10.6342/NTU202000533

“你怎麼可以那麼過份！”(You have gone too far!)

“你怎麼可以那麼任性!” (You are too willful!)

A professional voice trainer recorded the 16 passages in two versions – one with a monotonous tone and the other with a lively tone. The recording was made in a quiet room in a classroom of National Taiwan University (NTU) with the speech software PRAAT (Version 6.0.43; Boersma, P. & Weenink, D.,2018) installed on a laptop.

F0 Range of the Passages

In line with the works of Ahrens (2004), and Holub (2010), intonation is defined as the f0 range. The f0 range is obtained by calculating the difference between the F0 max and the F0 min in each passage. Figure 2 shows the spectrogram of the passage “*In the forest lives a fox*” read in a lively tone. The blue line shows the pitch track. As demonstrated in the figure, the F0 max value was extracted from point a, and the F0 min value was extracted from point b. Figure 3 shows the spectrogram of the passage “In the forest lives a *macaque”* read in a monotonous tone, which is a minimal pair passage of the one in Figure 2.

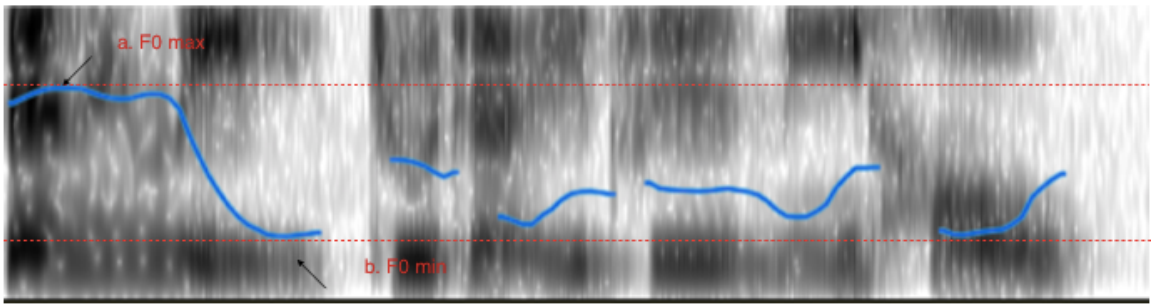


Figure 2. PRAAT graph of a lively version of “森林裡住著一隻狐狸”.

doi:10.6342/NTU202000533

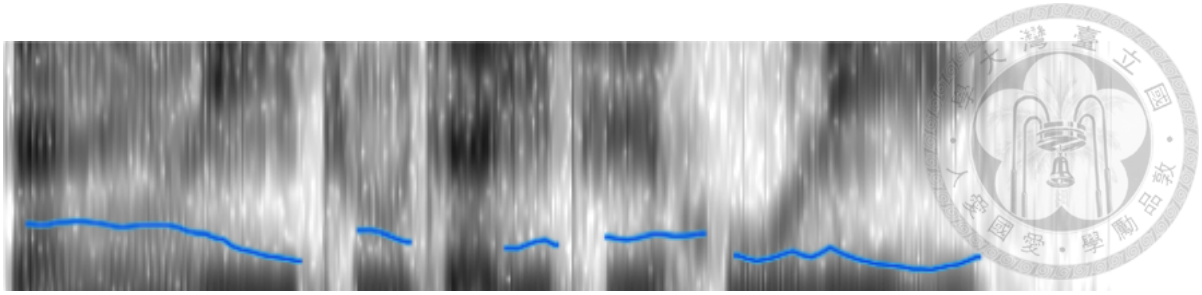


Figure 3. PRAAT graph of a monotonous version of "森林裡住著一隻獼猴".

The patterns of the pitch tracks in the two figures are distinctive – the one of the lively version (Figure 2) shows more fluctuation and variation while the one of the monotonous version (Figure 3) is flatter. The F0 max of the lively version is 242.04 Hz and the F0 min is 100.4 Hz (F0 range = 141 Hz). The F0 max and F0 min are 136.87 Hz and 54.9Hz (F0 range = 82 Hz) for the monotonous version.

Table 4 shows the average F0 Max, F0 Min and the f0 range for the 16 lively and 16 monotonous versions of speech used in the experiment. A paired-two sample t-test shows significant difference in f0 range for the two versions of, more specifically, lively speech has wider f0 range (= 168.65 Hz) than monotonous speech (= 59.62 Hz); t (31)=6.13, P<0.001

Table 4. The F0 values of the lively vs. monotonous versions.

	Lively	**Monotonous**
Maximum	302.36 Hz	157.34 Hz
Minimum	133.71 Hz	97.72 Hz
F0 range	168.65 Hz	59.62 Hz

Rated Liveliness of the Passages

doi:10.6342/NTU202000533

In order to further ensure the validity of the stimuli in listeners' ears, the recorded passages were randomized and rated by 30 raters. The raters were asked to listen to the 32 pairs of passages and rate on a four-point Likert Scale, 1 being monotonous and 4 being lively (see Apendix 1). The results are shown in Table 5 below. The rating results ensures that the two versions prepared, monotonous vs. lively, were valid in naïve native listeners' ears.

Table 5. The means and standard deviations of the monotonous and lively passages ratings.

Passages	Means	Standard Deviation
Monotonous	1.90625	0.857
Lively	3.404296875	0.723

3.2 Procedure

Subjects participated in a dichotic listening experiment in which they listened to two recordings simultaneously and their task was to shadow a specific ear. Loudness and duration of the two passages in the same trial were controlled. The loudness for all the recordings was set at 70 dB using PRAAT. The software AUDACITY (Version 2.3.0; Audacity Team, 2018) was used to edit the passages' duration and merge two mono tracks into one stereo recording. All the sentences were manipulated to ensure that the characters were played simultaneously controlled within 0.032 seconds time difference. A 1.18 second-pause took place between trials, giving subjects sufficient time to finish shadowing the previous sentence. The entire stereo-recording with 16 trials had a total length of one minute.

doi:10.6342/NTU202000533

3.3 Data Analysis

A points system was designed to record the accuracy of the shadowing results. One point was awarded to the subject if the speech matched the passage of the designated ear. On the other hand, if the subject did not shadow any characters from the assigned passage, or only shadowed part of the keywords, then no points were awarded. For example, if a subject was exposed to the passage pair: "東京航空明年將結束營業/京都航空明年將結束營業" （Tokyo Airline will stop operating next year/Kyoto Airline will stop operating next year) and was asked to shadow the ear in which the passage was "東京航空," then the subject would only receive one point if his or her response matched the keyword "東京." Any other answers such as 京都,東都, silence, or nonsense word that did not match the keyword would be regarded as failing to shadow the passage. The total number of points awarded was then calculated. Finally, a two-sample assuming equal variances t-test was used to determine whether subjects' performance score was affected by intonation.

Data analysis for accuracy was assessed by counting the number of answers which matched the keyword of the designated passage. If the subject's response matched the keyword, indicating he or she was shadowing from the right ear, the subject received 1 point. Any other answer which did not match the keyword would not be awarded any points. Subjects shadowed 16 pairs of passages and received a score between 0 and 16.

3.4 Subjects

doi:10.6342/NTU202000533

132 subjects were recruited from National Taiwan University in Taipei, Taiwan. Ninety-eight subjects are students from a course named Online English Program (OEP), twenty-three subjects are students attending a freshmen English course, and eleven subjects are students from an education practicum class. Data was incomplete for 26 subjects due to technical problems. 3 subjects were not included in the results because their native language was not Mandarin Chinese.

A total of 103 subjects' data were collected from four groups: Group ML (21), Group MM (34), Group LM (23), Group LL (25). All subjects had good hearing and no psychological illness would affect their reactivity to the stimuli. Since students with interpreting training would be able to shadow from a specific ear without being affected by the other ear, all subjects are students with no prior interpreting experiences.

3.5 Hypothesis

There were two general hypotheses concerning the performance of the dichotic test. The first hypothesis is that intonation would lower the threshold of the attention filter (Broadbent, 1958; Cherry, 1953; Cutler, 1997; Treisman, 1960) making the lively sentences easier for subjects to recognize and shadow. Subjects shadowing the lively ear while listening to the monotonous ear would score higher than other groups since it is easier to focus on the lively sentences and neglect the monotonous sentences. On the contrary, subjects listening to the lively sentences while shadowing the other ear would face most distraction since signals from the designated ear is more likely to be attenuated, allowing the lively sentences to pass through the filter. Therefore, it is hypothesized that Group LM

doi:10.6342/NTU202000533

and Group ML will have the highest and lowest scores respectively. Subjects listening to either two lively or two monotonous passages would be exposed to two stimuli with equal chances of passing the attention filter, resulting in no significant difference between the scores of Group LL and Group MM. It is predicted that the scores of the 4 groups would rank as follow: LM > LL = MM > ML. If the hypothesis is accepted, we could conclude that intonation, as a contextual cue, can lower the threshold of the attention barrier and that the audience will more likely pay attention to a livelier SI instead of a monotonous SI.

The second hypothesis relates to the scores of the emotional texts. Since intonation facilitates the passing of information through the selective filter due to listeners' "expectancy based on transition probabilities" (Treisman 1960, p.1), it is assumed that a lively intonation aligned with the context of the speech would be easier for listeners to process and therefore is more likely to be shadowed. The scores of the lively emotional texts would expect to be higher than the scores of the lively non-emotional texts (including both statements and questions). On the contrary, the scores of monotonous emotional texts will have the lowest score due to the mismatch between the intonation and the context. If the hypothesis is accepted, it suggests that intonation, as a contextual cue, plays a more important role in emotional texts compared to the non-emotional texts. This would highlight the audience tendency to listen more attentively to emotional texts than to the non-emotional texts when both were delivered in a lively manner.

3.6 Results

doi:10.6342/NTU202000533

This chapter presents the results testing the two hypotheses. Performance was assessed by counting the total number of responses which matched the correct keyword, henceforth, “accuracy”. Section 3.6.1 lays out the performance of the four subject groups (LL, MM, LM, ML) who participated in the dichotic test, shadowing 16 pairs of passages. Section 3.6.2 examines the accuracy results of the emotional and the non-emotional texts. Section 3.6.3 scrutinizes the common mistakes and the questionnaire completed by subjects after the experiment.

3.6.1 Group scores for dichotic test

The group average scores are shown in Figure 4. An ANOVA test showed significant difference between the scores of the four groups ($F(3,110)=5.923$, $p < 0.001$). As shown in the figure, Group LM received the highest average score ($M=0.86$, $SD=0.37$), followed by Group MM ($M=0.82$, $SD=0.2$), Group ML ($M=0.72$, $SD= 0.19$), while Group LL received the lowest score ($M= 0.67$, $SD= 0.11$). This result rejects the null hypothesis of the 4 group scores ranking as LM > LL = MM > ML.

In addition, the post hoc t-tests were conducted to observe the between group differences. The results revealed significant difference between LM and LL ($t(46) =4.437$, $p <0.001$) as well as between ML and MM ($t(53) =-1.6997$, $p <0.05$), but not between ML and LM.

doi:10.6342/NTU202000533

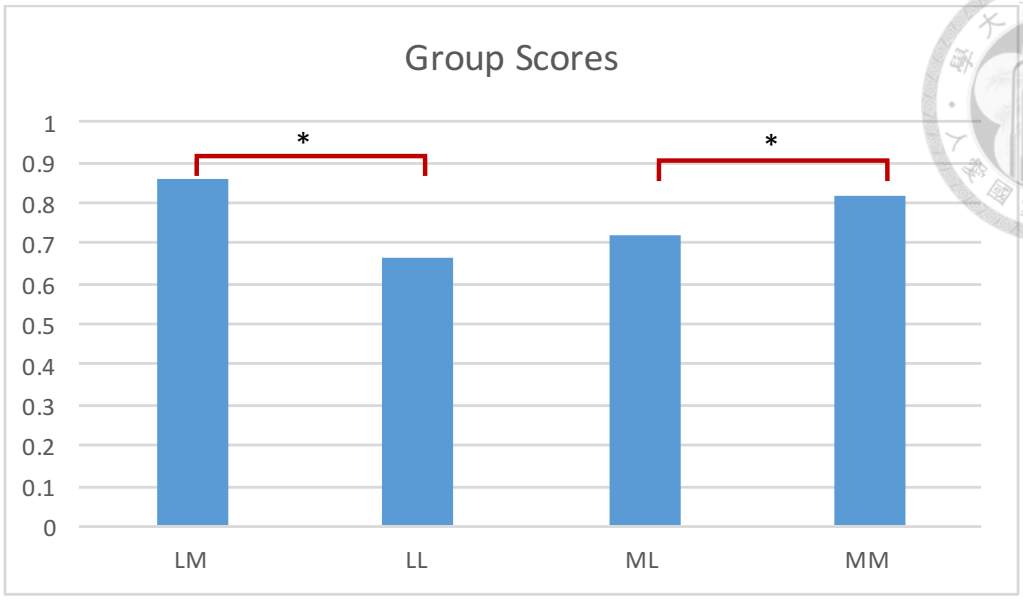


Figure 4. Average scores for Group LM, LL, ML, and MM.

Figure 5 and Figure 6 demonstrate the groups whose scores were significantly different from each other.

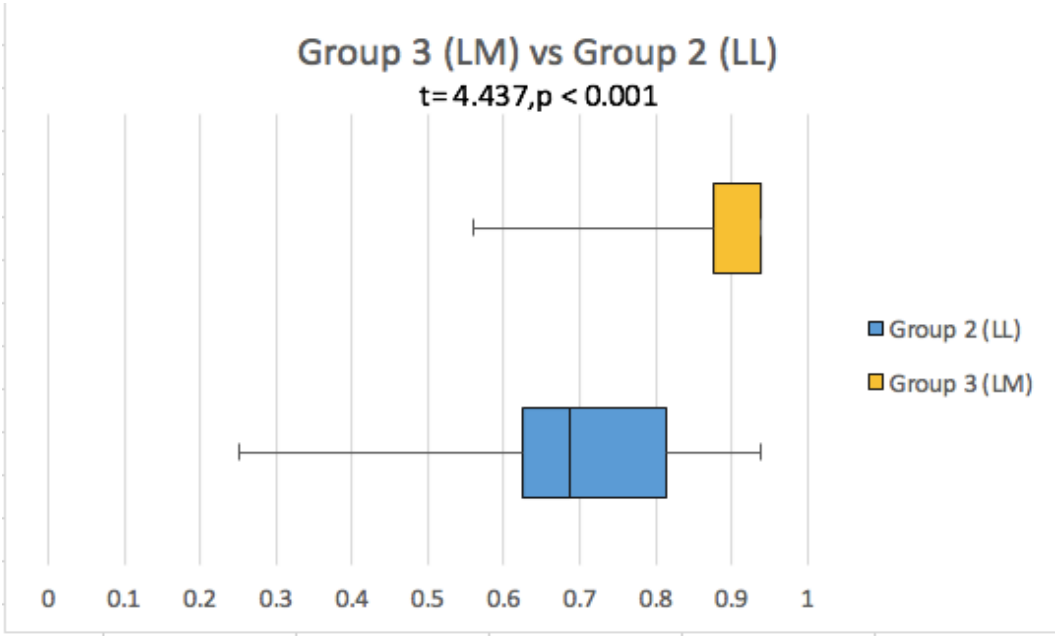


Figure 5. Average scores for Group LM and Group LL.

doi:10.6342/NTU202000533

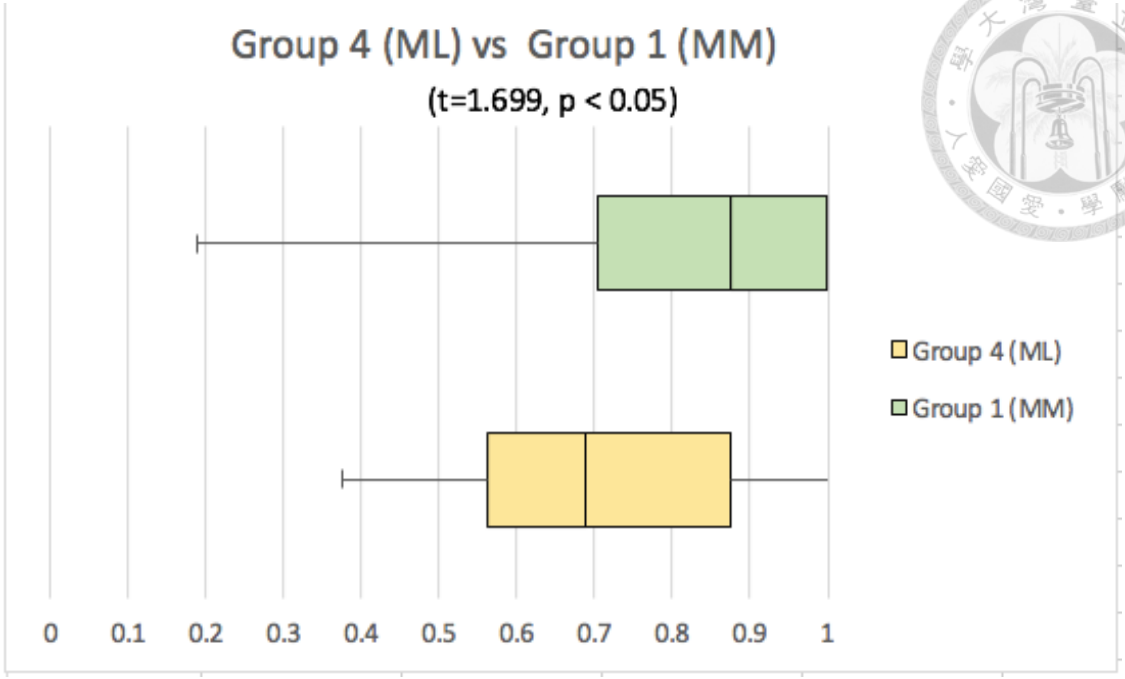


Figure 6. Average score of Group ML and Group MM.

The findings that there was no significant difference between the scores of Group ML and Group LM, and that the scores of Group LL was the lowest among the four groups reject the null hypothesis which predicted that Group LM and Group ML would have had the highest and the lowest scores respectively and that Group LL and Group MM would have shared a similar score.

3.6.2 Group Scores for Emotional Texts and Non-emotional Texts

In this section, the scores for emotional texts and non-emotional texts are examined separately. The general pattern in scores of these two texts are demonstrated in Figure 7. It shows that the average score of emotional texts is higher than the score of the non-emotional texts (including both statement and question sentences) in all four groups. For the emotional texts, Group LM show the highest score (92%) followed by Group MM (84%), Group ML (77%), and Group LL (76%). For the non-emotional texts, Group LM

doi:10.6342/NTU202000533

show the highest score (79%) followed by Group MM (78%), Group ML (67%), and Group LL (62%).

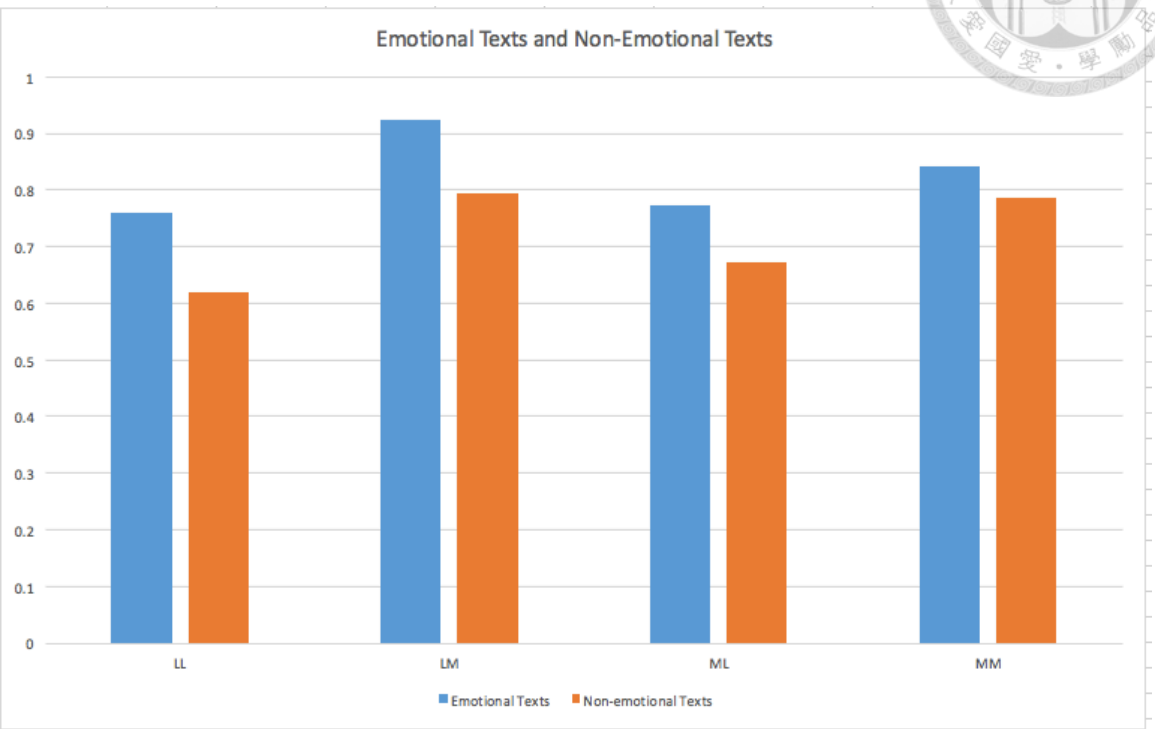


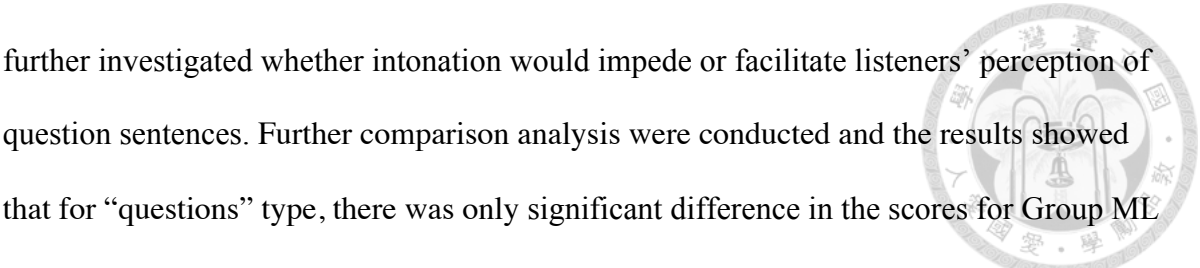
Figure 7. Average group score of emotional texts and non-emotional texts.

Non-emotional text

Statistical results of the non-emotional text scores (i.e., the blue bars in Figure 7) showed that there is a significant difference between the experimental groups (LM, ML) and the control groups (LL, MM). A series of t-tests was conducted and the result revealed significant difference between Group LM and Group LL ($t(382) = 3.775$, $p < 0.001$) and between ML and Group MM ($t(438) = -2.6782$, $p < 0.01$).

There are two sentences types in the non-emotional text, namely "statements" and "questions". Unlike English, statements and question sentences often have the same structure in Mandarin Chinese. Since statements are often changed into questions by simply adding a sentence particle or raising the intonation at the end of the sentence, the researcher

doi:10.6342/NTU202000533



further investigated whether intonation would impede or facilitate listeners' perception of question sentences. Further comparison analysis were conducted and the results showed that for "questions" type, there was only significant difference in the scores for Group ML and Group MM ($t(218)=0.172$, $p<0.05$) (see Figure 8). As for the "statement" type, significant differences were found between Group LM and Group LL ($t(190)=4.342$, $p<0.01$) and between Group ML and Group MM ($t(218)=3.578$, $p<0.01$) (see Fig.9).

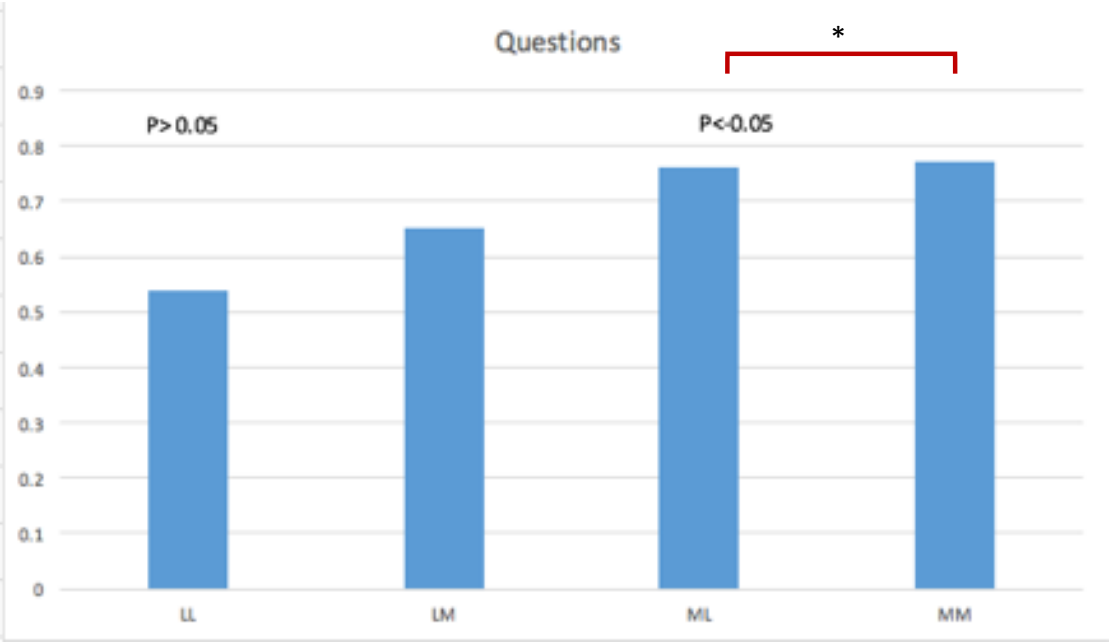


Figure 8. Average group score of question sentences.

doi:10.6342/NTU202000533

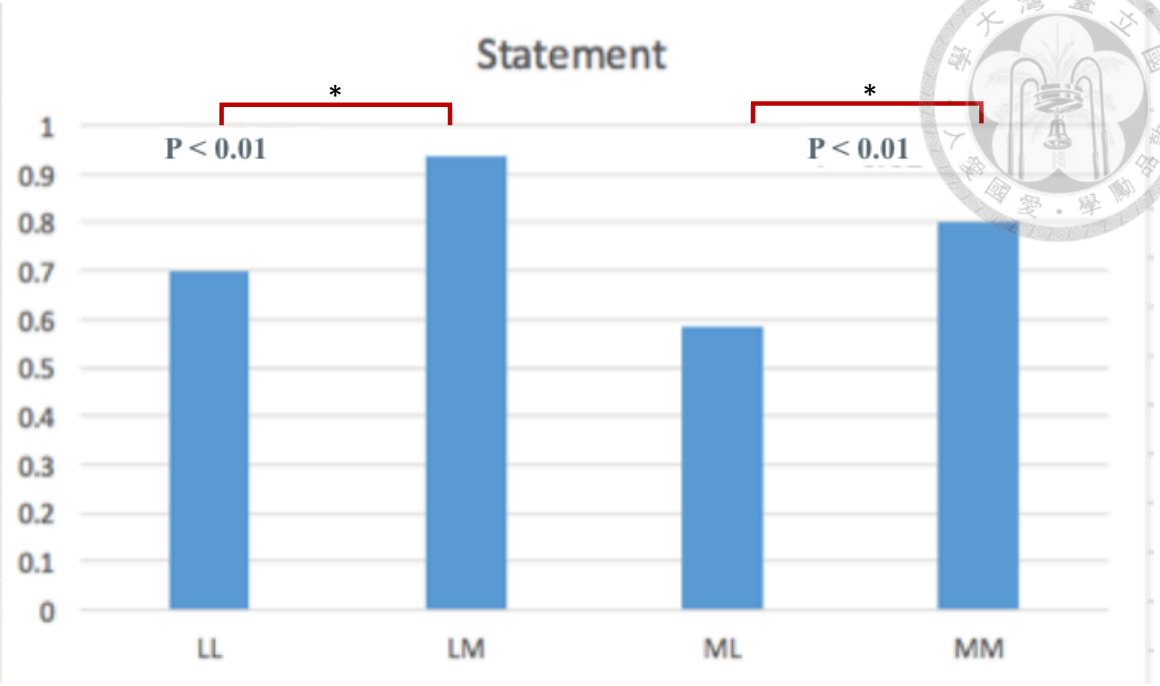


Figure 9. Average group score of statement sentences.

Emotional texts

Statistical results showed that there was a significant difference between the two experimental groups (LM, ML) and the control groups (MM, LL). There was a significant difference between Group LM and Group LL ($t(382)=4.461$, $p<0.001$) and between Group ML and Group MM ($t(438)=-1.793$, $p<0.05$.)

doi:10.6342/NTU202000533

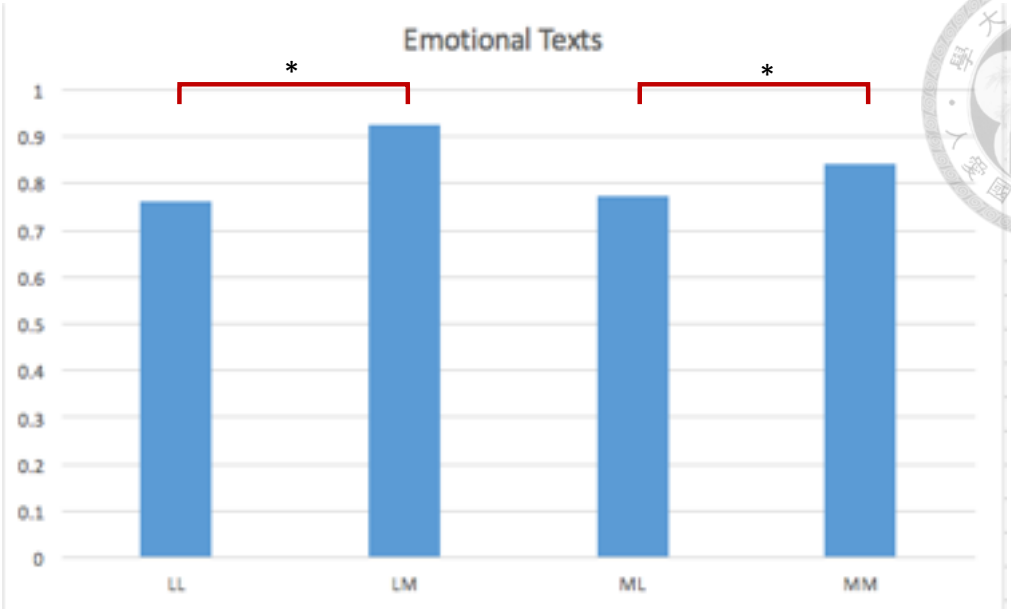


Figure 10. Average group score of emotional texts.

In order to see if scores of a particular emotion affected the results of the emotional text in general, scores of the four emotions were examined individually below.

Joy

A further analysis on the different sentences within the emotional texts indicated that only the LL and LM group showed significant score difference in terms of sentences with the emotion joy. Test revealed the first group's result to be Group LM (M=0.93, SD=0.44) and Group LL (M= 0.76, SD= 0.24); t (94) =2.42, p <0.01 while the result of the second group revealed no significant difference; Group ML (M=0.73 , SD= 0.43) and Group MM (M=0.80, SD=0.39); t (108)=-0.867, p> 0.05 (see Fig. 11).

doi:10.6342/NTU202000533

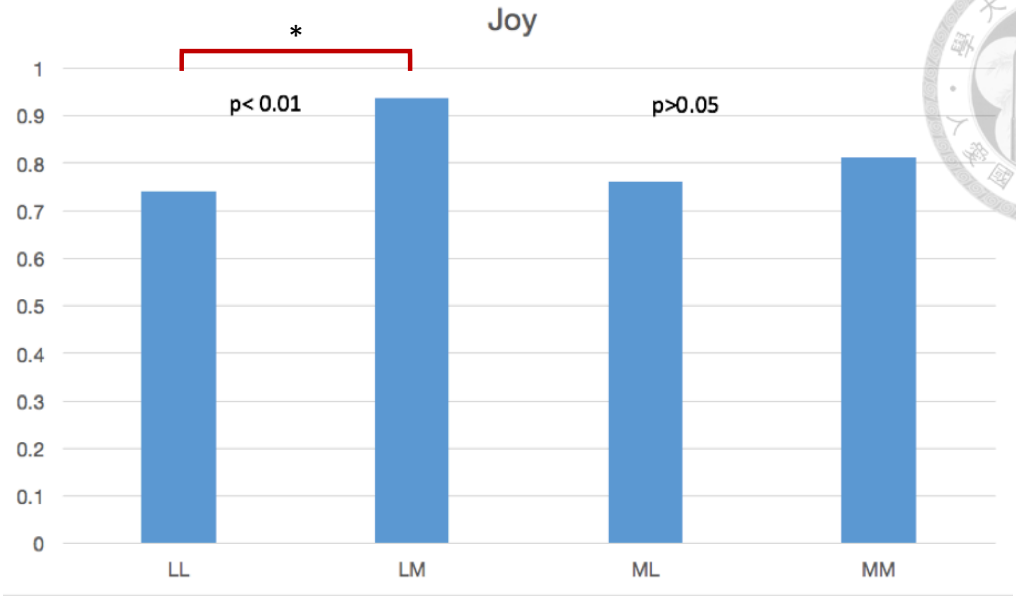


Figure 11. Average group score of emotional texts with the emotion joy.

Sadness

The results were also similar for sentences with the emotion sadness, scores indicated that only the LL and LM group showed significant score difference, Group LL (M= 0.82,SD= 0.38) and Group LM (M=0.93, SD=0.24); t (94)=1.71, p <0.05; while Group ML and MM, Group ML (M=0.9 , SD= 0.29) and Group MM (M=0.88, SD=0.32); t (108)=0.36, p> 0.05 showed no significant difference (see Fig. 12).

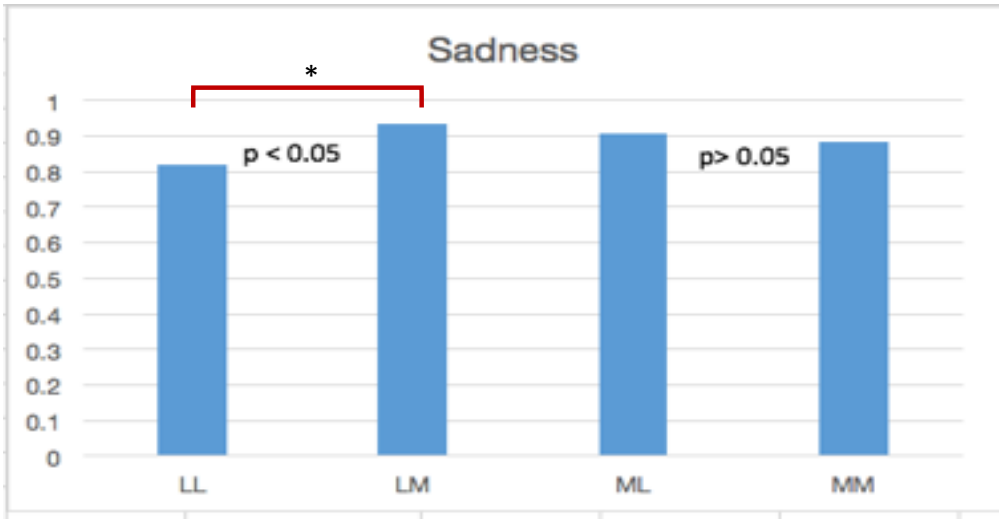


Figure 12. Average group score of emotional texts with the emotion sadness.

doi:10.6342/NTU202000533

Fear

As shown in Figure 13, there were no significant difference in the two control and experimental groups for sentences with the emotion fear. Group LL (M= 0.74,SD= 0.44) and Group LM (M=0.82, SD=0.38); t (94)=1.01, p >0.05; Group ML (M=0.76 , SD= 0.43) and Group MM (M=0.86, SD=0.34); t (108)=1.42, p> 0.05.

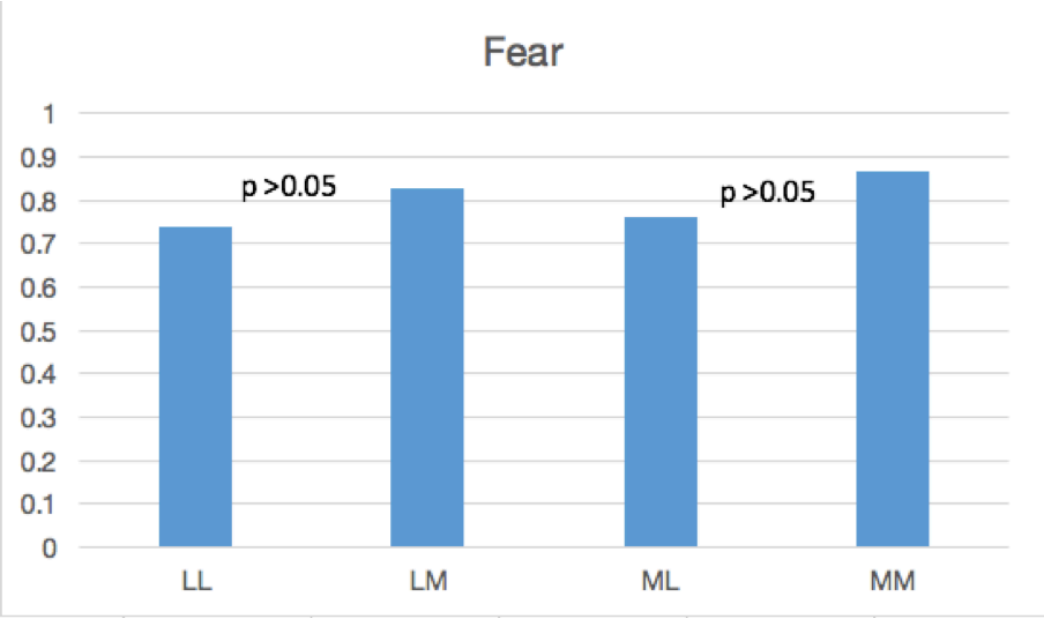


Figure 13. Average group score for emotional texts with the emotion fear.

Anger

For sentences with the emotion anger, the results were extremely significant for Group LL (M= 0.72,SD= 0.45)and Group LM (M=1, SD=0); t (94)=4.18, p <0.001; and insignificant for Group ML (M=0.69 , SD= 0.46) and Group MM (M=0.80, SD=0.39); t (108)=1.41, p> 0.05 (see Fig. 14).

doi:10.6342/NTU202000533

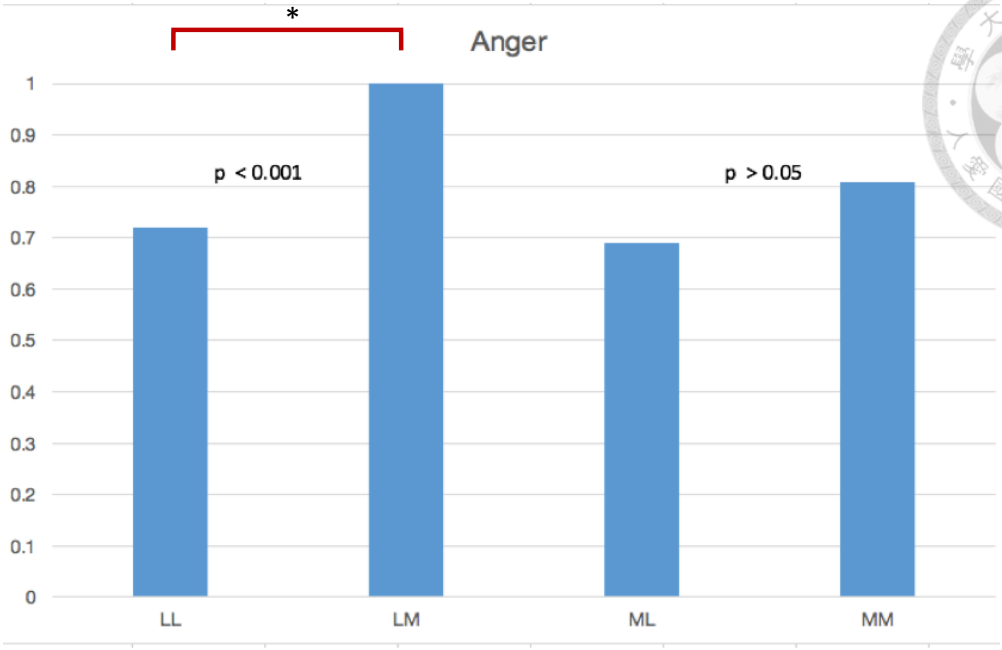


Figure 14. Average group score for emotional texts with the emotion anger.

3.6.3 Observed common group mistakes

Group LM

Subjects from Group LM all unanimously shadowed "**你們什麼時候開始呀?"** (When did you two begin?) when exposed to the passages "**你們什麼時候開學呀?"** and **"你們什麼時候吃完呀?"** Two subjects left the first four questions blank but then accurately shadowed every question afterwards. This group performed significantly better than all the other groups, with subjects responding more instantaneously, accurately, and confidently. It was also observed that some subjects would even mimic the emotion of the passage which they were shadowing.

Group LL

Subjects from Group LL also responded saying, "你們什麼時候開始呀?" (When did you two begin?) when exposed to the passages "你們什麼時候開學呀?" and "你們什麼時

doi:10.6342/NTU202000533

候吃完呀?" similar to Group LM. In addition to this shared mistake, it was observed that subjects merged the keywords "前頭"(ahead) and "十樓" (tenth floor) into "前樓" or the keywords "跳繩" (jump rope) and "作文" (writing) into "跳文". Subjects would often start shadowing the attended ear yet eventually switch to the wrong ear; in addition to this observed interference of the unattended messages, most of the unmatched responses were silent responses or nonsense words that did not pertain to keywords from either passages.

Group ML

The phenomena of failing to ignore the unattended message was most commonly observed in Group ML as most unmatched answers were either keywords from the wrong ear or merged nonsense keywords. This Group seemed more susceptible to the stimuli of the irrelevant channel as subjects took more time to offer their responses and would sometimes hesitate in the middle of their speech. Various subjects from this group appeared unaware of the fact that they were shadowing from the wrong ear. Despite efforts to focus on the designated ear, they would switch back to the irrelevant lively channel after accurately shadowing one or two monotonous passages.

Group MM

Little mistakes have been observed in this group and most of the unmatched responses were keywords from the wrong channel or words that did not belong to either of the passages. For example, one subject responded saying, "森林裡住著一隻老虎" (In the forest lives a *tiger*) instead of saying 獼猴 (Macaque) or 狐狸(Fox). Another subject said, " 我今天挑選比賽得了第一名" (I won first in a *selection* competition) instead of saying 跳繩(Jump rope) or 作文 (Writing). It is interesting to point out that the mistake of merging keywords together was not observed in this group.

doi:10.6342/NTU202000533

3.6.4 Part 1 questionnaire results

All 103 subjects who participated in the experiment answered a post-experiment questionnaire of twenty questions (see Appendix 3) regarding their basic information and their general reflection of the dichotic test. Subjects were asked to select the passages they found most difficult to shadow, explain their decisions, answer whether they found the task difficult and if they believe intonation affects their listening attention. Ninety-two subjects agreed that the variance of intonation or prosody affects their will to listen while eleven subjects disagreed (see Fig.15). The following paragraphs present the survey results for Part 1 of the respective groups.

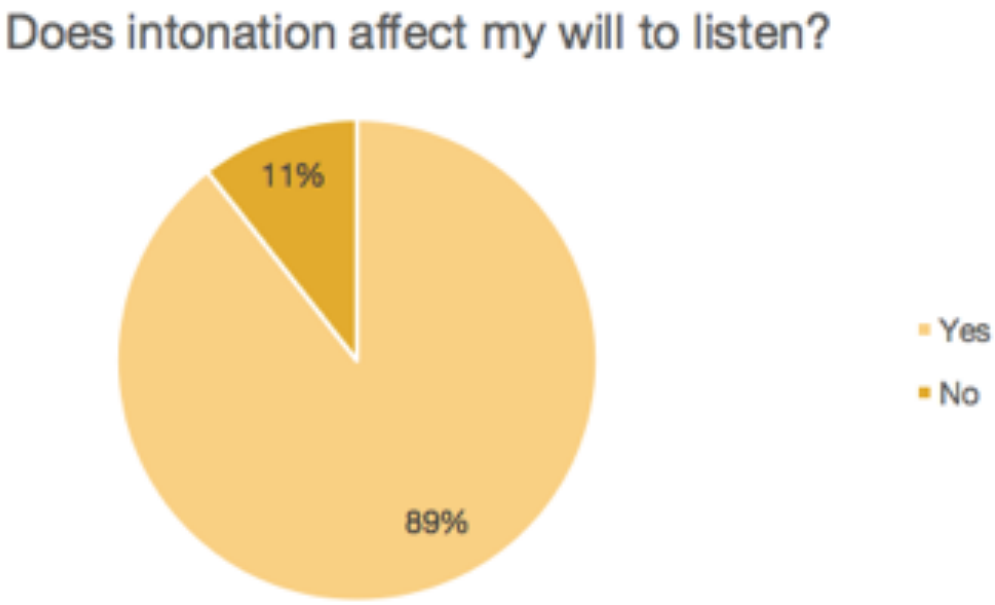


Figure 15. Subjects' responds regarding intonation and their will to listen.

Group LM

doi:10.6342/NTU202000533

More than two third of the subjects responded saying that they found no difficulty listening or shadowing the right ear since they could easily differ the two sound tracks due to the speaker's distinct voice quality. The same subjects also commented saying that they found the lively ear more pleasant to listen to and it was very natural and intrinsic for them to shadow the accurate ear. When asked which sentences where most difficult to shadow, nearly one fourth of the subjects selected the option "none," while various answered saying that they found question sentences like 請問公司是在前頭/十樓嗎？ *(*Is the company *ahead/ on the tenth floor*) or "你們什麼時候開學呀/吃完呀?" (When will you *begin school*/ *finish eating*?) with more difficulty due to the phonetic similarity between the two keywords. The two subjects who left the first four questions blank responded saying that they were unable to speak due to nervousness at first yet they could clearly hear the right ear after a short period of time. When asked which sentence Group LM found easiest to shadow, subjects answered that sentences with emotions or longer sentences were preferable. Many responded saying that sentences with intense emotions of sadness and anger were fun and easier to shadow because the speaker's tone enhanced the difference between the two ears, allowing them to clearly recognize the responding ear. In addition, longer sentences were also less challenging because it gave them more time to digest the context and a few subjects said that they somehow already knew how certain sentence would finish.

Group LL

Slightly more than half of the subjects in this group agreed that both identifying the right sound track and shadowing were difficult tasks. Overall, most subjects answered saying that for some reason, the phrases sounded too identical and that the keywords

doi:10.6342/NTU202000533

sounded all “mixed up,” making it hard for them to react on time. Many subjects also replied feeling constantly distracted and disturbed or not being able to hear clearly during the experiment, especially when they were asked to shadow longer sentences. Similar to subjects in Group LM, subjects in this group found question sentences most difficult to shadow due to phonetic similarity. When asked which passages were easiest to shadow, many subjects selected the passage “哇！這個地方讓人覺得好浪漫！/快樂！”(Wow! This place is so *romantic/pleasant*) explaining that the short pause in the middle of the sentence allowed them to take a short breath, dividing the sentence into a shorter phrase, permitting them to prepare for the rest of the sentence.

Group ML

All the subjects in Group ML agreed that listening to a specific ear is not an easy task and fifteen subjects agreed that shadowing was an uneasy task. No obvious trend was observed from subjects’ responses of which sentences where easier to shadow. When asked which sentences were most difficult to shadow, the passages “跟你說一個好消息，我今天跳繩/作文比賽得了第一名！” (I won first in a *jumping/writing* competition), and 你怎麼可以那麼過分！/ 任性！(How can you *go so far/ be so stubborn*?) were most selected. Subjects stated that it was difficult to shadow the same ear because certain passages were spoken at a faster pace or sounded louder than the other ear. Many subjects also responded saying that the other ear sounded more attractive and therefore it was very disturbing.

Group MM

With the exception of five subjects, all participants in Group MM responded saying that they found no difficulties focusing on a specific ear. However, half of the respondents

doi:10.6342/NTU202000533

agreed that shadowing was a challenging task. In terms of passages which was easiest to shadow, nearly half of the subjects selected the option “none” and explained that as long as they remained focused and paid attention it was easy to distinguish the two ears. Subjects in this group also said that shadowing was easier when the sentences were shorter or more commonly used. It is also interesting to point out that many subjects selected passages with the emotion fear or sadness saying that sentences with more emotion and tone or conveying a negative feeling were easier to focus on.

doi:10.6342/NTU202000533

Chapter Four Part 2

Part 2 was a questionnaire focused on interpreters, investigating their perception towards intonation's role in Mandarin Chinese SI and whether interpreters, as listeners, are also affected by monotony. The purpose of this quantitative experiment is to acquire a general understanding of practicing interpreters' subjective evaluation of: (1) intonation's role in their interpreting tasks (2) the effect of speakers' intonation on their performance and (3) the strategies and tactics interpreters apply when interpreting a monotonous speaker. This quantitative study also serves as a follow up experiment to explore whether the findings and theories proposed from the results of Part 1 were in line with real-life practice and to what extent it may benefit future interpreting practice and training.

4.1 Method

4.1.1 Procedure

The questionnaire was created using google survey and the invitation of the survey was sent via Facebook. The survey was sent to potential subjects and they accepted the invitation by clicking the google form link to answer 16 questions online. Subjects participated in the experiment voluntarily and received no rewards. The completion of the survey took about 10 minutes.

4.1.2 Material

The questionnaire contained sixteen questions asking respondents about their basic background, and how they perceived the essential elements in interpreting such as prosody,

doi:10.6342/NTU202000533

language, and accuracy. The second part of the questionnaire focused on interpreters' perception of monotony, the strategies they adopt when interpreting monotonous speakers, and whether they believe intonation should be included in interpreting training. The questionnaire consisted mainly of closed questions or Likert scale questions, but also included open-ended questions such as:

1. If the speaker is conveying an emotional message (sounding joyful, sad, angry, or scared), I will try to deliver my interpretation in a way that reflects the according emotion because…

2. I do/ do not find it more difficult to interpret a monotonous speaker because....

3. What are some strategies that I apply when interpreting a monotonous speaker?

For full questionnaire please refer to Appendix 4.

4.1.3 Subjects

A total of 29 interpreters completed the online questionnaire. Seventeen responses were received from the experienced interpreters' cohort while twelve responses were received from the novice interpreters' cohort. From the experienced interpreters' cohort, more than 47.1 % of respondents had more than 15 years of practicing experience, 29.4% had more 10-15 years of experience, and 23.5% had 5-10 years of experience. On the other hand, all respondents from the novice interpreters' cohort had less than 5 years of interpreting practice experience, 33.3% had 1-2 years of experience, 41.7% had 2-3 years of experience, and 25% of the novice interpreters had 3-4 years of experience.

doi:10.6342/NTU202000533

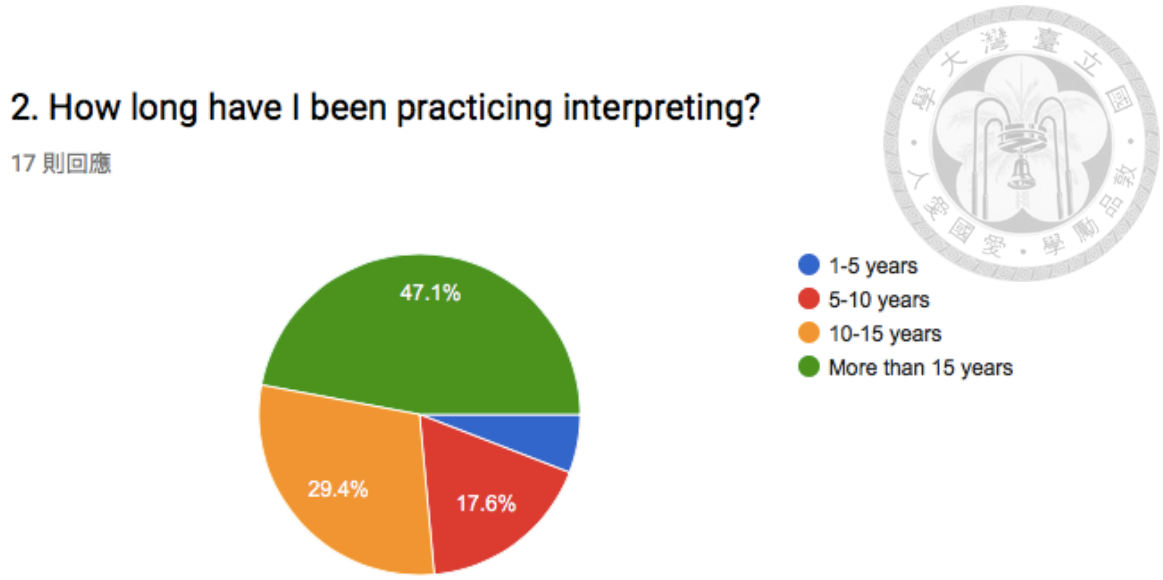


Figure 16. Experienced interpreters' years of practice.

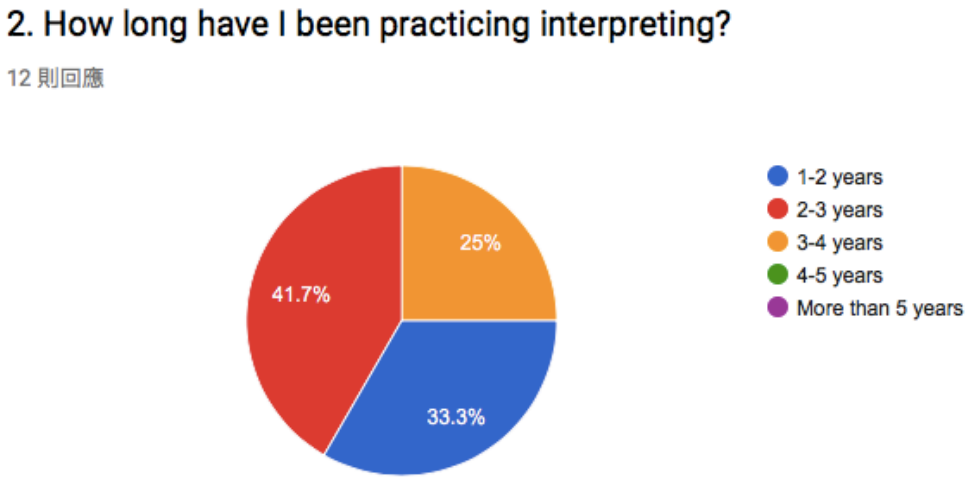


Figure 17. Novice interpreters' years of practice

doi:10.6342/NTU202000533

4.2 Results

4.2.1 Interpreters' View on Intonation's Role in Mandarin Chinese SI

When asked to rate essential elements such as accuracy, language, and prosody (intonation) in SI on a five-point Likert Scale (one being strongly disagree and five being strongly agree), experienced interpreters rated an average of 4.94, 4.52, and 4.47 respectively while novice interpreters rated an average of 4.75, 4.5, and 4.25 respectively (see Fig.18).

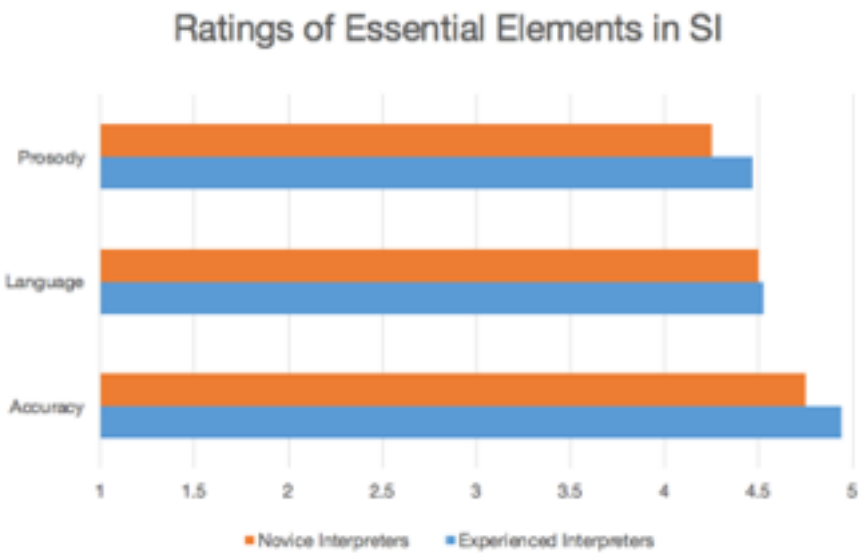


Figure 18. Ratings of Essential Elements in SI.

When asked whether a lively intonation was crucial for good speakers and whether interpreters should deliver in a way that conveys the speaker's emotion, experienced interpreters gave higher average scores than novice interpreters and rated 4.58, 4.23, 4.47 out of five respectively while novice interpreters rated a higher 4.16, 3.83, 4.41 respectively (see Fig.19).

doi:10.6342/NTU202000533

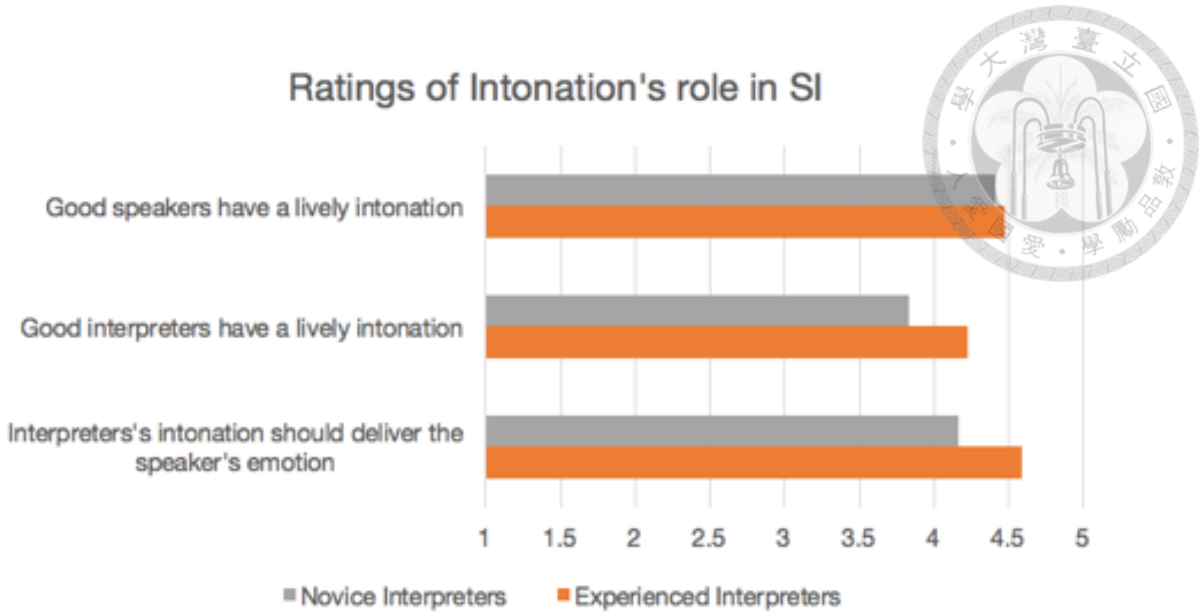


Figure 19. Ratings of Intonation's role in SI.

When asked to explain the reasons why one would or would not deliver interpretation in a way that conveys the speaker's emotional message, most experienced interpreters' responses emphasized the importance of using intonation to achieve authenticity, communication effect, and delivery efficiency. Five subjects answered saying that emotion is part of the message and that non-verbal messages must be delivered in addition to the verbal ones. Subject E13 simply stated that non-verbal message is key to full comprehension while Subject E7, an interpreter with more than 15 years of experience, responded saying:

> *"Interpreters are vessels for knowledge/messages to flow through and serve the audience. Verbal communication only accounts for 30%, with the other 70% of the messages delivered by non-verbal communication" (E7).*

doi:10.6342/NTU202000533

Similarly, Subject E5 and Subject E4 responded saying, *"Because that's part of the message and also a language in and of itself,"* and "*How you express message is just as important as what you say,"* respectively.

Subjects E1, E3, and E12 explained that intonation helps create a communicate effect on the audience; while Subjects E10 and E6 believe intonation helps deliver in a more efficient manner. Subject E10 responded, *"words alone can be deceiving and meaning is often comm*unicated *in the tone*," Subject E6 answered stating, "[intonation serves] *As a means to get the ideas across more effectively".* Unlike other experienced interpreters, Subject E*2,* responded saying *"Depends on the context and purpose of the speech and occasion".*

The responses received from novice interpreters were very similar to the answers provided by experienced interpreters. Subjects N3, N6, N9, N10, and N12 all support the notion that emotions are a part of the message and that interpreters should render them faithfully because they are relevant to the delivered content. According to Subject N12:

> *"Emotions are part of the message of the speaker, not conveying them would be like missing part of the message. For example, if the speaker is very upbeat and trying to sell something or get people excited about an event, if the interpreting is monotonous, then the audience will just be bored" (N12).*

On the other hand, Subjects N2, N6 and N11 believe that non-verbal languages should be used differently depending on the occasion. Subject N6 commented saying:

> *"Because emotions are a part of the message that the speaker tries to get across to the audience. But when the speaker gets too emotional, the interpreter should tone down a bit so as not to disrupt a clear delivery" (N6).*

doi:10.6342/NTU202000533

Similarly, Subject N2 responded saying, *"It depends on the occasion. Some occasions intonation or emotion is not to be rendered fully (e.g. medical interpreting)."* and Subject N11 said, "[It is] *Not always appropriate because emotions transcend languages. It would be really jarring to those listening to the interpretation.*".

On the other hand, Subjects N1 and N7 stated that audiences who rely on interpretation have the right to experience the same emotions as those who do not interpretation services; thus an intonation that is consistent with the source text is necessary. Subject N8 emphasized interpreters' role as speakers and the priority of intonation in SI:

> *"Interpreters are also speakers. They need to be accurate in delivery and able to sustain their audience's attention through their voice and intonation. Sentence structure of TL may be affected by SL, which may make the sentence sound slightly awkward, but as long as it does not change the meaning, perfect syntax is not necessarily a primary concern" (N8).*

Subject N10's answered, "*Though the audience can't understand source language, they can detect the speaker's emotion. Not following could mean bad interpretation to some audience,"* emphasizing how intonation helps maintain the perceived authenticity of the translation.

4.2.2 Monotony's Effect on Interpreters During SI

When subjects were inquired whether it was more challenging to interpret a monotonous speaker than a lively speaker, 94.1% of the experienced interpreters and 83.3% of the novice interpreters agreed. Subjects from both groups responded saying that the lack of intonation makes it harder for them to grasp the essence of the speech and that it requires

doi:10.6342/NTU202000533

extra attention in order to identify the "key message point" (E3, E6, E17, N2, N3). Subjects E6 and E17 both pointed out that when speakers lack intonation, all words are given equal weight, thus making it hard for one to stay focused on the important information. Similarly, Subjects E13, E14, and N7 responded saying that the lack of non-verbal cues means that they can only focus on verbal language, and must pay attention to every single word they hear in order to decipher the context of the speech. Monotony also makes it more challenging for subjects to anticipate the information and direction of the speech, making it harder for them to "be in sync" or "to get into the flow" while interpreting (E1, E12, N4, N5, N12). Subject N7 also responded by saying that monotony interferes with her understanding and processing of the speech. In addition, a monotonous speaker is also harder for interpreters to stay focused and engaged when the speaker is monotonous, resulting in boredom and distraction (E5, E7, E8, E10, N9). According to Subject E10, "*It is harder to focus for a long stretch of time on someone who speaks like a robot,"* and Subject E5 states that *"Monotonous deliveries cause audiences to zone out and it's the job of interpreters to keep their audience engaged".*

4.2.3 Strategies Applied when Interpreting Monotonous Speakers

When asked what strategies are adopted when interpreting monotonous speakers, Subjects E6, E13, E14, and N4 said that they would distribute more effort into listening. On the other hand, Subjects E3, E4, E9, N5, N8 would rely on other contexts such as slides, the structure of the speech, and the speaker's logic or stance in order to anticipate the upcoming information. Subjects N7 would focus on the speaker's body language, Subject N2 would visualize the content in order to stay engaged, and Subjects E7 would rely on eye contact. Subjects N1, N3, and E11 responded saying that they would try to enhance their level of

doi:10.6342/NTU202000533

focus and pay more attention. Some subjects would maintain or adjust their voices when interpreting monotonous speakers, for example, Subject E2 would speak in a softer manner in order to listen better, Subject N6 would maintain her usual intonation, and Subject E10 responded saying "I try to speak in a livelier manner myself to stay focused."

4.2.4 Should Interpreting trainers focus more on voice training?

When asked whether interpreting trainers should focus more on voice training and the importance of prosody, 58.8% of the experienced interpreters agreed, while 41.2% disagreed; and the number of those who agreed and disagreed were equal among novice interpreters. Subjects N7, E1, E2, E3, E13, E16 agreed that intonation related training is necessary since intonation is an important part of delivery. Subject E10 answered, *"A lively and rhythmic delivery inspires confidence and interest in the listener,"* and Subjects N2 and N7 also agreed that a listener-friendly approach to interpreting could enhance the engagement of the audience. Similarly, Subject E14 commented, "*When the audience stops listening to interpreters due to monotony, even if the content is accurate, it probably wouldn't make much of a difference, because the message is not delivered".* In addition, subjects also pointed out the importance of intonation to clients and the general audience,

> *"In actual practice, clients often pay more attention to the quality of interpreters' voice rather than the content of the message. So one of the reasons voice training is important is to help interpreters better manage their professional image and to meet clients' expectations (E4)."*

Subjects N8 and E12 also pointed out that the market's emphasis on intonation means that if interpreters can offer pleasant delivery, they are deemed more valuable and are more likely to receive positive feedbacks.

doi:10.6342/NTU202000533

On the other hand, those who disagree stated that despite the importance of intonation, trainers should still prioritize the training of accuracy over prosody due to the limited amount of time (Subjects E6, E8, E9, N5, N6). Subjects N3 and N4 responded that it is not necessary for trainers to focus on intonation, since the manner of expression cannot be trained. On the other hand, Subjects E15 and E5 believe that prosody and voice training is another area of expertise and it is best for experts to teach.

doi:10.6342/NTU202000533

Chapter Five General Discussion

In this chapter, the researcher will summarize and discuss the results reported above, address the research questions of the study and also present the implications of the findings.

5.1 Traffic Jam Theory: How Intonation Affects Listeners' Attention

In 1960, Treisman's Attenuation Theory proposed the idea that when two signals simultaneously entered one's auditory system, the signal which carried more contextual cue would be prioritized, reaching the working memory, while the other signal would be attenuated and less likely to be further processed. Treisman included physical proprieties, diction, and relevance as factors that acted as contextual cues. However, how intonation, as a contextual cue, would affect attention was never studied in previous literatures. In order to study whether intonation, as a contextual cue, would enhance or weaken one's attention, the researcher carried out an experiment with four groups of subjects.

It is observed that Group LM (86%) had the highest average score, followed by Group MM (82%), Group ML (72%), while Group LL received the lowest score (67%) among the 103 subjects. The experiment results were significant for the two control and experimental groups Group LM and LL ($p < 0.001$) and Group ML and MM ($p < 0.05$). More specifically, the fact that LM accuracy is greater than LL shows that when the inputs contained two distinctive speeches (monotonous and lively), it is easier for the listeners to shadow the lively ear and ignore the monotonous ear. This indicates how intonation, as a contextual cue, does help lower the threshold of listeners' attention filter and facilitate the passing of stimuli while the lack of intonation obstructs the process.

doi:10.6342/NTU202000533

This experiment proves how intonation may serve as a deciding factor along with semantic, words, syntax, and relevance, in determining whether a signal can pass through the attention filter and enter one's working memory (Treisman 1960). In other words, a passage which is more intoned, carrying more contextual cue, is therefore more likely to be paid attention to and listened to by listeners.

The following paragraphs presents the Traffic Jam Theory, a concept proposed by the researcher to explain the different ways intonation may affect listeners' attention based on the data collected from Part 1. The Traffic Jam Theory compares the attention filter as a road junction that decides the speed and direction which the signal is processed based on the amount of contextual cue it carries. It is proposed that signals with more contextual cues are granted access to a "highway" which rapidly leads to the working memory while signals with insufficient cues are attenuated, heading to a "dead end street," making it less likely to be further processed. However, when the attention filter is overloaded with signals, a traffic jam occurs, impeding the signals to efficiently reach the working memory system.

Lively channel vs. Monotonous channel (Group LM)

When subjects are exposed to two sound channels, one being lively and the other being monotonous, stimuli from both ears will pass through the sensory store and reach the attention filter. Due to the brain's limited cognitive resources, once both signals reach the attention filter, the brain will have to decide which channel to prioritize based on the amount of contextual cues each signal contains. In this case, both signals have the same pitch, diction, and semantic meaning; the only difference is the amount of contextual cue their intonation offer. Subjects in this group all shadowed more instantaneously and

doi:10.6342/NTU202000533

confidently compared to other groups. Judging from the experiment result, it is believed that the lively channel which carries more contextual cue lowers the threshold of the attention filter, speeding up the time it takes for the speech to enter the working system. In addition, the monotonous channel is attenuated and in most cases unattended since subjects did not report to face any distractions or interference.

It can be said that the attention filter acts like a road junction, categorizing signals into two separate ways: a highway to the working memory system or a dead end street. The contextual cue which intonation provides grants the lively channel access to a highway directly heading towards the working memory system, shortening the required time and path. On the other hand, the monotonous channel is directed towards a path whose destination is unlikely to reach, explaining how subjects were almost unaffected by the monotonous ear since none of the unmatched answers in this group were merged words or keywords from the wrong ear.

It is suggested that the greater the difference there is between the two channels, the more facile it is for the attention filter to decide which signal to prioritize. This explains why subjects believe it is easier to shadow when the difference between the two ears is easier to discern. Subjects in this group have the highest scores since shadowing the lively ear is in align with the nature of the attention filter (see Figure 20).

doi:10.6342/NTU202000533

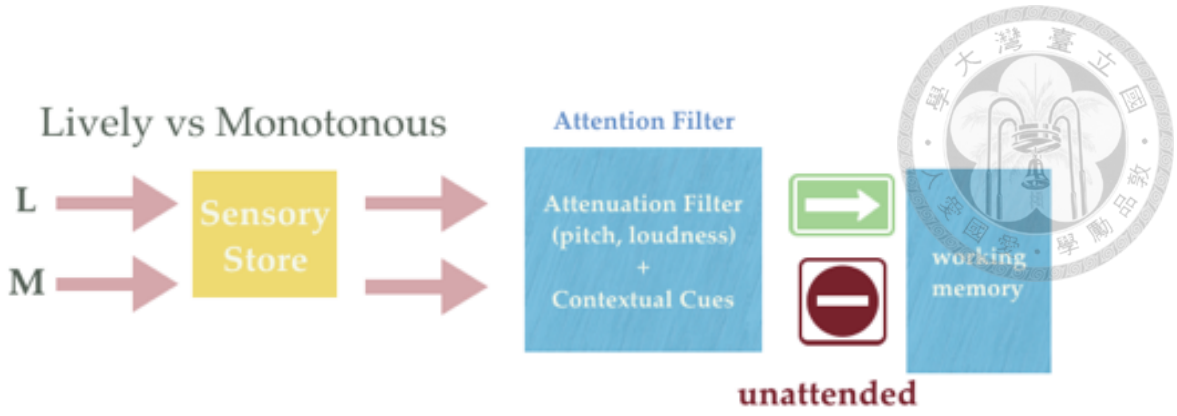


Figure 20. Diagram illustrating how subjects in Group LM process signals.

Lively channel vs. Lively channel (Group LL)

The results of Part 1 show how subjects in Group LL would frequently shadow passages from different ears – an evidence that they struggle with blocking signals from the wrong ear. In addition to this observed interference of the unattended messages, most of the unmatched responses were silent responses or nonsense words that did not pertain to keywords from either passage. This suggests that when subjects are exposed to two lively passages with the same amount of contextual cues, the attention filter will not attenuate neither of the passages, but grant both signals access to the "working memory highway". Due to the fact that one's working memory has a limited capacity to process mental activities, as signals reach the working memory system simultaneously, a "traffic jam" occurs in which the excessive amount of information leads to a cognitive overload (see Figure 21). This cognitive overload results in subjects being overwhelmed by information, thus no longer capable to finish the shadowing task. This offers an explanation of why subjects in this group reported that they were unable to hear the passages clearly and that all the sounds appeared "mixed-up" and merged together.

doi:10.6342/NTU202000533

As established above, lively intonation acts as a contextual cue, facilitating the passing of signals through the attention filter, shortening the time for the signal to reach one's working memory. In most cases, a lively intonation will aid the listener, reducing the cognitive resources required for the listener to pay attention. However, based on the experimental results of Group LL, the researcher believes that when one is exposed to two lively passages, both rich in contextual cues, the excess amount of contextual cues may hinder one's listening attention more than a monotonous passage with scarce contextual cue.

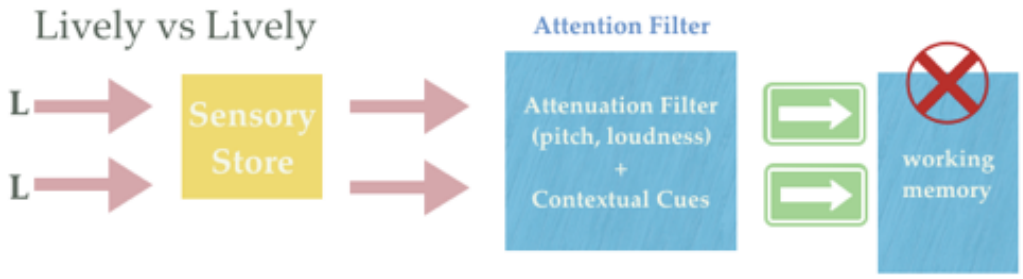


Figure 21. Diagram illustrating how subjects in Group LL process signals.

Monotonous channel vs. Lively channel (Group ML)

Nearly all of the subjects reported shadowing being a difficult task and commented saying that the lively ear appeared louder, faster, and more attractive. Subjects in Group ML share the same mechanism as subjects in Group LM, however, the outcome differs greatly due to the different ear subjects have to shadow. In the case of Group ML, a competition seems to occur between one's attention, the ability to selectively stay concentrated on a stimulus, and one's attention filter. Subjects had to maintain their focus on the attenuated signals that were constantly being overridden by the prioritized lively signals. This is indicated by the fact that subjects would consciously try to switch back to

doi:10.6342/NTU202000533

the monotonous ear after being distracted by the lively ear. In addition, subjects in these group were also more hesitant, more likely to pause in between sentences, and less aware that they switched to the wrong ear.

It can be surmised that one's ability to stay focused and concentrated will largely determine how well one can focus on the less-attended message and resist the signals from the lively ear. It is also reasonable to suspect that if the time of the task exceeds one's attention span, the amount of time one can stay concentrated, the subject's performance will significantly deteriorate. In short, subjects in Group ML were extremely susceptible to the wrong ear and responded with merged keywords and hesitation, showing how the lively channel outweighs the monotonous channel in most cases; and how it requires a paramount amount of volition to withstand the intuitive mechanism of our attention filter (see Figure 20).

Monotonous channel vs. Monotonous channel (Group MM)

Subjects who were asked to shadow one of the two monotonous ears received the second highest score among the four groups, only ranking behind Group LM. Subjects in this group did not report any passages to be especially challenging yet they found shorter sentences easier to shadow. Similar to Group ML, subjects in Group MM are required to exert more voluntary attention in order to process monotonous passages which lack contextual cues (see Figure 22). However, compared to Group ML, subjects in this group did not face the distraction of a lively passage.

It was observed that none of the subjects merged any of the keywords together and that most mistakes were keywords from the wrong ear or words that did not appear in any

doi:10.6342/NTU202000533

of the passages. It is therefore reasonable to suggest that these mistakes were due to factors irrelevant to intonation such as distractions or word frequency. Subject's performance was largely decided by their attention span and their volition to complete the task. In conclusion, it can be said that the process of shadowing the monotonous ear is more associated with action (i.e., other distractions) rather than a perceptual phenomenon.

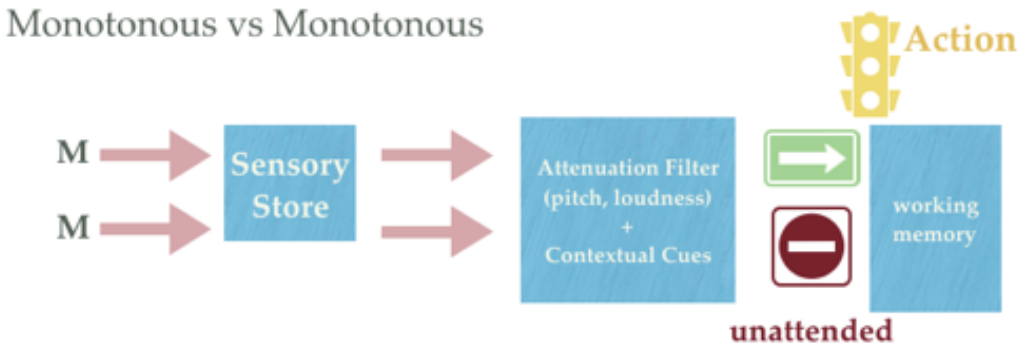


Figure 22. Diagram illustrating how subjects in Group MM process signals.

5.2 How Intonation Affects Listeners' Attention in SI

Why should interpreters focus on intonation?

Previous literatures did not elaborate on why listeners place such emphasis on prosody, yet the findings of this paper reveal how a lively interpreting affects both the audience's preference of the interpreting as well as their ability to pay attention to the interpretation. According to the survey of Part 1 (see Figure 17), 89% of the subjects agreed that the level of liveliness of a speech affects their will to listen. In addition to this subjective evaluation, the results of the dichotic test also indicate that a lively passage is more easily perceived by the audience while monotony creates a negative effect on listening comprehension.

doi:10.6342/NTU202000533

As explained in the Traffic Jam Theory, a lively SI which is rich in contextual cue, is more likely to pass through the attention filter, thus making it more "hearable" and "perceivable" by listeners. This reduces audience's listening burden, allowing them to spend more cognitive effort on the higher level processing of the incoming information, thus increasing their comprehension. Since information can enter the working memory more efficiently, the audience's chances to comprehend and recall heard information is greatly enhanced. As listeners' comprehension of the information improves, so does their satisfaction of the interpreting quality.

On the other hand, if the interpreter's interpretation is monotonous, the audience will have to spend more energy on autonomously attending the speech. A monotonous interpretation short of contextual cues is at constant risk of being attenuated by the attention filter and less prone to enter the working memory system. This means that the listening process is more attention-consuming and listeners are more likely to grow weary of listening. In addition, most headsets used in conference settings do not have noise-cancelling functions, thus any sound that is more likely to pass through the attention filter will distract the audience from focusing on the interpretation. Monotonous interpreting makes it more tiring and difficult for the audience to listen to, as a result, negatively affecting their perception of the interpreting quality.

When should interpreters especially focus on intonation?

It was hypothesized that intonation plays a greater role in emotional texts compared to non-emotional text. However, the results presented in 3.6.2 rejects this hypothesis since significant difference of scores in emotional vs. non-emotional texts was not found in any

doi:10.6342/NTU202000533

of the groups. This shows that a monotonous speech which is not aligned with the speech's emotion does not necessarily impede comprehension. As observed in the previous chapter, all groups scored higher on emotional texts and most subjects believed emotional texts were easier to shadow. A possible explanation for this phenomenon may be that emotional texts carry more semantic meaning (or contextual cues) at a lexical level compared to non-emotional sentences, thus making them easier to listen to.

A closer analysis of the two control and experimental groups, Group ML and MM and Group LM and LL revealed no significant trend of how intonation affects texts with emotions of joy, anger, and sadness, but not fear. However, the scores of the monotonous emotional texts does seem to suggest that the intonation characteristics (Yuan 2006) of different emotions in Mandarin Chinese affects whether subjects perceive the certain passages as being "lively" or "monotonous". Despite the lack of intonation in the monotonous passages, many subjects in Group MM stated that passages with the emotion fear or sadness were much easier to shadow and more memorable because it conveyed emotion and tone. As mentioned in the literature review, the emotions fear and sadness have the narrowest F0 fluctuation in Mandarin Chinese (Li et al., 2011; Gu & Lee, 2007). The small difference between the monotonous and lively versions of fear and sad sentences thus makes it easier for subjects to perceive monotonous fear and sad sentences as more lively and emotional. As shown in Figure. 23, Group ML and MM scored highest in fear and sad sentences; and intonation had no significant effect on fear sentences in both control and experimental groups.

doi:10.6342/NTU202000533

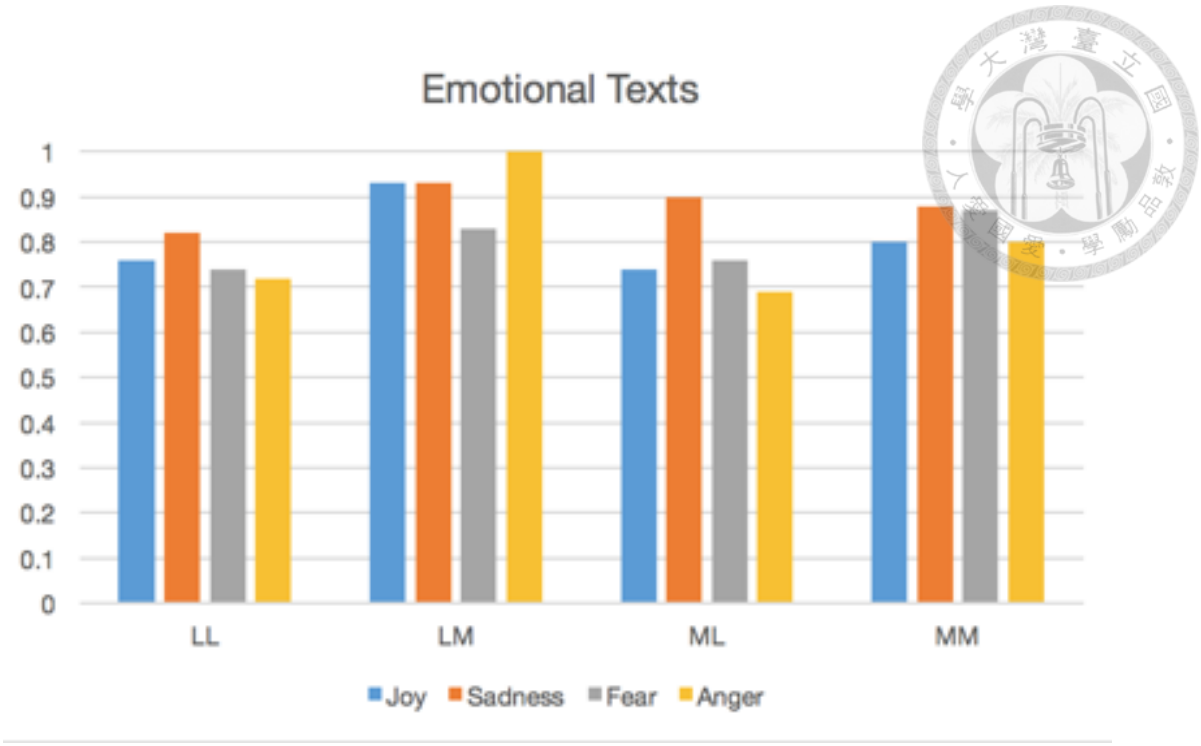


Figure 23. Group scores for the different emotional texts.

When it comes to sentence types, it is observed that the effect of intonation in statement sentences is significant in both groups, indicating how monotony has the most negative effect on statement sentences. This suggests that intonation is very important when interpreting statement passages which declare facts, opinions or ideas. On the other hand, the effect of intonation on question sentences were significant for Group ML and MM yet insignificant for Group LM and Group LL. Almost all the subjects shadowing the lively passage "你們什麼時候開學呀" (When will you *begin school*) responded saying "你們什麼時候開始呀？" (When will you *begin?*) A possible explanation for this common mistake is "word familiarity". Since the experiment took place long after the semester has begun, it is less likely for students to think of the word "school" and more natural for them to say "開始" (begin), a question that is used to ask when people started dating or started

doi:10.6342/NTU202000533

working on a new project. As mentioned in the literature review, one of the most prominent features of Mandarin Chinese is that question sentences are often expressed by the change of the intonation at the sentence-final position. The insignificant difference between Group LM and Group LL seem to reveal that intonation which occurs in the middle of sentences are not as crucial for listeners when it comes to question sentences.

Intonation, as a contextual cue, can greatly facilitate the listening process of the non-emotional texts while monotony might further deteriorate listeners' attention. However, intonation's role is proven to be less significant when it comes to emotional texts. This is because the lexical words of emotional texts can also serve as contextual cues, making it easier to process. In addition, intonation is not as crucial for passages involving fear and sadness. Since these emotions are characterized by a narrower F0 range in Mandarin Chinese, a neutral voice is sufficient for listeners to perceive passages.

5.3 How Interpreters Perceive the Role of Intonation in Mandarin Chinese SI

In the survey study, all interpreters agree intonation is an important non-verbal element which helps interpreters deliver in a more effective and authentic manner. As Subject E6 mentioned, "*words can be deceiving and meaning is often communicated in tone,*" intonation not only makes the interpreting more pleasant acoustically but also provides rich information which improves the audience's comprehension of the interpreting. Various subjects stated that intonation is also a part of the speaker's message and that failing to render the emotion or attitude of the speaker could be regarded as a "mistranslation". In addition, subjects also mentioned that a lively interpreting helps keep

doi:10.6342/NTU202000533

the audience engaged and motivated. These responses are all in line with previous research, proving intonation's critical role as a non-verbal cue, serving to convey messages more efficiently and holistically (Cutler,1997; Levinson, 2003; Steling, 1992).

On the other side, some subjects approached this matter from a client's perspective. They believe clients expect interpreters to be good speakers; a monotonous interpreter appears to be uninterested of the matter and may be deemed as unprofessional. Furthermore, though the audience and clients may not be able to comprehend the source language, a mismatch between the speaker's and the interpreter's intonation may also be regarded as unfaithful to the source language. These results echo the survey results of Moser (1996), demonstrating how intonation may affect audience's perception of the interpreter's professional ability.

It is interesting to point out that experienced interpreters with more than 10 years of experience appear to value intonation more than novice interpreters. In addition, experienced interpreters place more emphasis on interpreter's intonation and regard it as an essential part of the delivery. Novice interpreters tend to focus more on accuracy and less on prosody due to the nature of their training. Since interpreting trainers mostly focus on accuracy and the linguistic features in the discourse, less emphasis is placed on intonation. As a result, novice interpreters who have just completed their training and have just begun practicing in the market may undermine clients' expectations in terms of voice quality or prosody. On the contrary, experienced interpreters are more familiar with client expectations, and tend to be more aware of how audience engagement and non-verbal elements complement accuracy in real-life interpreting.

doi:10.6342/NTU202000533

Nonetheless, some interpreters also mentioned that the use of intonation must not disrupt the clarity of the overall delivery and that interpreters should only convey speaker's emotions to a certain extent. Interpreters must take into consideration the occasion, purpose, and atmosphere of the working setting when deciding whether they should render the speaker's emotions since it might not always be appropriate to do so.

5.4 Implications of the Study

Selection of SI Materials

The experiment in this paper reveals how intonation, as a contextual cue, can significantly lower the attention filter, thereby facilitating the process of listening. In order words, listeners are more likely to pay attention to a lively passage rather than to a monotonous passage. Trained professional interpreters are capable to overcome this perceptual phenomenon because they are trained to allocate cognitive resources to attend the source text. However, when interpreter trainees first begin with their trainings, they are not yet equipped with these necessary skills and practice.

It is therefore strongly suggested that interpreter trainers and trainees take into consideration the prosodic features of the source text before selecting shadowing or SI practice materials at the early stages of the training. Based on the results of the experiment, it is advised to select materials of speakers who have a livelier intonation, as contextual cues in the source text may alleviate students' cognitive loading spent on listening and increase their chances of accurately shadowing the source text. On the contrary, if the source text lacks prosodic cues, students will be required to increase their volition and

doi:10.6342/NTU202000533

attention in order to focus on the source text, possibly leading them to more exhaustion and a degrading performance.

Traffic Jam Theory as an interpreting strategy

Though the subjects of Part 1 were not interpreters, it was inferred from the experiment that monotony will also create a negative impact on interpreters' listening attention. The results of Part 2 confirmed this assumption, with twenty-six out of the twenty-nine interpreters (89.6%) admitting that interpreting for monotonous speakers posed more challenges. Subjects in both groups responded saying that monotony encumbers their ability to grasp the main thrust of the speech and that it was more attention-consuming to follow the speaker's flow. Many experienced interpreters explained that when all words are given equal emphasis, the lack of non-verbal cues force them to solely rely on verbal cues; thus, making it harder to anticipate the speech. Subjects also pointed out that they were more likely to lose interest and be distracted when interpreting monotonous speakers. These findings suggest that though interpreters are trained listeners, intonation still plays a crucial role in helping them identify speaker's intention and emotion, and highlight the emphasis or focus of the speech (Cohen, Douaire, & Elsabbagh, 2001).

Interestingly, strategies interpreters adopt in order to overcome the challenge of monotony seem to echo the explanations provided in the Traffic Jam Theory. From the subjects' answers, one can conclude that interpreters mostly adopt three types of strategies. The first strategy is to allocate more cognitive resources to listening in order to comprehend the information. This once again proves that intonation is a deciding factor in determining whether a signal can pass through the attention filter, and the lack of intonation forces one to allocate more cognitive resources in order to maintain the same level of listening

doi:10.6342/NTU202000533

attention. The need to distribute extra cognitive resources also justifies why interpreters are more vulnerable to fatigue and distraction when interpreting monotonous speakers.

Another strategy that interpreters adopt is to focus on the speaker's body language, stance, and eye contact or rely on the visual aids and information provided in the presentation slides. All these actions can be regarded as interpreters trying to recompense the missing prosodic cues in the speech with other forms of contextual cues in order to facilitate the listening process. Last but not least, two experienced interpreters (E2, E11) with more than 15 years of experience stated that they would adjust their way of delivery when interpreting monotonous speakers. Subject E2 stated that he would speak in a softer manner in order to hear himself better, while Subject E11 said she would speak in a livelier manner in order to keep herself focused. As mentioned in the previous chapter, the attention filter prioritizes or attenuates stimuli based on pitch, loudness and contextual cues. When interpreters are working, they listen to the source language with one ear and monitor their delivery with the other. Therefore, the act of lowering one's volume and speaking in a more neutral voice (narrower F0 range) can be deemed as reducing the noise or distraction from one ear in order to prioritize the information of the other ear. On the other hand, if the interpreter chooses to intensify his or her delivery, then the monotonous source text is more prone to be attenuated. In other words, it can be assumed that Subject E2 "attenuates" his output in order to better listen to the source language while Subject E11 opt to place more focus on her delivery in order to stay engaged to the speech.

In conclusion, monotonous speakers are more difficult to interpret because speeches with little contextual cues are less likely to be perceived. Consequently, interpreters are forced to expend extra attention span in order to comprehend the information. Based on the

doi:10.6342/NTU202000533

ideas presented in the Traffic Jam Theory, it is believed that (1) enhancing one's listening attention (2) compensating the insufficient contextual cues through other means and (3) utilizing one's delivery to facilitate one's monitoring or listening task in order to stay engaged may help interpreters overcome the barriers presented by monotonous speakers.

Voice Training and Intonation Awareness

Since this research has proven intonation to play a crucial role in interpreting practice, it is suggested that interpretation institutions place more emphasis on the role and effect of intonation in the training process. Based on the responses collected from Part 2, most experienced interpreters acknowledge the importance of prosody in interpreting and admit that in actual practice, interpreters with good voice quality and pleasant delivery are regarded as more professional and valuable. Subject E5, an experienced interpreter with 10 to 15 years of experience stated: "*When the audience stops listening to interpreters due to monotony, even if the content is accurate, it probably wouldn't make much of a difference, because the message is not delivered"*. It is interesting to point out that those who disagreed with providing more voice related training simply questioned the feasibility of covering prosodic elements in addition to accuracy and language during classroom lectures.

Considering the significant results of the study, it is suggested that (1) instructors emphasize how intonation can provide semantic, syntactic, and psychological meaning in interpretation as well as enhance listeners' attention and (2) training schools invite voice training experts to share their knowledge with trainees or encourage students to attend voice training lessons to enhance their expression. It is believed that these teachings will allow interpreter trainees to be more aware of the prosodic features of their working

doi:10.6342/NTU202000533

languages, help increase the interpreting quality and efficiency, as well as enhance their client's satisfaction rate (Moser, 1996).

doi:10.6342/NTU202000533

Chapter Six Conclusion

6.1 Answering the research questions

Intonation is an essential non-lexical paralanguage which can facilitate or disrupt spoken communication (Cutler, 1991; Ahrens, 2005). The literatures reviewed in this paper have demonstrated how intonation can influence the audience's assessment of SI quality and their understanding of the interpretation (Holub, 2010; Holub & Rennert, 2011). However, little interpreting studies have investigated intonation's role as a contextual cue in speech communication and its effect on listeners' attention. Part 1 was devised to explore how intonation, as a contextual cue, can influence listener's attention in SI; Part 2, a follow-up experiment, was designed to investigate interpreters' subjective perception of intonation's role in Mandarin Chinese SI.

The present study aims to answer the following two research questions:

1. Can intonation, a wider F0 range, serve as a contextual cue as defined in Treisman's attenuation theory (1960) and help increase listeners' attention in Mandarin Chinese SI. If yes, then how?

2. How will this study benefit future interpreting training and practice?

In terms of the first research question, Part 1 revealed that intonation, as a contextual cue, had a significant effect on listeners' attention— a discovery which echoes Treisman's Attenuation Theory (1960). The author proposed the Traffic Jam Theory to illustrate how lively stimulus and monotonous stimulus interact with the attention filter and thus affect the

doi:10.6342/NTU202000533

listening process. Lively passages facilitate the passing of signals, making it easier for listeners to perceive. However, if listeners are exposed to an excessive amount of signals which can easily pass through the attention filter, a "traffic jam" occurs in which the working memory is overwhelmed by information and no longer capable to function efficiently. On the other hand, monotonous passages make listeners vulnerable to fatigue and distraction due to the extra effort spent to process the signals. This finding revealed that a lively interpretation is not just more appealing to the ears but also enhances listeners' listening attention and improves their comprehension while a monotonous interpretation is more likely to lose the audience's interest.

From the analytical data and observations acquired from Part 1, it can be concluded that though monotony had the most negative effect on listeners' attention when listening to statement sentences, it did not seem to affect one's listening of question sentences or emotional texts. It has also been observed that emotional texts expressing fear and sadness are more likely to be perceived as conveying emotions even when spoken with a monotonous tone due to its intonation characteristics.

Regarding the second research question, Part 1 and Part 2 results both show that listeners are generally more willing to pay attention to lively speeches and that intonation helps keep the audience engaged and motivated. The responses collected from the interpreters who participated in Part 2 further revealed intonation's crucial role in Mandarin Chinese SI, helping interpreters deliver in a more effective, authentic, and holistic manner. Almost all interpreters agreed that intonation is a crucial non-verbal element and that inappropriate intonation may be regarded as unprofessional to clients.

doi:10.6342/NTU202000533

The answers received from Part 2 also support the belief of using the concepts presented in the Traffic Jam Theory as an interpreting strategy. There are three possible strategies that interpreters may apply when interpreting for monotonous speakers. Firstly, they may train their attention span in order to increase their level and duration of focus. Secondly, interpreters may compensate the insufficient contextual cues through other means such as visual aids, body language, or the speech's structure. Thirdly, interpreters may utilize the intonation of their delivery to prioritize either their monitoring or listening task in order to reduce their cognitive work load. Finally, since both experiments prove that even trained listeners such as interpreters still find monotonous speeches more difficult to process, it is advised that students be provided with materials delivered by lively speakers in order to help lessen their listening cognitive loading.

6.2 Contributions of the study

In summary, the present study shows how intonation plays an integral part in selective listening and also explains why a lively SI will increase the audience's comprehension and satisfaction rate while a monotonous SI will create the opposite results. The findings of this study offer pedagogical contributions, suggesting interpreting trainees and trainers should factor in intonation when selecting practice materials for raising intonation awareness and considering using intonation as a strategy while performing Chinese SI.

6.3 Suggestions for future research

This study did not recreate or simulate an actual conference, therefore, it is suggested that future studies conduct a follow-up experiment simulating a real-life conference similar

doi:10.6342/NTU202000533

to Holub's experiment (2010) to further justify how intonation may impact listeners' comprehension and the assessment of Mandarin Chinese simultaneous interpretation. Though this study has proven the crucial role that intonation plays in retaining listener's attention and how it may aid interpreting tasks, the results provided in this study is only a preliminary step towards learning how interpreters perceive the element of intonation in interpreting and their subjective evaluation of how monotony affects their performances. It is therefore suggested that interpreting trainers focus on how interpreters can capitalize on intonation and to what extent they can leverage this tool without deviating from the original speaker in a more intricate and comprehensive manner.

Another direction that would merit future studies are investigations into whether the ideas mentioned in the Traffic Jam Theory can serve as interpreting strategies and significantly help interpreters better monitor their interpretation or focus on the source text. Despite the mentioned limitations, the researcher hopes that this study was able to highlight the importance of intonation in simultaneous interpreting, not just as an accessory, but as an integral factor which may affect the listeners' comprehension of the information as well as perception of the interpreting quality.

doi:10.6342/NTU202000533

References

Ahrens B. (2004). *Prosodie beim Simultandolmetschen*, Frankfurt,Germany: Peter Lang.

Ahrens, B. (2005). Analysing prosody in simultaneous interpreting: difficulties and possible solutions. Retrieved from http://hdl.handle.net/10077/2465

Audacity Team (2018). Audacity(R) Version 2.3.0.

Auer, P., & Di Luzio, A. (1992). *The contextualization of language*. Amsterdam, Philadelphia: John Benjamins.

Bachorowski, J.A. (1999). Vocal Expression and Perception of Emotion. *Current Directions in Psychological Science*, 8(2), 53-57. doi:10.1111/1467-8721.00013

Beckman, M. E. (1986). Stress and non-stress accent. Dordrecht, the Netherlands: Foris Publications. doi:10.1515/9783110874020

Boersma, P. & Weenink, D. (2018). Praat: doing phonetics by computer (Version 6.0.43)

Broadbent, D. E. (1958). *Perception and communication*. New York, NY: Pergamon Press. doi: 10.1037/10037-000

Cherry, E.C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975-979. doi: 10.1121/1.1907229

Chun, M. M., & Wolfe, J. M. (2001). Visual attention. In E. B. Goldstein (Ed.), *Handbook of experimental psychology series* (pp. 272-310). Blackwell Publishing.

Cohen, H., Douaire, J., & Elsabbagh, M. (2001). The role of prosody in discourse

doi:10.6342/NTU202000533

processing. Brain and Cognition, 46(1), 73-82. doi: 10/d4xj4j

Collados Aís, A. (2001). Efectos de la entonacion monotona sobre la recuperacion de la informacion en receptores de interpretacion simultanea. doi: 10/dkvz

Collados Aís, A. (1998). *La Evaluacion de la calidad en interpretacion simultanea : la importancia de la comunicacion no verbal*. Granada, Spain: Comares.

Crystal, D. (1975). *The English tone of voice: Essays in intonation, prosody and paralanguage.* London, England: Edward Arnold.

Cutler A.& Ladd D.R. (1983). Speakers' Conceptions of the Function of Prosody. In Cutler A., Ladd D.R. (Eds.), *Prosody: Models and Measurements*. Berlin, Germany: Springer

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the Comprehension of Spoken Language: A Literature Review. *Language and Speech*, 40(2), 141-201. doi:10.1177/002383099704000203

Driver, J. (2001). A selective review of selective attention research from the past century. *British journal of psychology,* 92, 53-78.

Eady, S. J. (1982). Differences in the F0 Patterns of Speech: Tone Language Versus Stress Language. *Language and Speech*, 25(1), 29-42. doi: 10/dkv2

Epstein, W. (1961). The influence of syntactical structure of learning. *The American journal of psychology*,74, 80-85.

doi:10.6342/NTU202000533

Fantini, A. E. (1995). Introduction - Language, Culture and World View: exploring the nexus. *International Journal of Intercultural Relations*, 19, 143-153.

GÂrding, E. (1987). Speech Act and Tonal Pattern in Standard Chinese: Constancy and Variation. *Phonetica Phonetica*, 44(1), 13-29.

Grice, M., & Baumann, S. (2007). An Introduction to Intonation – Functions and Models. Retrieved from shorturl.at/lnACK

Gu, W., & Lee, T. (2007). Quantitative analysis of F0 contours of emotional speech of Mandarin. Paper presented at the 6th ISCA Workshop on Speech Synthesis, Bonn, Germany, August 22-24, 2007. Retrieved from shorturl.at/xCKLQ

Halliday, M. A. K. (1967). *Intonation and grammar in British English*. The Hague, Paris: Mouton.

Hargrove, P. M., & McGarr, N. S. (1994). Prosody Management of Communication Disorders. San Diego, California: Singular Publishing Group.

Hirschberg, J. and J. Pierrehumbert. (1986). The Intonational Structuring of Discourse. In *Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics*, 24,136-144.

Holub, E. 2010. “Does intonation matter? The impact of monotony on listener comprehension”. *The Interpreters’ Newsletter*, 15,117-126. Retrieved from http://www.openstarts.units.it /dspace/bitstream/10077/4753/1/HolubIN15.pdf.

Holub, E. & Rennert, S. (2011). Fluency and Intonation as Quality Indicators. Paper

doi:10.6342/NTU202000533

presented at the Second International Conference on Interpreting Quality, Casa de la Cultura de Almuñécar, Spain.

Houston, D. M., & Bergeson, T. R. (2013). Hearing versus Listening: Attention to Speech and Its Role in Language Acquisition in Deaf Infants with Cochlear Implants. *Lingua. International review of general linguistics*, *139*,10-25.

Ladd, G. J. D. D. R. (1992). Papers in laboratory phonology II: gesture, segment, prosody. *Phonology*, 11(1), 200-204. doi:10.1017/S0952675700001913

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception & Psychophysics*, 56(2), 183-197. doi:10.3758/BF03213897

Lenglet, C. (2012). The Impact of Simultaneous Interpreting Prosody on Speech Perception: Proposal for an Experiment. Retrieved from shorturl.at/nwyF3

Lenglet, C. (2013). The impact of prosody on the comprehensibility and quality of simultaneous interpreting: A pilot study. Presented in the 7th European Society for Translation Studies Congress. Retrieved from shorturl.at/bq245

Levinson, S.C. (2003). Contextualizing 'contextualization Cues'. Retrieved from http://vm136.lib.berkeley.edu/ANTH/emeritus/gumperz/gumpb.pdf

Liu, D. (1995). Sociocultural transfer and its effect on second language Speakers' communication. *International Journal of Intercultural Relations*, 19(2), 253-265.

doi:10.6342/NTU202000533

doi: 10.1016/0147-1767(94)00008-L

Li, A., Fang, Q., & Dang, J. (2011). Emotional Intonation in a Tone Language: Experimental Evidence from Chinese. Retrieved from shorturl.at/imCEO

Martin, J. G. (1968). Temporal word spacing and the perception of ordinary, anomalous, and scrambled strings. *Journal of Verbal Learning and Verbal Behavior*, 7(1), 154-157. doi: 10.1016/S0022-5371(68)80181-1

Mehrabian, A., & Wiener, M. (1967). Decoding of inconsistent communications. *Journal of Personality and Social Psychology Journal of Personality and Social Psychology*, 6(1), 109-114.

Möbius,B. & Patzold, H. & Hess, H. (1993). Analysis and synthesis of German F0 contours by means of Fujisaki's model. *Speech Communication*, 13(1-2), 53-61.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology*, 11(1), 56-60.

Moser, P. (1996). Expectations of users of conference interpretation. *Interpreting*, 1(2), 145-178. doi: /10.1075/intp.1.2.01mos

Musiek, F. E., & Chermak, G. D. (2015). Psychophysical and behavioral peripheral and central auditory tests. *Handb Clin Neurol*,129, 313-332. doi: 10/dkv3

Murray, I.R. and J.L. Arnott (1993). Toward the simulation of emotion in synthetic

doi:10.6342/NTU202000533

speech: A Review of the Literature on human vocal Emotion. *Journal of Acoustical Society of America*, 93, 1097-1108

Nygaard, L. C., Herold, D. S., & Namy, L. L. (2009). The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning. *COGS Cognitive Science*, 33(1), 127-146.

Nygaard, L. C., & Lunders, E. R. (2002). Resolution of lexical ambiguity by emotional tone of voice. *Memory & Cognition*, 30(4), 583-593. doi:10.3758/BF03194959

Pan, Q. (2012). The Comparative Study on English and Chinese Intonation. *Theory and Practice in Language Studies*, 2(1), 161-164.

Pierrehumbert, J. B. (1980). The phonology and phonetics of English intonation. Retrieved from http://worldcat.org /z-wcorg/ database.

Pierrehumbert, J. H. J. (1986). The intonational structuring of discourse. Retrieved from shorturl.at/mnquP

Posner, M. I. (1982). Cumulative Development of Attentional Theory. *American Psychologist*, 37(2), 168-179. doi:10.1037/0003-066X.37.2.168

Poyatos, F. (1997). *Nonverbal Communication and Translation: New perspectives and challenges in literature, interpretation and the media.* John Benjamins Publishing Company. doi: 10.1075/btl.17

doi:10.6342/NTU202000533

Roda, C., & Thomas, J. (2006). Attention aware systems: Theories, applications, and research agenda. *CHB Computers in Human Behavior*, 22(4), 557-587.

Seeber, K. G. (2001). Intonation and anticipation in simultaneous interpreting: an experiment. Retrieved from shorturl.at/ikuWZ

Shih, C., & Kochanski, G. P. (2000). Chinese tone modeling with Stem-ML. Retrieved from shorturl.at/yIY59

Shlesinger, M. (1994). Intonation in the production and perception of simultaneous interpretation. doi: 10.1075/btl.3.17shl

Speer, S. R., Crowder, R. G., & Thomas, L. M. (1993). Prosodic Structure and Sentence Recognition. *Journal of Memory and Language*, 32(3), 336-358. doi: 10.1006/jmla.1993.1018

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology*, 12(4), 242-248. doi:10.1080/17470216008416732

Tzeng, C. Y., Duan, J., Namy, L. L., & Nygaard, L. C. (2018). Prosody in speech as a source of referential information. *Language, Cognition and Neuroscience*, 33(4), 512-526. doi:10.1080/23273798.2017.1391400

VaissiËre, J. (1983). Language-Independent Prosodic Features. In: Cutler A., Ladd D.R.

doi:10.6342/NTU202000533

(Eds). *Prosody: Models and Measurements* (pp.*53*-66). Heidelberg, Berlin: Springer.

Yang, C. (2011). The acquisition of Mandarin prosody by American learners of Chinese as a foreign language (CFL). Retrieved from shorturl.at/wxPXZ

Yuan, J. (2006). Mechanisms of Question Intonation in Mandarin. Retrieved from https://www.ling.upenn.edu/~jiahong/publications/j01.pdf

Zwischenberger, C. (2010) Quality Criteria in Simultaneous Interpreting: An International vs. a National View, *The Interpreters' Newsletter*,15, 127-142

doi:10.6342/NTU202000533

Appendix 1 Rating Survey for Part 1

姓名：_____

性別：_____

年紀：_____

您的母語是：_____

您是否有任何精神相關的疾病： _____

請針對聽到的音檔選出您最同意的答案
1=非常不同意; 2=不同意; 3=同意; 4=非常同意

	您認為此音檔聽起來語調變化豐富、有抑揚頓挫	最符合這個音檔的情緒是
第一個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第三個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第四個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第五個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第六個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第七個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第八個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第九個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十一個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十二個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十三個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒

doi:10.6342/NTU202000533

第十四個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十五個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十六個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十七個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十八個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第十九個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十一個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十二個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十三個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十四個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十五個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十六個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十七個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十八個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第二十九個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第三十個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第三十一個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒
第三十二個音檔	1 2 3 4	開心 生氣 悲傷 害怕 無情緒

doi:10.6342/NTU202000533

Appendix 2 Passages used for Part 2

The effect of intonation on statement sentences **Does a wider F0 range affect listeners' attention in statement sentences?**		
	Lively Version A	**Monotonous Version B**
	台南今年的經濟成長了一倍	宜蘭今年的經濟成長了一倍
	森林裡住著一隻狐狸	森林裡住著一隻獼猴
	Lively Version D	**Monotonous Version C**
	手機越來越便宜	手機越來越豪華
	東京航空明年將結束營業	京都航空明年將結束營業
The effect of intonation on emotional text **Can intonation serve as a contextual cue in Mandarin Chinese?**		
	Lively Version A	Monotonous Version B
Joy	哇！這個地方讓人覺得好浪漫！	哇！這個地方讓人覺得好快樂！
	Lively Version D	Monotonous Version C
	跟你說一個好消息，我今天跳繩比賽得了第一名！	跟你說一個好消息，我今天作文比賽得了第一名！
	Lively Version A	Monotonous Version B
Anger	他竟然會說出這樣的話，實在太爛了！	他竟然會說出這樣的話，實在太壞了！
	Lively Version D	Monotonous Version C
	你怎麼可以那麼過分！	你怎麼可以那麼任性！
	Lively Version A	Monotonous Version B

doi:10.6342/NTU202000533

Sadness	會發生這件事情，我真的很難過	會發生這種事情，我真的很無奈
	Lively Version D	Monotonous Version C
	我這麼愛你，你怎麼可以利用我	我這麼愛你，你怎麼可以背叛我
	Lively Version A	Monotonous Version B
Fear	拜託 不要碰我	拜託 不要弄我
	Lively Version D	Monotonous Version C
	我不知道 我該怎麼做	我不知道 我該怎麼辦
Intonation's effect on question sentences **Does a wider F0 range affect listeners' attention in question sentences?**		
	Lively Version A	Monotonous Version B
	我們一起去那裡逛逛好嗎？	我們一起去那裡看看好嗎？
	你們什麼時候開學呀？	你們什麼時候吃完呀？
	Lively Version D	Monotonous Version C
	請問公司是在前頭嗎？	請問公司是在十樓嗎？
	你認識他是不是？	你認識他對不對？

doi:10.6342/NTU202000533

Appendix 3 Part 1 Questionnaire

Part 1 Survey

謝謝您願意撥冗參加本實驗，本實驗是與聲音及聽力相關的研究。

*必填

1. 姓名 *

2. **我的所屬班級:** * *單選。*

- Class A
- Class B
- Class C
- Class D
- Class E

3.我的座位號碼： *

4. 性別 *

單選。

- 女
- 男

5. 年齡： *

6.我的母語是中文（國語）嗎？ * *單選。*

- 是
- 不是 （請填下題）

7.我的母語是

8. 我是否患有任何心理健康方面的疾病？ * *單選。*

- 是
- 否

9.我是否曾被診斷有聽力問題？ * *單選。*

- 是
- 否

10.我覺得要專注聆聽特定耳朵的內容很困難 *

單選。

- 同意

doi:10.6342/NTU202000533

- 不同意

11.我覺得同時聽、說很困難 * *單選。*
- 同意
- 不同意

12.下列哪一組句子讓我覺得特別難專心聽、說 *

- 台南今年的經濟成長了一倍/ 宜蘭今年的經濟成長了一倍
- 森林裡住著一隻狐狸/ 森林裡住著一隻獼猴
- 手機越來越便宜/手機越來越豪華
- 東京航空明年將結束營業/京都航空明年將結束營業
- 我們一起去那裡逛逛好嗎？/ 我們一起去那裡看看好嗎？
- 你們什麼時候開學呀？/你們什麼時候吃完呀？
- 請問公司是在前頭嗎？/請問公司是在十樓嗎？
- 你認識他是不是？/你認識他對不對？
- 無任何一組句子讓我覺得特別難聽、說

13. 為什麼？ *

14. 下列哪一組句子讓我覺得特別容易專心聽、說 *
- 台南今年的經濟成長了一倍/ 宜蘭今年的經濟成長了一倍
- 森林裡住著一隻狐狸/ 森林裡住著一隻獼猴
- 手機越來越便宜/手機越來越豪華
- 東京航空明年將結束營業/京都航空明年將結束營業
- 我們一起去那裡逛逛好嗎？/ 我們一起去那裡看看好嗎？
- 你們什麼時候開學呀？/你們什麼時候吃完呀？
- 請問公司是在前頭嗎？/請問公司是在十樓嗎？
- 你認識他是不是？/你認識他對不對？
- 無任何一組句子讓我覺得特別容易聽、說

15. 為什麼？ *

16. 下列哪一組句子讓我覺得特別難專心聽、說 *
- 哇！這個地方讓人覺得好浪漫！/哇！這個地方讓人覺得好快樂！
- 跟你說一個好消息，我今天跳繩比賽得了第一名！/跟你說一個好消息，我今天作文比賽得了第一名！

doi:10.6342/NTU202000533

- 他竟然會說出這樣的話，實在太爛了！/ 他竟然會說出這樣的話，實在太壞了！
- 你怎麼可以那麼過分！/ 你怎麼可以那麼任性！
- 會發生這件事情，我真的很難過 /會發生這種事情，我真的很無奈
- 我這麼愛你，你怎麼可以利用我/我這麼愛你，你怎麼可以背叛我
- 拜託 不要碰我/ 拜託 不要弄我
- 我不知道 我該怎麼做/我不知道 我該怎麼辦
- 無任何一組句子讓我覺得特別難聽、說

17. 為什麼？ *

18. 下列哪一組句子讓我覺得特別容易專心聽、說 *

- 哇！這個地方讓人覺得好浪漫！/哇！這個地方讓人覺得好快樂！
- 跟你說一個好消息，我今天跳繩比賽得了第一名！/跟你說一個好消息，我今天作文比賽得了第一名！
- 他竟然會說出這樣的話，實在太爛了！/ 他竟然會說出這樣的話，實在太壞了！
- 你怎麼可以那麼過分！/ 你怎麼可以那麼任性！
- 會發生這件事情，我真的很難過 /會發生這種事情，我真的很無奈
- 我這麼愛你，你怎麼可以利用我/我這麼愛你，你怎麼可以背叛我
- 拜託 不要碰我/ 拜託 不要弄我
- 我不知道 我該怎麼做/我不知道 我該怎麼辦
- 無任何一組句子讓我覺得特別容易聽、說

19. 為什麼？ *

20. 我覺得聲音的變化及豐富程度會影響我聆聽的意願 * *單選。*

- 是
- 否

doi:10.6342/NTU202000533

Appendix 4 Part 2 Questionnaire

A Survey on Intonation and Interpretation

*必填

1. **1. Name** *

2. **2. How long have I been practicing interpreting?** *

   單選。

   - ◯ 1-5 years
   - ◯ 5-10 years
   - ◯ 10-15 years
   - ◯ More than 15 years

3. **3. I think accuracy (completeness, correct terminology, sense consistency with the original) is important when interpreting simultaneously into Chinese.** *

   單選。

	1	2	3	4	5	
Disagree	◯	◯	◯	◯	◯	Agree

4. **4. I think language (grammar, accent, appropriate style) is important when interpreting simultaneously into Chinese.** *

   單選。

	1	2	3	4	5	
Disagree	◯	◯	◯	◯	◯	Agree

5. **5. I think prosody (lively intonation, pleasant voice, fluency of delivery) is important when interpreting simultaneously into Chinese.** *

   單選。

	1	2	3	4	5

doi:10.6342/NTU202000533

6. **6. If the speaker is conveying an emotional message (sounding joyful, sad, angry, or scared), I will try to deliver my interpretation in a way that reflects the according emotion.** *

單選。

	1	2	3	4	5	
Disagree	◯	◯	◯	◯	◯	Agree

7. **7. Why?** *

8. **8. I believe a good speaker should have a lively intonation.** *

單選。

	1	2	3	4	5	
Disagree	◯	◯	◯	◯	◯	Agree

9. **9. I find it more challenging to interpret a monotonous speaker than a lively speaker.** *

單選。

- ◯ Yes
- ◯ No (please skip question 11)

10. **10. I do/ do not find it more difficult to interpret a monotonous speaker because....** *

11. **11. What are some strategies that I apply when interpreting a monotonous speaker?**

doi:10.6342/NTU202000533

12. **12. I believe a good interpreter should interpret with a lively intonation.** *

單選。

	1	2	3	4	5	
Disagree	◯	◯	◯	◯	◯	Agree

13. **13. I believe people place too much emphasis on intonation when listening to SI and should focus more on accuracy of content.** *

單選。

	1	2	3	4	5	
Disagree	◯	◯	◯	◯	◯	Agree

14. **14. It is proven that a lively intonation will increase the audience's attention and comprehension of the passage. Does this change your perception of intonation's role in SI?** *

單選。

- ◯ Yes
- ◯ No

15. **15. I believe interpreting trainers should focus more on voice training and the importance of prosody.** *

單選。

- ◯ Disagree
- ◯ Agree

16. **16. Why?** *

_____

_____

_____

doi:10.6342/NTU202000533