



**MONASH** University

**The Gift of the Tongue  
- Investigating Spoken Voice Intervention  
for Interpreting Students**

by

**Yingzhi Gu, B.A.; M.A.**

**Dissertation**

Presented to the School of Languages, Literatures, Cultures  
and Linguistics, Faculty of Arts of Monash University

in Fulfilment

of the Requirement

for the Degree of

**Doctor of Philosophy**

**Monash University**

**July 2022**

## **Copyright Notice**

© The author 2022.

I certify that I have made all reasonable efforts to secure copyright permissions for third-party content included in this thesis and have not knowingly added copyright content to my work without the owner's permission.

## **Abstract**

Interpreters are bilingual communication professionals, and voice and speech are fundamental communication tools. It is possible, however, that spoken-voice management and development may not always feature prominently in some training institutions' professional interpreting training programs. This study examines the effects of spoken-voice intervention on students' vocal performance in relation to the delivery of interpreting. Learners from 3 intervention groups and one control group, each group with different levels of exposure to voice interventions, were recorded at two points in their general interpreter training. This study adopted a mixed-method design with both quantitative and qualitative approaches in the description of learners' performance. Three external assessors – a voice coach, an interpreting instructor and a lay assessor – each with a distinct professional profile, evaluated the students' vocal performances. Data was also collected from the learners themselves in order to explore their self-reported perceptions of skill level and if or how they were able to effect positive changes in their interpreting performance. The data analysis shows that the spoken-voice intervention groups recorded significantly higher levels of vocal performance than the non-intervention group. It can be concluded that the spoken-voice intervention approach yielded a highly positive intervention outcome, resulting in significant levels of improvement in vocal performance in relation to the delivery of interpreting delivery for the 3 intervention groups. However, the level of improvement varied due to the different intervention activities, regimens and instruction approaches.

## **Declaration**

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature: .....

Print name: Yingzhi Gu

Date: July 2, 2022

## Acknowledgements

Many people have contributed to this dissertation.

First, I would like to express my sincere gratitude to my supervisors, Dr Jim Hlavac and Dr Hui Huang, whose expertise was invaluable in formulating the research questions and methodology. Their insightful feedback has pushed me to sharpen my thinking and brought my work to a higher level. Without their expert guidance, warm encouragement and enthusiastic support, this dissertation would not have been possible.

I would also like to express my appreciation to my milestone panel members, Dr Uldis Ozolins, Dr Marc Orlando, Dr Robyn Spence-Brown, Dr Rita Wilson and Dr Marko Pavlyshyn, for their valuable insights that have helped strengthen this dissertation and for the inspiration and guidance they have provided.

The dissertation would not have been accomplished without the help of many colleagues and friends. I would particularly like to single out my colleague, Dr Lola Sundin: thank you for your patient support and for taking busy time out of your busy schedule to read my thesis and provide valuable feedback.

I would also like to thank my tutor, Ms Marg Mills, for her valuable guidance throughout my studies in voice training. You provided me with the tools that I needed to choose the right direction and successfully complete my dissertation.

I would also like to thank Ms Hanner Pepper for helping me conduct the voice training seminars with the student participants, and the participants of this study for generously sharing their time and expertise with me.

My heartfelt appreciation also goes to accredited professional editor Mary-Jo O'Rourke AE who provided copyediting and proofreading services according to the national university-endorsed 'Guidelines for editing research theses' (Institute of Professional Editors, 2019).

Last but not least, my greatest appreciation is reserved for my family. I would like to thank my parents for their wise counsel and sympathetic ears. You are always there for me. Finally, I could not have completed this dissertation without the support of my daughter, who provided stimulating discussions as well as happy distractions to rest my mind outside of my research.

# Table of content

<b>COPYRIGHT NOTICE .....</b>	<b>I</b>
<b>ABSTRACT.....</b>	<b>II</b>
<b>DECLARATION.....</b>	<b>III</b>
<b>ACKNOWLEDGEMENTS.....</b>	<b>IV</b>
<b>TABLE OF CONTENT .....</b>	<b>VI</b>
<b>LIST OF FIGURES .....</b>	<b>X</b>
<b>LIST OF TABLES .....</b>	<b>XI</b>
<b>ABBREVIATIONS .....</b>	<b>XII</b>
<b>CHAPTER 1 INTRODUCTION TO RESEARCH PROBLEM .....</b>	<b>1</b>
<b>1.1 MOTIVATION AND BACKGROUND OF SPOKEN-VOICE RESEARCH FOR INTERPRETING PROFESSION.....</b>	<b>1</b>
<b>1.2 RESEARCH FOCUS AND GAPS.....</b>	<b>3</b>
<b>1.3 SIGNIFICANCE OF THE STUDY .....</b>	<b>5</b>
<b>1.4 THESIS OUTLINE.....</b>	<b>6</b>
<b>CHAPTER 2 LITERATURE REVIEW .....</b>	<b>7</b>
<b>2.1 INTERPRETING PEDAGOGY IN RELATION TO SPOKEN-VOICE SKILLS AND PRESENTATION .....</b>	<b>7</b>
<b>2.1.1 Modes of interpreting.....</b>	<b>8</b>
2.1.1.1 Simultaneous interpretation .....	8
2.1.1.2 Consecutive interpretation .....	10
2.1.1.3 Conference interpreting.....	12
2.1.1.4 Community interpreting .....	13
2.1.1.5 Remote interpreting .....	14
<b>2.1.2 Bilingualism, bilingual competence and spoken-voice skills.....</b>	<b>15</b>
<b>2.1.3 Training of interpreters’ spoken-voice presentation skills .....</b>	<b>17</b>
2.1.3.1 Information content and packaging.....	17
2.1.3.2 Component-task training and automaticity .....	19
2.1.3.3 Training of interpreters’ performance and audience awareness .....	21
<b>2.1.4 Interpreting quality standards .....</b>	<b>23</b>
2.1.4.1 Different perspectives on interpreting quality .....	23
2.1.4.2 Interpreting quality assessment benchmarks .....	27
2.1.4.3 Self-evaluation in quality assessment .....	29
2.1.4.4 Language directionality and interpreting quality .....	31
<b>2.2 SPOKEN-VOICE PEDAGOGY APPLIED IN INTERPRETING SETTINGS.....</b>	<b>33</b>
<b>2.2.1 Definitions of voice and speech.....</b>	<b>34</b>
<b>2.2.2 Voice training as an area of interdisciplinary pedagogy and research.....</b>	<b>37</b>
<b>2.2.3 Language directionality and vocal features .....</b>	<b>41</b>
<b>2.2.4 Spoken-voice training approaches relevant to interpreting training.....</b>	<b>42</b>
2.2.4.1 Intensive voice training model.....	43
2.2.4.2 Perceptual learning and optimal training regimens .....	44
2.2.4.3 The visual auditory kinaesthetic learning model.....	45
2.2.4.4 Prosodic feature awareness training.....	47
2.2.4.5 Alexander technique training .....	49
2.2.4.6 Berry’s approach – integration of voice and text .....	50
2.2.4.7 Linklater’s psycho-physical approach.....	52
2.2.4.8 Shewell’s approach – integration of science and art .....	53
2.2.4.9 Implicit and explicit instruction .....	53

<b>2.2.5 Spoken-voice categories and attributes .....</b>	<b>56</b>
2.2.5.1 Spoken-voice production .....	56
2.2.5.1.1 Posture and alignment .....	57
2.2.5.1.2 Body movement .....	58
2.2.5.1.3 Physical tension balance .....	59
2.2.5.1.4 Breathing technique .....	60
2.2.5.1.5 Release of speech muscles .....	60
2.2.5.1.6 Vocal care .....	62
2.2.5.2 Spoken-voice quality .....	63
2.2.5.2.1 Voice projection .....	63
2.2.5.2.2 Vocal resonance and timbre .....	64
2.2.5.2.3 Pace and pauses .....	65
2.2.5.2.4 Articulation, pronunciation, accent, enunciation .....	66
2.2.5.2.5 Modulation of pitch range .....	68
2.2.5.3 Spoken-voice dynamism .....	69
2.2.5.3.1 Intonation and tone .....	70
2.2.5.3.2 Rhythm and fluency .....	72
2.2.5.3.3 Expressiveness .....	73
2.2.5.3.4 Vocal confidence .....	75
2.2.5.3.5 Audience connection .....	76
<b>2.3 METACOGNITION, SELF-REGULATION AND THEIR APPLICATION TO INTERPRETERS' VOCAL PERFORMANCE ..</b>	<b>77</b>
<b>2.3.1 Metacognition .....</b>	<b>77</b>
<b>2.3.2 Self-regulation .....</b>	<b>79</b>
<b>2.3.3 Intersections of metacognition with self-regulated learning and motivation .....</b>	<b>80</b>
<b>2.3.4 Metacognition and self-regulation in interpreting pedagogy and spoken-voice training .....</b>	<b>82</b>
<b>CHAPTER 3 RESEARCH METHODOLOGY .....</b>	<b>90</b>
<b>3.1 MATERIAL PREPARATION STAGE .....</b>	<b>92</b>
<b>3.1.1 Recruiting research participants .....</b>	<b>92</b>
3.1.1.1 Selecting student participants .....	93
3.1.1.2 Contacting student participants .....	95
3.1.1.3 Selecting external assessors .....	96
3.1.1.4 Contacting external assessors .....	97
3.1.1.5 Researcher's positionality and qualification .....	97
<b>3.1.2 Setting up experimental groups and control group .....</b>	<b>99</b>
<b>3.1.3 Preparing Mandarin source speeches for pre- and post-intervention assessment .....</b>	<b>100</b>
<b>3.1.4 Preparing intervention for spoken-voice awareness-raising group .....</b>	<b>102</b>
3.1.4.1 Pilot testing interview questionnaires .....	103
3.1.4.2 Devising interview questions and self-evaluation sheet for main study .....	104
Part 1 – Interview questions .....	104
Part 2 – Self-evaluation sheet .....	105
<b>3.1.5 Preparing spoken-voice training for voice-training groups .....</b>	<b>105</b>
3.1.5.1 Designing spoken-voice training schedule for LTG .....	106
3.1.5.2 Designing spoken-voice training schedule for STG .....	108
3.1.5.3 Planning spoken-voice exercises for interpreting students .....	109
<b>3.1.6 Preparing assessment materials for external assessors .....</b>	<b>114</b>
<b>3.2 DATA COLLECTION METHOD .....</b>	<b>118</b>
<b>3.2.1 Collecting video recordings for pre- and post-intervention assessments .....</b>	<b>119</b>
<b>3.2.2 Collecting interview responses and self-reflection sheets from ARG students .....</b>	<b>119</b>
<b>3.2.3 Collecting reflective journals from voice-training groups .....</b>	<b>120</b>
<b>3.2.4 Collecting rating scores and comments from assessors .....</b>	<b>121</b>
<b>3.3 DATA ANALYSIS METHOD .....</b>	<b>122</b>
<b>3.3.1 Quantitative analysis .....</b>	<b>122</b>
3.3.1.1 Within-group improvement .....	122
3.3.1.2 Cross-group comparison .....	123
3.3.1.3 Pearson correlation coefficient analysis .....	124
<b>3.3.2 Qualitative content analysis of interview responses and reflective journals .....</b>	<b>124</b>
<b>3.3.3 Magnitude coding method for assessors' comments .....</b>	<b>126</b>
<b>CHAPTER 4 PRESENTATION OF RESULTS .....</b>	<b>134</b>



<b>4.1 QUANTITATIVE RESULTS .....</b>	<b>134</b>
<b>4.1.1 Results for comparison of vocal performance before and after intervention.....</b>	<b>134</b>
<b>4.1.2 Results for multiple group comparison .....</b>	<b>136</b>
<b>4.1.3 Results of correlational analysis .....</b>	<b>140</b>
<b>4.1.4 Comparison of vocal categories before and after intervention.....</b>	<b>144</b>
<b>4.2 QUALITATIVE RESULTS.....</b>	<b>147</b>
<b>4.2.1 Spoken-voice attributes that recorded high improvement.....</b>	<b>147</b>
4.2.1.1 Vocal confidence .....	149
4.2.1.2 Audience connection and expressiveness.....	150
4.2.1.3 Speaking clarity .....	151
4.2.1.4 Posture and alignment.....	152
4.2.1.5 Voice projection.....	153
4.2.1.6 Rhythm and fluency .....	154
4.2.1.7 Intonation and tone .....	155
4.2.1.8 Breathing support and modulation of pitch range .....	156
4.2.1.9 Pace and pauses .....	158
<b>4.2.2 Factors that contributed to high performance.....</b>	<b>160</b>
4.2.2.1 Kinaesthetic awareness .....	160
4.2.2.2 Vocal care awareness .....	162
4.2.2.3 Prosodic feature awareness .....	164
4.2.2.4 Audience awareness.....	166
4.2.2.5 Changed perceptions of spoken-voice training for interpreters.....	168
4.2.2.6 Awareness of dealing with tension.....	170
4.2.2.7 Capacity for self-monitoring and self-reflection .....	172
4.2.2.8 Motivation and self-regulated practice.....	176
<b>CHAPTER 5 DISCUSSION OF RESULTS .....</b>	<b>185</b>
<b>5.1 INTERVENTION METHODS AND INTERVENTION EFFECTS.....</b>	<b>186</b>
<b>5.1.1 Intervention activities and improved spoken-voice awareness.....</b>	<b>189</b>
5.1.1.1 Kinaesthetic awareness and spoken-voice production .....	189
5.1.1.2 Prosodic feature awareness and spoken-voice quality .....	191
5.1.1.3 Audience awareness and spoken-voice dynamism.....	192
<b>5.1.2 Spoken-voice training regimens and motivation for regular self-practice .....</b>	<b>193</b>
<b>5.1.3 Metacognitive instruction and self-regulation .....</b>	<b>196</b>
<b>5.2 SPOKEN-VOICE PERFORMANCE AND INTERPRETING PRESENTATION QUALITY .....</b>	<b>199</b>
<b>5.3 HIGHLY IMPROVED SPOKEN-VOICE CATEGORIES AND ATTRIBUTES.....</b>	<b>200</b>
<b>5.4 PERSPECTIVES OF MULTIPLE EXTERNAL ASSESSORS.....</b>	<b>202</b>
<b>CHAPTER 6 CONCLUSION .....</b>	<b>206</b>
<b>6.1 SUMMARY OF RESULTS .....</b>	<b>206</b>
<b>6.2 PEDAGOGICAL AND METHODOLOGICAL IMPLICATIONS.....</b>	<b>209</b>
<b>6.3 LIMITATIONS AND FUTURE RESEARCH DIRECTIONS .....</b>	<b>212</b>
<b>REFERENCES.....</b>	<b>215</b>
<b>APPENDIX A-1 EXPLANATORY STATEMENT AND CONSENT FORM .....</b>	<b>247</b>
<b>APPENDIX A-2 EVALUATION INSTRUCTION.....</b>	<b>252</b>
<b>APPENDIX A-3 GLOSSARY FOR INTERPRETER VOICE SKILLS EVALUATION SHEET1F1F1F1F .....</b>	<b>253</b>
<b>APPENDIX B-1 PRE-INTERVENTION INTERPRETING SPEECH .....</b>	<b>255</b>
<b>APPENDIX B-2 POST-INTERVENTION INTERPRETING SPEECH.....</b>	<b>259</b>
<b>APPENDIX C-1 INITIAL AND FOLLOW-UP INTERVIEW QUESTIONS .....</b>	<b>263</b>
<b>APPENDIX C-2 SELF-EVALUATION SHEET .....</b>	<b>267</b>
<b>APPENDIX D-1 VOICE TRAINING SCHEDULE FOR LTG .....</b>	<b>269</b>

APPENDIX D-2 VOICE TRAINING SCHEDULE FOR STG .....	271
APPENDIX D-3 SPOKEN VOICE EXERCISES FOR INTERPRETING STUDENTS .....	275
APPENDIX D-4 SAMPLE OF VOICE TRAINING HANDOUT .....	292
APPENDIX E-1 INTERPRETER VOICE SKILL EVALUATION SHEET .....	294
APPENDIX E-2 INTERPRETER VOICE SKILL EVALUATION SHEET .....	300
APPENDIX F SAMPLES OF INTERVIEW RESPONSES.....	306
APPENDIX G SAMPLES OF REFLECTIVE JOURNALS .....	315
APPENDIX H RELIABILITY STATISTICS .....	317
APPENDIX I-1 TEST OF WITHIN-SUBJECT EFFECTS FOR FOUR GROUPS.....	318
APPENDIX I-2 TEST OF WITHIN-SUBJECT EFFECTS FOR FOUR GROUPS.....	319
APPENDIX I-3 TEST OF WITHIN-SUBJECT EFFECTS FOR FOUR GROUPS (INTERPRETING INSTRUCTOR ASSESSOR) .....	320
APPENDIX J-1 TEST OF WITHIN-SUBJECT EFFECTS FOR THREE CATEGORIES OF FOUR GROUPS (VOICE TRAINER ASSESSORS).....	321
APPENDIX J-2 TEST OF WITHIN-SUBJECT EFFECTS FOR THREE CATEGORIES OF FOUR GROUPS (LAY ASSESSORS).....	325
APPENDIX J-3 TEST OF WITHIN-SUBJECT EFFECTS FOR THREE CATEGORIES OF FOUR GROUPS (INTERPRETING INSTRUCTOR ASSESSOR) .....	329
APPENDIX K-1 BETWEEN-GROUP TEST RESULTS FOR FOUR GROUPS FROM THREE ASSESSORS .....	333
APPENDIX K-2 BETWEEN-GROUP TEST RESULTS OF THREE CATEGORIES.....	339
APPENDIX L-1 CORRELATIONAL ANALYSIS INTERPRETING PERFORMANCE AND VOCAL PERFORMANCE (PRE AND POST).....	356
APPENDIX L-2 CORRELATIONAL ANALYSIS ATTRIBUTES AND OVERALL INTERPRETING PERFORMANCE .....	357
APPENDIX M SAMPLE OF TRANSCRIBING OF ASSESSOR’S COMMENTS .....	390
APPENDIX N-1 MATRIX OF NUMERICAL RATINGS OF EVALUATIVE CONTENT (STG).....	393
APPENDIX N-2 MATRIX OF NUMERICAL RATINGS OF EVALUATIVE CONTENT (LTG) .....	394
APPENDIX N-3 MATRIX OF NUMERICAL RATINGS OF EVALUATIVE CONTENT (ARG).....	395
APPENDIX N-4 MATRIX OF NUMERICAL RATINGS OF EVALUATIVE CONTENT (COG) .....	396

## **List of Figures**

Figure 2.1 Quality standards for the product and service of interpreting.....	29
Figure 3.1. Procedure of Analysing the Students' Journals (STG & LTG) .....	125
Figure 3.2. Procedure of Analysing Assessors' Comments by QCA.....	127

## List of Tables

Table 2.1 Attributes of spoken language interpreting quality standards from ISO and ASTM. ....	28
Table 2.2 Vocal and non-vocal elements of communication .....	74
Table 3.1 Summary of student participants' background information. ....	94
Table 3.2 Assessors' profiles. ....	97
Table 3.3 List of spoken-voice exercises for interpreting students. ....	109
Table 3.4 Student Speech and Interpreting Self-Evaluation Sheet.....	114
Table 3.5 Data collection timelines. ....	118
Table 3.6 Rating scale descriptions for assessors' comment text.....	- 129 -
Table 3.7 Matrix of numerical ratings of evaluative content (STG). ....	- 131 -
Table 4.1 Descriptive data and within-group comparison (pre- and post-intervention scores from 3 assessors).....	135
Table 4.2 Tests of between-subjects effects (using scores from 3 assessors). ....	137
Table 4.3 Cross-group post-hoc LSC pairwise comparisons of post-test results. ....	138
Table 4.4 Correlation coefficients between spoken-voice scores and interpreting performance scores ..... (pre and post). ....	141 141
Table 4.5 Correlational coefficients between scores for spoken-voice attributes and interpreting performance (pre and post). ....	142
Table 4.6 Descriptive data and within-group comparison for 3 categories (pre- and post-category scores from 3 assessors for STG). ....	144
Table 4.7 Descriptive data and within-group comparison for 3 categories (pre- and post-category scores from voice trainer assessor for LTG). ....	145
Table 4.8 Descriptive data and within-group comparison for 3 categories (pre- and post-category scores from 3 assessors for ARG).....	146
Table 4.9 Summary of spoken-voice attributes that recorded high improvements. ....	148
Table 5.1 Intervention methods and intervention effects. ....	188

## Abbreviations

<b>AIC</b>	International Association of Conference Interpreters
<b>ARG</b>	Spoken-voice awareness-raising group
<b>CALD</b>	Culturally and linguistically diverse
<b>CI</b>	Consecutive interpretation
<b>CoG</b>	Control group
<b>CIUTI</b>	Conférence Internationale d'Instituts Universitaires de Traducteurs et Interprètes
<b>LTG</b>	Long-interval voice-training group
<b>MTS</b>	Master of Translating and Interpreting Studies
<b>MRI</b>	Metacognition and reflective inquiry
<b>MTG</b>	The long-interval voice-training group
<b>NES</b>	Non-English-Speaking
<b>OPI</b>	Over-the-phone interpreting
<b>PBL</b>	Problem-based learning
<b>RI</b>	Remote interpreting
<b>SI</b>	Simultaneous interpretation
<b>SL</b>	Source language
<b>STG</b>	Short daily voice-training group
<b>TL</b>	Target language
<b>VCA</b>	Victorian College of the Arts
<b>VRI</b>	Remote video interpreting

# **Chapter 1 Introduction to Research Problem**

The first chapter of this thesis provides an overview of the study. It begins with examining the research purpose and objectives and this is followed by a presentation of the research background. The research focus and gaps, the key research questions and the study's significance are provided. The chapter concludes with a description of the organisation of the thesis.

## **1.1 Motivation and background of spoken-voice research for interpreting profession**

My experience as a former interpreting student, an interpreting practitioner and an interpreting instructor served as a prime motivation for pursuing this study. When I (the author researcher, hereafter referred to as the researcher) was an interpreting student in China many years ago, interpreting pedagogy was still a new discipline. Interpreting training was focused mainly on acquisition of bilingual language skills, knowledge acquisition, memory retention and linguistic transfer. The verbal performance aspect of interpreting and spoken-voice skills was not given its due importance.

After being a professional interpreter for a few years in China and moving to Australia, the researcher was suddenly enriched with the added attributes of being a Chinese migrant and an interpreter working in the Australian community and in conference settings. However, she encountered many formidable challenges due, in part, to insufficient training related to interpreting delivery and vocal performance. She met many clients who commented that interpreters' voice quality and speaking skills were critical to them as audience members. However, for various reasons students in interpreting training programs may hold different views from the interpreting audience in the real market. A few years ago, Bartłomiejczyk (2014) undertook an empirical study to understand which quality components interpreting students perceived as crucial during self-evaluation. The results demonstrated that, in contrast to interpreting audiences' expectations of interpreting quality, most interpreting students still regarded completeness and faithfulness to the original message as the only important qualities in interpreting. As Bartłomiejczyk (2014) pointed

out, “Issues of presentation (including monotonous intonation, hesitant voice, and long pauses), on the other hand, were hardly ever mentioned” (p. 247).

The researcher does not contest that faithful and complete rendition is a key attribute of interpreting performance. However, she maintains that for professional interpreters who are committed to providing quality interpreting services in both community and conference interpreting settings, the spoken-voice skills are as important as the faithful rendition. This is evident by studies on interpreting quality assessment conducted from the mid-1980s (Bühler, 1986; Meak, 1990; Kurz, 2001) that vocal quality influences the audience’s perception and actual assessment of the message in an actual interpreting assessment. Previous research findings, both from observational studies (Namy, 1978; Gile, 1991; Iglesias Fernández, 2006, 2007) and experimental studies (Collados Aís, 1998, 2002; Pöchhacker & Shlesinger, 2002; Garzone, 2003; Collados Aís et al., 2007; Iglesias Fernández, 2007) also provide strong support for this argument.

To exemplify this, the researcher recounts the following story from a colleague, a senior conference interpreter, who told her something that greatly aroused her interest in this topic. During a conference interpreting break, a Chinese delegate approached him and commented (in Mandarin), “Mister Interpreter, you are doing well, but I think it would be better if you could sound like a radio announcer”. The audience did not seem to understand the strenuous nature of interpreting work and the request he made seemed infeasible. Nevertheless, from the perspective of a monolingual audience – “they listen to the speaker, but it is the interpreter they hear” (Gebhard, 2000, p.143). It is legitimate for the audience to request a quality interpreting service which affords the maximum listening pleasure to be gained through accurate linguistic transfer. This demonstrates how an unpleasant delivery in interpreting negatively impacts the audience’s assessment of the message to the extent that fidelity is assessed as compromised and not fulfilled (Iglesias Fernández, 2007).

Herbert, who is one of the first generation of well-noted professional interpreters, also suggested, “A good interpreter must be a trained public speaker” (1952, p. 59). Shewell, a spoken-voice

teacher and speech and language therapist with long experience in voice work with a wide range of clients, pointed out that the voice is “a two-way psychosomatic phenomenon” (2009, p. 4). This means that interpreters’ voices are shaped by their own physical and psychological states, and this will, in turn, affect their audience’s physical sensations and emotions. The ultimate success of an interpreter lies in whether they can use their voice and speaking skills to command the attention of their audience and produce the maximum effect on them as the original speech does on the audience. Both spoken-voice skills and interpreting skills are crafts. According to Rodenburg (2015), spoken-voice training is something that everyone who aspires to improve the sound of their voice and the way they speak should consider so that they can address problems that may exist with their habitual spoken-voice patterns.

The purpose of this research is to examine methods of spoken-voice intervention that are appropriate for the interpreting profession. This study does not seek to discover a new view of vocal knowledge but to examine it from different perspectives, particularly those of interpreting instructors, interpreting trainees and interpreting audiences. The researcher’s professional interpretation expertise has provided her with a distinct perspective from which to undertake this enquiry.

## **1.2 Research focus and gaps**

Spoken-voice pedagogy has been developing for over 200 years, mainly for professional voice users such as actors. There are many similarities between the work of interpreters and of actors. For both, voice and speech are their most important technical equipment. Both consist of “much more than the delivery of his lines ... take away the element of voice, and very little is left to him” (Turner, 2003, p. 1). However, there are fundamental differences between the roles of the two professions (Roloff, 1973). For example, interpreters do not need to learn specific acting techniques as actors do. Much of their attention is focused on bilingual linguistic transfer and delivery; interpreters do not take on the role of a character in a play and usually do not have significant body



movement; interpreters use their voice and body language to represent the intention and mood of others, but do not become another character. As a result, voice-training methods created for actors or other professional voice users may need to be adapted for usage by interpreters.

Gebhard's (2000) article on voice management published in the International Association of Conference Interpreters (AIIC) bulletin, marks a call for voice trainers to facilitate voice training workshops for professional conference interpreters. To be sure, such training had already been a feature of some training institutions' programs before then. Gebhard's (2000) support for voice training can be seen as a reminder of its importance. However, there is very little information known so far on the results gained from these voice-training workshops due to a shortage of research on the subject.

Because of the absence of a systematic spoken-voice training pedagogy in current interpreting training programs, interpreters may hold different views towards their vocal instrument and spoken-voice training. Some may adopt a "get-by" attitude, feeling that interpreters always work in the second line as a shadow of the speakers and therefore it is a waste of time and energy to have voice and speech training. Others may acknowledge the importance of spoken-voice training but believe vocal awareness itself will make them perform better, and so they do not want to make much effort to practise spoken-voice skills. Still others may have realised the importance of spoken-voice skills to their career through various trials and errors, and intend to improve their vocal skills. However, they might find it difficult to identify and implement voice-training techniques for their actual work and to take consistent and conscious action to improve their spoken-voice skills and habitual vocal behaviours.

Given the different understandings of spoken-voice training for interpreters, this study has conducted experiments with the aim of exploring the effectiveness of spoken-voice intervention approaches for interpreting students and comparing the outcomes of those who

had spoken-voice intervention and those who did not. Given this scope, the researcher has conducted a research project to explore the following preliminary questions:

1. To what extent would interpreting students who receive spoken-voice intervention make progress and how different would the outcome be between conducting and not conducting spoken-voice intervention?
2. What specific spoken-voice intervention procedures are the most effective for interpreting students?
3. Would merely raising awareness of spoken-voice skills (without actual voice training) improve interpreters' presentation of vocal skills?

It is worth noting that the term “spoken-voice training” in this study does not only refer to well-known voice and speech attributes such as modulation, projection, speaking clarity and intonation, but also includes presentational and interactional performance attributes such as body movement, vocal confidence, audience connection (such as eye contact) and posture and alignment, collectively referred to as speech behaviour, all of which are likely to have an effect on the overall spoken-voice quality (Horváth, 2012).

### **1.3 Significance of the study**

This research project seeks to make a valuable contribution to both interpreting literature and interpreting pedagogy. Research-wise, this study uses a mixed research method to evaluate the outcome of the spoken-voice intervention approaches and understand the factors behind the outcome. Pedagogically, this study seeks to bridge the gap between spoken-voice pedagogy and interpreting pedagogy by conducting an empirical experiment. The researcher hopes that this study will assist interpreting instructors in identifying the factors that contribute to students' better vocal performance while undertaking interpreting duties, as well as assisting interpreting students in increasing their vocal awareness, changing their habitual vocal behaviour, applying their spoken-

voice skills in future professional settings and ensuring them of a competitive edge in their training and professional endeavours.

#### **1.4 Thesis outline**

The thesis consists of 6 chapters in total.

**Chapter 1** is an introductory chapter that has established the context for the thesis.

**Chapter 2** provides disciplinary contexts for interpreting pedagogy and spoken-voice pedagogy.

**Chapter 3** discusses the data collection methodology, which includes the research design, the data collection procedure, the tools used to collect data and information from participants, the quantitative and qualitative data analysis procedures and a detailed explanation of the spoken-voice intervention approach used with the intervention groups.

**Chapter 4** presents the findings of the data analysis.

**Chapter 5** discusses the findings.

**Chapter 6** summarises the thesis, discusses the implications of the findings along with the study's contributions and limitations, and makes recommendations for further research.

## **Chapter 2 Literature Review**

This chapter is divided into 3 sections. Section 2.1 offers a review of the literature on interpreting training for spoken-voice delivery, interpreting performance quality and interpreting performance evaluation criteria for spoken-voice delivery. Section 2.2 discusses the fundamental notions of spoken-voice pedagogy and how they are applied to interpreting instruction. Section 2.3 discusses metacognition and self-regulation theories and their application to interpretation and spoken-voice training.

### **2.1 Interpreting pedagogy in relation to spoken-voice skills and presentation**

Section 2.1 provides a literature review on interpreting training for spoken-voice delivery, interpreting performance quality, and interpreting performance evaluation criteria. Interpreting is a process of conveying messages and “understanding” (Nolan, 2008, p. 2) verbally between people who speak different languages or using sign languages for people who have hearing impairments (Roy, 2000). This study looks only at the former – the interpreting activity performed through the instrument of voice and speech – which is often referred to as “oral translation” (Lee & Buzo, 2009, p. 3). The language of a speech to be interpreted into another language is commonly referred to as the source language (SL) and the language of the interpreted message is called the target language (TL).

Spoken language interpreting (hereafter referred to as interpreting) is a live bilingual performance. The interpreting process cannot be described straightforwardly as decoding and encoding between the SL and the TL. The interpreters listen to, comprehend and mentally analyse the SL speech and re-express the SL speech in the TL speech in real time. Although each stage of the interpreting process is vital, the effort that interpreters require to listen to, analyse and retain (via short-term memory) the content of a source speech (Gile, 2009) cannot be perceived directly by the audience. The audience can only see, listen to and interact with the interpreters’ production effort at the final stage of the interpreting process. It is at this stage, through the interpreters’ presentation of spoken-

voice skills, that “live and immediate oral messages” (Lee & Buzo, 2009, p. 3) are conveyed across different languages and the speakers and audiences have “dynamic communicative interaction” between various backgrounds and cultures. The following subsections introduce the relevant spoken-voice skills required in the major modes and settings of interpreting and discuss the current interpreting training literature in relation to spoken-voice delivery.

### **2.1.1 Modes of interpreting**

Interpreting can be categorised according to different modes of delivery, settings and subject matter. The divisions often overlap. Based on delivery mode, interpreting falls into two major categories: simultaneous interpretation (SI) and consecutive interpretation (CI). According to the subject matter and interpreting setting, interpreting can be generally categorised into 3 types: conference interpreting, community interpreting and remote interpreting (RI). As interpreting is highly specialised work, interpreters can choose to specialise in certain modes and interpreting assignments according to their preferences. However, regardless of the adjective preceding the word “interpreter”, practitioners of this profession worldwide perform the same service and should meet the same standards of competence (Mikkelsen, 2009). A detailed explanation of these modes and types of interpreting and how spoken-voice skills are related to each type is presented in the following subsections.

#### **2.1.1.1 Simultaneous interpretation**

Due to an increasing number of multilingual conferences conducted in the SI mode, people tend to equate conference interpretation with SI. Weber (1989, p. 162) stated, “ninety-eight per cent of assignments of conference interpreters surveyed by the International Association of Conference Interpreters (AIIC) is in the simultaneous mode.” However, as Jones (1998, p. 6) pointed out, “all conference interpreters should be able to work in both modes, as one can never rule out being called upon to work in consecutive modes”.

The SI mode has been used on a large scale since it was found to be workable at the Nuremberg trials (Nolan, 2008, p. 3). A feature of this mode is dependence on high-tech booth facilities (Yin, 2011). The biggest advantage of the SI mode is that it is not as time-consuming as CI, which doubles the presentation time but enables a genuine multilingual interpretation to take place with as many as 6 languages (United Nations) to 11 languages (European Union) (Jones, 1998, p. 6). The SI mode allows the speakers and audience to communicate directly across linguistic and cultural boundaries through using technical equipment such as a soundproof booth, an interpreter station, a wireless transmission system and receivers. Interpreters who can do SI are specially trained to listen to one language while speaking (into a microphone) almost at the same time (simultaneously) in another language. For this reason, SI is often regarded as the most advanced level of interpreting and trained SI interpreters are described as being at the pinnacle of the industry (Mikkelsen, 2009). SI interpreters must render the message as immediately and smoothly as they can while listening continuously to the incoming information from the SL speaker (Kirchhoff, 1976/2002).

The disadvantage of the SI mode is that interpreters are technology dependent and usually need to work with a team of two or three interpreters per booth taking turns every 20 to 25 minutes. The time constraint of the SI mode – “interpreters cannot take longer than the original speaker except for a few odd seconds” (Jones, 1998, p. 7) – puts constant pressure on the interpreters and requires them to have “considerable practice and presence of mind” (Nolan, 2008, p. 3).

Simultaneous interpreters generally retain an anonymous voice, with little or no personal contact with the delegates they are working for (Jones, 1998, p. 7). Hence, SI interpreters’ spoken-voice skills play a distinctive role in their SI performance because their voice is the only thing that the audience with headphones can count on in a conference. The interpreters need to convey the original speaker’s intention through their voice and speech, maintaining clear and accurate delivery to avoid any disconnect in the listeners’ mind between the interpreted speech and how they perceive the speaker.

Whispered SI, also known as “chuchotage”, is done when required amenities such as a booth or portable interpreting equipment are not available. The interpreters are expected to sit next to one or a maximum of two people who require interpreting services, lower their voices and whisper the interpretation simultaneously into the person’s ears to facilitate the communication (Jones, 1998; Lee & Buzo, 2009). In community interpreting, whispering is often used when the situation requires it, for example, during court procedures where a barrister is presenting a matter to a judge or jury, or between a psychiatrist and a patient with a mental illness who is often not able to pause for the interpreter and the psychiatrist wants to know the “stream of discourse” of the patient. These situations may pose a challenge for interpreters.

Research has found that excessive whispering tightens and strains vocal cords, and may lead to vocal fatigue and laryngitis (Vox Daily, 2009). When interpreters perform whispered interpretation, they may be more likely to suffer from vocal fry (where the voice is low-pitched and has a characteristic rough or creaking sound). Whispering is particularly taxing on vocal cords because it is “a ragged use of the vocal cords that will only irritate an infection and produce damage” (Rodenburg, 2015, p. 256).

### **2.1.1.2 Consecutive interpretation**

CI is “a process in which adequate information is orally presented and transferred into another linguistic and cultural system” (Hu, 2006, p. 3). With time constraints and the rising number of interpreting languages, organisers of international conferences often have no choice but to opt for the SI mode over CI. However, CI is by no means second-class to SI or a less challenging mode. As a matter of fact, CI was the standard mode long before SI was invented (Nolan, 2008, p. 3). It has several advantages. Firstly, it has a relatively low cost and a moderate requirement for technical support. Secondly, it can provide higher accuracy than the SI mode. Therefore, it is often used when high interpreting accuracy is required, such as in confidential business negotiations, press conferences, round tables, legal settings, medical consultations, scientific committees and many

other important settings with an interpreter on site or via video or telephone conferencing (Kellett, 1995; Yin, 2011).

In the CI mode, an interpreter listens to a segment of a speech in the SL while taking notes and then reproduces the speech in the TL when the original speaker pauses for interpreting. The time pause depends on different speakers – some prefer to talk for just a few sentences and then invite interpretation, while others may talk for a few minutes. In either case, the interpreter is required to be equipped with a good short-term memory and skilful note-taking techniques to deal with any length of speech. Apart from dealing with a heavy retention load, interpreters in CI mode are expected to “react immediately” after the speaker has finished a segment and paused, and their interpretation must be “fast and efficient” (Jones, 1998, p. 7). On top of that, CI interpreters are often invited to stand beside the presenter in front of their audience and deliver the interpretation with or without a microphone. Hence, interpreters are constantly working under the stress of being judged, compared, monitored and questioned by their audience, who may understand little or part of the SL to be interpreted. As Arumí rightly (2012) stated, “Consecutive interpreting entails a large number of almost concurrent cognitive, psychomotor and affective processes, all of which pose major challenges for the interpreter who has to deal with them simultaneously” (p. 812).

As such, CI interpreters, as reproducers of the original speech in another language, must be well presented and bring their audience both accuracy and enjoyment in listening to a speech by using their voice and speech presentation skills. Apart from conducting linguistic transfer, they are also required to interpret using appropriate gestures, eye contact with the audience and pleasant voice projection and speech techniques, whereas interpreters doing SI do not have to bear such an “extra-interpretation” burden as they usually sit in a little booth at the back of a conference room. As Lim (2008) pointed out, “in bilingual situations, consecutive interpretation is still widely used. A good consecutive interpretation is a great public relations tool since participants can see the marvels of interpretation being re-enacted for their very eyes” (p. 126).



With such a highly demanding nature of work, Weber (1989) proposed that the training of public speaking skills for CI interpreters is essential:

Good public speaking skills are the interpreter's safety net when he gets in difficulty. He can then give his full attention to solving the problem at hand and even possibly hide the existence of any such problem (p. 164).

Furthermore, CI is often utilised as an essential means of testing candidates' comprehensive aptitudes, including vocal presentation and general knowledge, by interpreting schools and potential employers (Kellett, 1995; Weber, 1989). Hence, this research project employs the CI mode in order to understand to what extent spoken-voice intervention could help students improve their spoken-voice performance, which will influence other aspects of their interpreting performance.

### **2.1.1.3 Conference interpreting**

Conference interpreting is "a professional communication service rendered in either the simultaneous or the consecutive mode of interpreting in a conference(-like) situation" (Pöchhacker, 2011, p. 50). The AIIC defines conference interpreting as the practice of conveying the meaning of a speaker's message verbally and in another language to listeners who would not otherwise understand. Often conference interpreting is used at high-level meetings where multiple languages are spoken. Conference interpreters help communicate the speakers' messages from one language into another or multiple languages.

For conference interpreters, voice quality is essential. A practical guide for professional conference interpreters published by the AIIC in 2016 reads:

Remember that the delegates' life is not always easy. Some have to wear headphones most of the time ... it is tiring enough to have to follow several days of discussions, and when this has to be done via interpretation through headphones, it becomes very demanding. Be helpful, by being clear and making your interpretation as easy and pleasant to follow as possible. Do not speak in sharp bursts followed by long pauses, nor in a deadpan monotone, nor in a sing-song (AIIC, 2016).

Although interpreters are not required to possess a voice like a professional TV or radio announcer to please their audience, they are still required to deliver messages verbally in which meaning is conveyed not only through words but also through paralanguage – voice quality, intonation and the speed, volume, rhythm, and pitch of speech – all of which are fundamental components of oral communication (Gebhard, 2000).

#### **2.1.1.4 Community interpreting**

Community interpreting, also known as public sector interpreting or liaison interpreting (Gentile, 1997), refers to an interpreting mode that “enables people who are not fluent speakers of the official language(s) of the country to communicate with the providers of public services to facilitate full and equal access to legal, health, education, government, and social services” (Roberts, 1994, p. 127). Community interpreting often takes place in medical, legal, government agency, immigration and school settings. Telephone interpreting, video interpreting and telehealth interpreting fill gaps where physical face-to-face interpreting is not an option.

Unlike the clients in the conference interpreting setting, who are primarily experts, academics, enterprise executives and government officials, a significant number of community interpreting clients are allophone speakers with a very wide variety of educational backgrounds and perspectives that they bring to interpreter-mediated interactions. In the context of this study conducted mainly in Australia, allophone speakers here relate to those with limited English proficiency. One of the ‘clients’ that interpreters work with are non-English-speaking background (NESB) migrants and refugees who come from culturally and linguistically diverse (CALD) backgrounds and who have come to settle in Australia. They often find it difficult to access health, legal and welfare services, struggling to comprehend the jargon and terminology used by professionals (Appel, 2019). As language is closely related to a person’s identity and impacts on people’s power (Fidan, 2017), many NES clients are at a distinct disadvantage in health and legal literacy and in social cohesion, a consequence of their language disparities and socioeconomic disadvantage.

As such, in Australia community interpreters are greatly needed for NES clients who cannot speak English confidently to help them connect with services. Due to the lower level of formality of the community interpreting setting, in the past interpreters who performed community interpreting were once considered amateurs with limited formal interpreting education (Mikkelson, 2009). With growing awareness of its importance, particularly in multilingual and multicultural societies such as Australia, the United Kingdom and the United States, community interpreting training is being offered up to postgraduate level.

During interpreting interaction, interpreters are required to use the first person instead of the third person. As such, their work involves “enacting” or “voiceover” for both parties in another language; for example, an interpreter may alternatively play the roles of a doctor and a patient, or a police officer and a suspect. Hence, their work entails more than merely transmitting information between the clients, but needs to reflect the clients’ intentions, aspirations, position and values through “vocal role play”. Appropriate spoken-voice modulation (Iglesias Fernández, 2007) of the discourse is required to reflect the clients’ intentions and emotions, and to help establish rapport between the clients and the services. In working with CALD clients, interpreters help professionals to understand more about potential cross-cultural barriers and to develop the skills necessary to prevent the adverse effects of poor communication. At the same time, interpreters empower the clients, improve their wellbeing and enable communication in general.

#### **2.1.1.5 Remote interpreting**

RI is distance communication in real time and an alternative method to on-site interpreting. Moser-Mercer (2011) defined RI as a bi- or multi-lingual telephone or videoconference call where interpreters are physically remote from the meeting room and thus do not have face-to-face interaction with the clients. RI started with an over-the-phone interpreting (OPI) service. Countries such as Australia, the United States and the United Kingdom started to use telephone interpreting services from the late 1970s (Kelly, 2008). Australia was the first country to introduce telephone

interpretation, in 1973, as a fee-free service to respond to its growing immigrant communities (Mikkelson, 2003). Since then, OPI has seen fast development owing to the rapid evolution of telecommunication technologies and globalisation.

Currently, OPI and remote video interpreting (VRI) are used worldwide. Through telephonic links and video links, interpreting services are provided for clients who are not physically in the same location. RI covers settings such as medical and social services, business, court cases and various government agencies.

Interpreters who perform RI face several challenges coping with the increasing complexity of RI scenarios compared to on-site interpreters. Moser-Mercer concluded these challenges to be “cognitive, psychological and physiological factors, such as virtual presence, multisensory integration, multi-tasking, emotions, and psychological stress and fatigue” (2011, p. 132). Braun (2015) also reported that VRI is a stressful mode of interpreting and interpreters’ VRI performance was seemingly inferior to on-site interpreting. RI also requires a high level of spoken-voice quality. The absence of visual cues poses challenges for OPI. The clients rely on the interpreters’ spoken-voice skills to convey messages. As Kelly (2008) pointed out, “With telephone interpreting, body language is not observable by the interpreter. A great deal of non-verbal information is reflected in the tone of voice, breath patterns, inflection, vocal volume, and other auditory cues” (p. 3).

This study examines students’ voice quality while performing CI. Nonetheless, as the above shows, successfully interpreting in any of the communication modes discussed requires a firm command of spoken-voice skills and management.

### **2.1.2 Bilingualism, bilingual competence and spoken-voice skills**

The term “bilingualism” is defined as using two different languages with a range of linguistic abilities (speaking, reading and writing) in different contexts (Simmonds et al., 2011). There are different levels of bilingualism. Native-like proficiency in both languages, referred to as “balanced bilingualism” or “equilingualism” (Beardsmore, 1986, p. 9), is rare. Most bilingual and multilingual

people have a dominant language (Summer, 2009). This is why some international organisations such as the AIIC maintain that interpreters should interpret into their mother tongue in pursuit of higher quality work. Languages from the same language family, for example Spanish and French, have similar characteristics and features, whereas languages from different language families, such as English and Chinese, will be dissimilar (Crystal, 2005). TLs that are distinctly different from the first language (L1) require more effort to learn. Given the inherent hurdles that bilingualism presents to them, some interpreters who have acquired a second language (L2) later in life, in particular for those languages from different families, face challenges in their career.

Professional interpreters are required to have higher proficiency in bilingual language comprehension and production than ordinary bilinguals. They must analyse an SL message for its meaning and search for equivalent vocabulary in the languages they use along with semantic equivalence above word level in the TL to ensure the success of the bilingual communicative interaction (Napier et al., 2005). Nonetheless, however necessary bilingual linguistic competence is to interpreters' work, this attribute alone does not lead to a person becoming a competent interpreter. Among many skills and capabilities, such as cross-language linguistic transfer skills (Liddicoat, 1993) and solid professional ethics, spoken-voice skills are vital for interpreters, involving the ability to consciously engage their audience and command its attention by using and managing their voice and speech skills under a high level of performance pressure.

Spoken-voice skill is a verbal expression of language by articulating sounds. The vocal sound results from the physical process of air passing from the lungs through the larynx (containing the vocal folds or vocal cords that open and close to produce sound). The air then passes into the pharynx and out through the mouth and nose. This process is called phonation.

Interpreters' bilingual competence and spoken-voice skills are distinct from each other, although they are closely connected. Interpreters' spoken-voice skills are inevitably influenced by their bilingual proficiency, especially in their L2, and vice versa. However, the terms "bilingual

competence” and “spoken-voice skills” are not interchangeable. The admission tests that many interpreting schools administer emphasise candidates’ bilingual or trilingual language proficiency, with very little focus on the candidates’ vocal performance, including their spoken-voice management ability (Lim, 2008). Even with the best requisite skills and aptitude to become professional interpreters, interpreting students may have issues with their voice and speech without proper voice and speech training.

As such, some interpreters may unwittingly develop some poor speaking habits even though they have successfully built bilingual skills through simultaneous or consecutive acquisition of an L2. Good spoken-voice skills add value to interpreters’ bilingual skills. This study analyses the major sources of vocal problems for interpreting students and looks for the best solution.

### **2.1.3 Training of interpreters’ spoken-voice presentation skills**

While there is abundant literature on notetaking, retention and linguistic transfer skills, literature on interpreters’ spoken-voice delivery is scarce. This subsection examines a few valuable pieces of literature regarding the training of interpreters’ oral presentation skills in the delivery phase of interpreting, one of the most critical phases in interpreting production.

#### **2.1.3.1 Information content and packaging**

The way that a speaker achieves their communication aims at both general and specific levels is to utilise verbal signals, which consist of “informational content and packaging which interact and produce effects” (Gile, 2009, p. 38). The term “packaging” refers to linguistic and paralinguistic (Gile, 2009, p. 29) choices made by the speaker. “Content” means the message that the speaker would like to deliver. From the speaker’s point of view, communication is successful if both the content and the packaging are delivered to their audience. This also applies to interpreters. For interpreters, spoken-voice quality is part of the packaging in aiming to achieve the same effect as the speaker.

The speakers select and adapt both the content and packaging of their speech to their intended audiences depending particularly on how strongly they wish their message to be received, based on their knowledge of the intended audiences (Gile, 2009). In a conference that many specialists attend, the speaker may wish to speak with authority and confidence to show their expertise. The speaker may deliberately use technical terminology to express their ideas to their colleagues. The speaker's tone and pacing also play essential roles in getting their message across – that they understand and believe what they are saying and know the information is of value to the audience. In this context, the speaker may use a serious tone, a slow rate of speed and a relatively low pitch, deliberately pausing in their sentences to add emphasis.

In other situations, the intentions may change, and so does the form. For the same topic but where the audience is from the general public, the speaker may use a completely different strategy by choosing less specialised terminology to express the same ideas (content) or using exaggerated gestures, humour or even slang in their discourse so as to present an image of a humble person and to narrow the gap between a specialist and the public. Their rate of speech may be a bit faster, their tone cheerful.

For interpreters, packaging is a concept that needs to be used extensively in interpreting training along with content. Through training, interpreting students can distinguish the packaging from the content and recognise which packaging, including linguistic register, delivery style, tone and pace, suits the content. Then the interpreting students can use packaging with the audience that has the same impact as the speaker's packaging while retaining the content faithfully. This has been confirmed by Gile (2009) as follows:

While an accurate rendition of information is essential, it is also important to make students aware of the weight of the packaging ... the packaging in speeches is composed of words and linguistic structures of the speech and features of the voice and delivery, plus non-verbal signals and information, including body language ... Good vocalization and pleasant delivery and style does more toward convincing a listener than the quality of the idea that

is formulated or the information that is delivered. Conversely, good content is weakened by poor voice or poor delivery of a speech (p. 38).

### **2.1.3.2 Component-task training and automaticity**

An interpreting task can be divided into several components and sub-components: a listening and analysis component, a short-term memory component and a speech-production component (Gile, 2009). In addition, the speech production may include the following sub-components: the content of the TL speech, including errors and omissions, and the packaging, for example the linguistic output and delivery. These components and sub-components can be processed simultaneously, with each requiring sufficient cognitive processing capacity to provide optimum performance of the whole task.

In interpreting training tutorials, often interpreting trainees are required to perform or practise a whole interpreting task. This, however, according to Schneider (1985), will often lead to problems such as resource overload and consequential frustration and fatigue. When an interpreter is under stress or suffers from a cognitive overload, interpreting quality deteriorates. Currently, there are some studies in the field of cognitive psychology that have provided data to indicate that component training is more effective than whole-task training, showing that component training will not only transfer to performance on the whole task but may even be more effective than whole-task training with inexperienced students (De Groot, 2000).

Skill acquisition theories suggest that effective and efficient learning requires the “automaticity” of the component parts (Muller-Townsend, 2017). According to Harpe (n.d.), automaticity is classified into two broad categories based on how an idea or behaviour is initiated, namely, unconscious automaticity and conscious automaticity. Unconscious automaticity is when some automatic processes are initiated completely spontaneously, frequently in response to environmental stimuli, while conscious automaticity requires a conscious act of will to initiate it. In 1899, William and Harter proposed that lower level processes need to be automatized before the acquisition of higher level components can be attempted. They claimed that non-automatic skills make greater attentional



demands and inhibit the ability to attend to higher order tasks. Automaticity of base-level skills frees up attentional capacity to deal with the attentional demands of more difficult tasks (William & Harter, 1899). Thus, automaticity is considered an essential component of skill acquisition.

Automaticity is the ability to perform a task without occupying the mind with the low-level details required, allowing the task performance to become an automatic response pattern or habit through learning, repetition and practice (Muller-Townsend, 2017). For example, some tasks are carried out by “muscle memory” such as walking, speaking, bicycle-riding and driving a car. After an automatised activity is sufficiently practised, it is possible to focus the mind on other activities or thoughts while undertaking the activity, such as driving while chatting with passengers. Reading comprehension skills can be automatised through practice to develop oral or silent reading fluency (Paige et al., 2014). Therefore, automaticity may be considered essential in learning to facilitate higher order cognitive thinking and task execution.

Through component and sub-component training and exercises, students will be likely to maximise automatic processing of the skills and then have more processing capacity available for the components that cannot be automated and need more processing capacity. In interpreting training, Gile (2009, p. 158) also claimed that some “non-automatic mental operation” requires attentional resources or processing capacity from the limited available supply. Automatic mental operations, however, do not require such attention and can be processed very rapidly. For example, in the component of speech production, speech output is generally a non-automatic operation because interpreters must follow the speech chosen by another speaker. However, interpreting students can become automatic in vocal skills and delivery fluency through proper voice training and repeated practice to reduce the multi-tasking challenge of the speech-production component and the whole interpreting task. Gile (2009) also mentioned that the distinction between automatic and non-automatic operations can be difficult to make because non-automatic operations may become automatic after enough repetition and practice.

### **2.1.3.3 Training of interpreters' performance and audience awareness**

A few decades ago, interpreters were educated to play an invisible role or act as a “translating machine” in their job (Archie, 1949, p. 7). Today, although the myth of interpreters' invisibility has been largely dispersed, voice training may not always be as widespread a feature of training, as evidenced by Lim (2008, p. 3), who asserts that “verbal performance is still one aspect of interpreting training that is often neglected by schools of interpretation and their instructors”, despite the fact that some interpreting scholars and educators have started to recognise the importance of the performative aspects of interpreting and have conducted various experiments and training activities in the hope of improving interpreting performance quality. Following are a few examples.

Naimushin (2014) pointed out that an interpreter who demonstrates good interpreting technique, but little expressive performance can only be regarded as an “experienced non-expert”. As Naimushin suggested (2014):

If we are to prepare real professionals, it is essential to cover everything from the interpreter's posture to controlling anxiety and developing audience interaction, public speaking and performance skills, thus giving fledgling interpreters insight that might have taken years of trial and error otherwise (p. 283).

Cho and Roger (2010) pointed out the important commonalities that exist between CI and acting in the theatre – both are the art of conveying a message using effective techniques of communication and both are performing for an audience. Hence, they designed some workshop activities including verbal and non-verbal types that drew on techniques commonly associated with acting training for the theatre. A warm-up phase at the beginning of each workshop included breathing and voice exercises to help students to relax and prepare mentally and physically for the complex activities that followed. The students were then engaged in various acts of communication. The results showed that most students improved their stress-coping ability and impromptu techniques, and could apply what they learnt to their interpreting practice.

Makarová (1994) was a strong proponent of training students' impromptu performance skills. She proposed to "teach improvisation in interpreting" as a form of extralinguistic preparation, claiming that no matter how much interpreters specialise, there will always be gaps in their knowledge because in many real work situations interpreters do not receive written materials beforehand, or only shortly before an assignment. This causes a lot of anxiety. Therefore, she believed that improvisation must be trained as a component of interpreting courses to prepare students for unexpected situations and as one of the possible methods of crisis management. However, there is no research showing how effective this training method is.

Lim (2008) believed that public speakers and interpreters overlap in many aspects of verbal delivery and therefore public speaking skills should be applied to CI and interpreters trained to become convincing speakers in CI and render interpretation with the same effect as the speaker. Lim indicated that impromptu speaking training can be effective for training interpreting students to help them overcome their stage fright and be mindful of their vocal delivery and audience connection (Lim, 2008).

Kellett (1995) also argued that more attention in CI training should be paid to the quality of vocal performance. Following is her observation from when she worked as an examiner of CI from Italian into English:

I was struck by the anxiety students conveyed through the voice or various gestures and mannerisms. Experience has taught me how important the teaching of delivery and verbal presentation techniques is, especially in the second half of a consecutive course after much time and effort has been dedicated to the techniques of listening, memorizing, note-taking and perfection of language skills. But these factors alone are not sufficient, the interpreter must refine his/her skills in public speaking (p. 43).

The preceding examples demonstrate what Kellett (1995) and Naimushin (2014) claimed: that interpreting students must consider their appearance and performance, which include features such as posture, anxiety management, audience interaction, public speaking and vocal performance skills,

if they are to become accomplished professionals and that interpreting training must develop these attributes.

Nevertheless, there is little empirical research on spoken-voice training in interpreting, nor is there sufficient information regarding the results of spoken-voice training. For one thing, interpreting is a highly specialised profession. For another, voice trainers experienced in training other professional voice users might not be aware of the specific challenges that interpreters face or provide specific exercises and advice on applying this knowledge and expertise to interpreters' fieldwork. This is why some interpreters who eventually seek professional assistance from voice coaches or speech therapists find it challenging to apply the taught voice techniques to their work.

In conclusion, these are beneficial training approaches for improving an interpreter's presentation skills, which sheds light on this study.

#### **2.1.4 Interpreting quality standards**

The interpreting industry and researchers have long debated the quality criteria for interpreting services. This section explores the various perspectives on interpreting quality standards and how these perspectives converge. Additionally, this section discusses the quality of interpreting in relation to directionality and the necessity of self-evaluation in interpreting practice.

##### **2.1.4.1 Different perspectives on interpreting quality**

Assessment of interpreting quality takes various forms due to the diversity of the quality standards of various stakeholders, including but not limited to assessment for interpreting accreditation or certification; aptitude tests before candidates are admitted to an interpreting training program; mid-term and end-of-program assessments by interpreting instructors to review students' progress; and self-evaluation by interpreting students and professional interpreters.

Traditionally, there are typically two types of quality assessment used in interpreting training. The first type is error analysis (Barik, 1975; Bartłomiejczyk, 2014; Kopczyński, 1980), which is a

complex scoring system employing various types of errors as indicators of deteriorating quality in interpreting. Error analysis is criticised for employing the standard of translation in measuring interpreting performance. The second type of quality assessment is propositional analysis (Dillinger, 1994; Kintsch, 1974; Tommola & Helevä, 1998), in which the “propositional analysis score” is calculated as the percentage of source text content successfully rendered in the target text, neglecting syntax and grammatical correctness. Both error analysis and propositional analysis focus on the semantic aspects of interpreting including correctness and completeness, overlooking communicative aspects which are crucial for interpreting quality as perceived by the interpreting audience.

In recent years, divergent opinions have been presented by researchers, interpreting trainers and professionals regarding the definition of interpreting quality and this has enabled the establishment and implementation of a benchmark for the interpreting industry. The fundamental diversity in interpreting modes, subject matter and interpreting settings is the first reason to explain this divergence (Kurz, 2008). Grbić (2008) advocated that different benchmarks must be established to fit different purposes and applied to specific contexts. For example, the delegates at a medical conference would have different expectations from those of the interpreting users in doctor–patient interview settings (Pöchhacker, 1994). In community interpreting settings, interpreters encounter professional clients and CALD clients from a variety of institutions and settings which represent diverging interests, including lawyers and clients, doctors and patients, police officers and suspects. Thus, interpreters need to fulfil quality requirements according to the individual situations.

Experimental studies (Edwards et al., 2005; Strong & Rudser, 1992) have demonstrated that CALD interpreting in the context of community interpreting settings (including sign language interpreters) frequently evaluates the quality of interpreting practitioners based on interpreters’ personality traits, first impressions and attitudes, instead of their interpreting skills.

On the other hand, conference interpreting has more stringent criteria for interpreting quality, as summarised by Déjean le Féal (1990):

What our listeners receive through their earphones should produce the same effect on them as the original speech does on the speaker's audience. It should have the same cognitive content and be presented with equal clarity and precision in the same type of language (p. 155).

Additionally, diverse stakeholders in the interpreting sector have disparate opinions and perceptions of quality. As Garzone (2002) emphasised:

The basic problem is that quality is the sum of several different, heterogeneous aspects, some of which involve different subjects – interpreters, interpreting users/clients, speakers – each with a different view and perception of quality ... There are variations even within the same group: among users, quality expectations tend to vary, but also individual tastes and criteria, often related to socio-linguistic factors (p. 107).

Among the different perspectives, the subjective evaluation of quality by the interpreting end users is the most prominent point of argument. Some researchers have argued that interpreting users who are outsiders to interpreting professions are not reliable judges of fidelity. They either overestimate the importance of smooth delivery of interpretation and may not even realise much of the message is distorted or missing when they rate an interpretation as high quality, or tend to blame the interpreter for lack of clarity when it is in fact the speakers who should be held accountable for the ambiguity (Gile, 1990; Shlesinger, 1997). Grbić (2008) also shared the argument that end users' opinions are often guided by personal judgement rather than professional norms.

Other researchers have contended that the quality of interpretation should always be defined by the end users because it is the chain of communication in the interpreting process that is critically important (Seleskovitch, 1968). Heskett et al. (1990) stated that "measurements of service quality that do not include customer expectations miss the point" (p. 6). Gile (1989) also shared the viewpoint that because the needs and expectations of the users of interpreting services are not

necessarily the same as the standards that interpreters set for themselves, a multi-parameter model for description and analysis is required. This means that the TL receivers or listeners are an essential element in the quality evaluation process (Pöchhacker, 1994).

In fact, the interpreting industry cannot avoid the evaluations from “lay interpreting users” merely because they may not be able to judge the interpreting fidelity. Rather than adopting profession-centred benchmarks, interpreting service providers and interpreting experts should seek to understand and meet end users’ expectations, as well as adhering to user-oriented professional standards (Kurz, 2009). Thus, research into audience preferences regarding the definition and evaluation of interpreting quality is of crucial importance for a profession whose aim is to establish effective communication between speaker and audience. As Marrone (1993, p. 35) pointed out, “checking our assumptions against our listeners’ feedback may provide useful orientation for practitioners, teachers and aspirant interpreters”.

Encouragingly, researchers and institutions have already made efforts in setting up quality standards to guide various assessment and evaluation activities based on multiple perspectives including the interpreting provider services, researchers, professional interpreters (Amini, 2015; Grbić, 2008; Kurz, 2001; Iglesias Fernández, 2013) and interpreting audiences (Pöchhacker, 1994; Shlesinger, 1994). For example, Shlesinger (1994) conducted an experiment on audiences’ perceptions of interpreters’ intonation. Gile (1990) conducted an experiment where 23 delegates at a medical conference were invited to assess the quality of interpreting into English and French (see Bartłomiejczyk, 2014, p. 250). Kellett (1995) pointed out that the interpretation process’s end-product is the primary basis of the audience’s evaluation of both SI and CI interpreters’ performance.

As mentioned previously, it is critical to establish an overarching professional quality guide based on the assessment of interpreting end users in order to improve service standards across the profession and to accomplish the goal of facilitating communication between interpreters and end

users. The next subsection reviews two quality assessment guides that have been published by worldwide standards organisations.

#### **2.1.4.2 Interpreting quality assessment benchmarks**

Several interpreting quality evaluation benchmarks have been produced and published over the last 4 decades by standards bodies including ASTM International and the ISO. ASTM International, formerly known as the American Society for Testing and Materials, developed a *Standard guide for language interpretation* (ASTM International, 2008) in order to establish a minimum standard for quality services in the language interpretation industry with respect to the distinctive characteristics of specific settings and to set up a process for making informed decisions about interpretation services. The guide identifies the components of quality language interpretation services and establishes criteria for each component, for example, the command of languages and the skills and abilities required in an interpretation process. These interpreting skills and abilities are a major quality component assessed in obtaining a degree, certification and professional experience. The ASTM criteria and definition of interpreting skills and abilities include attention skills, analytical skills, memory skills, language transfer skills, note-taking skills, cultural awareness, clear and audible speech, a demeanour appropriate to the setting, adaptability, stamina to provide high-quality interpretation for the duration of the assignment and the ability to cope with stress.

The ISO has released multiple documents in relation to interpreting. For example, ISO 18841 (International Organization for Standardization, 2018) is an international standard providing the basic requirements for interpreting services, with additional recommendations for best practices. ISO 18841 can be used by language service providers to demonstrate a commitment to providing a high-quality interpreting service. It was drafted to provide requirements and recommendations for the delivery of spoken or signed communication across languages and societal contexts and in all interpreting specialisations. This international standard may be used in conjunction with other more specialised interpreting standards. The general interpreting competence that is specified in the ISO



standard includes interpreters' linguistic competence, intercultural competence, interpersonal competence, technical competence, competences in research, and information acquisition and domain competence.

The criteria for interpersonal competence require interpreters to prove effective communication skills to cope with the work demands in challenging situations amid cultural and linguistic diversity. Interpreters are also required to demonstrate mastery of the skill of stress management along with other skills including sight translation, memory and note-taking in demonstrating linguistic competence by conveying a message from the SL accurately, faithfully and impartially into the TL appropriate for a given setting. In the standard of domain competence, interpreters are asked to show their ability to communicate and express ideas well and also their ability to self-monitor and self-correct and provide appropriate spoken delivery of interpretation. Table 2.1 outlines the quality standards that are required by the ISO and how these correspond to those listed by ASTM.

Table 2.1 Attributes of spoken language interpreting quality standards from ISO and ASTM.

<b>ISO (general interpreting competence)</b>	<b>ASTM (skills and abilities required by interpretation process)</b>
Linguistic competence	Attention skills, analytical skills, memory skills, language transfer skills, note-taking skills, ability to cope with stress
Intercultural competence	Cultural awareness
Interpersonal competence/domain competence	Clear and audible speech Demeanor (professional appearance) appropriate to setting, adaptability, stamina
Technical competence	Technological needs
Competence in research and information acquisition	Knowledge of subject matter

The attributes of the spoken language interpreting quality standards from the ISO and ASTM are aligned with Pöchhacker's interpreting service-product duality model (2001), which defines multidimensional criteria for interpreting products and services. As illustrated in Figure 2.1, the quality standards for interpreting range from the fundamental layer of accurate linguistic rendition of the source speech to the higher levels of adequate expression and equivalent intended effect. The

highest layer depicts a successful communicative interaction supported by all the competence criteria. In Figure 2.1 linguistic competence, including attention skills, analytical skills, memory skills, language transfer skills, note-taking skills and ability to cope with stress, is at the core of the accurate rendition of the source speech. Other criteria, such as intercultural competence and interpersonal/domain competence (i.e., clear voice and speech, professional appearance, adaptability to settings and stamina), appear in the following two layers of adequate expression and equivalent intended effect. The outermost layer that illustrates a successful communicative interaction consists of all the other requirements such as technical competence and competence in research and knowledge of the subject matter. This convergent quality standard suggests that interpersonal competence, including clear spoken-voice production, professional appearance and good coping strategies, is essential for interpreting production and creating the intended SL effect.

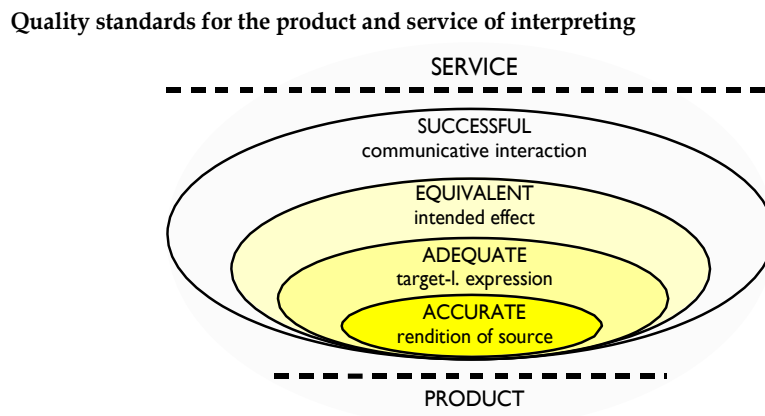


Figure 2.1 Quality standards for the product and service of interpreting

(Pöchhacker, 2001, p. 412)

#### 2.1.4.3 Self-evaluation in quality assessment

Two types of quality evaluation are distinguished in the context of interpreter training and education. The first type is external evaluation by various stakeholders, like the quality standard mentioned in Figure 2.2. The second type is self-evaluation by an interpreting trainee, which is seen as “a training tool designed to help the trainee in question in the further development of his or her

interpreting skills” (Bartłomiejczyk, 2014, p. 250). Self-evaluation is an essential tool to maintain interpreting quality.

Self-evaluation by interpreting students was first explored empirically by Russo (1995), who aimed to “encourage SI students to analyse their performance, discover their weaknesses and strengths and channel their resources during the training period accordingly” (p. 75). Many interpreting scholars have since noted that a clear explanation of the assessment criteria used in exams enhances learners’ autonomy and may have considerable influence on the quality of students’ work (Struyven et al., 2005). These assessment criteria can be made available and explained to students before an interpreting assessment (Bartłomiejczyk, 2014). The students then conduct self-practice and self-evaluation by listening to their own recorded interpretation and assessing their own performance based on the given assessment criteria.

Self-evaluation has been suggested to be a useful method of quality control to be practised both by interpreting trainees (Gillies, 2004; Heine et al., 2000; Van Dam, 1989) and by professionals (Déjean le Féal, 1990). Heine et al. (2000) stressed that interpretations should be recorded on a regular basis and evaluated by the student in accordance with the criteria introduced by the trainer in interpreting classes. Van Dam (1989), who offered an introductory course in SI, advocated homework assignments in which students are required to interpret the same speech several times, recording and listening to each interpretation until they are satisfied with their performance. Students are also supposed to prepare a written analysis of the last interpretation and hand it in together with the recording. Gillies (2004), in turn, advised trainee interpreters to listen to their output sentence by sentence or paragraph by paragraph and to transform each fragment, introducing all necessary changes in terms of grammar, idioms, intonation and structure, to arrive at “a version you might have expected to hear from an articulate native speaker giving their own speech” (p. 44). He also suggested making a transcript of one’s own output in which the spacing between words reflects the length of pauses between these words in the interpretation.

Interpreters' self-evaluation of their interpreting performance is one of the metacognitive and self-reflective learning strategies which helps students to become more conscious of their interpreting output. Déjean le Féal (1990, p. 156) enumerated the following aspects as the most noticeable in self-evaluation:

Since we can now fully concentrate on our output, we can evaluate its coherence, clarity and precision, as well as its language and oratory quality. Improper wording, unorthodox syntax, constant use of the same terms, slips of the tongue that remained uncorrected, numerous 'uh's' and 'um's' ... become glaringly obvious.

Despite the advantages of self-evaluation as a method of assessment, there remains the question as to how and to what extent this can be translated into actual improvement in the quality of students' interpretation. Perez et al. (2003) proposed that a structured assessment format is useful to increase the effectiveness of self-evaluation and suggested that the first step is to determine which quality components are most amenable to being improved by self-evaluation. For this purpose, they carried out an experiment which used responses from different groups of listeners to develop comprehensible and user-friendly "feedback grids" for students to evaluate their own and their peers' interpretations. The study found that students tended to focus on negative aspects in their linguistic transfer competence and devoted very little attention to their vocal presentation, which is one aspect that can be improved relatively easily if the student becomes aware of the relevant problems (such as filled or overlong pauses, a hesitant voice, monotonous intonation, false starts, etc.) and consciously strives to overcome them. Perez et al. (2003) pointed out that students' negative comments on their own output reflect the feedback they usually receive from their interpreting trainers, who tend to focus excessively on the negative aspects of interpretations and draw their students' attention to the errors while neglecting the presentation aspects.

#### **2.1.4.4 Language directionality and interpreting quality**

An A language, according to the AIIC definition (Glossary, n.d.), refers to the "interpreter's native language (or another language strictly equivalent to a native language) into which the interpreter

works from all the other languages in both modes of interpretation, simultaneous and consecutive”. A B language is defined as “a language other than the interpreter’s native language of which she or he has a perfect command and into which she or he works from one or more of her or his other languages”. In this study, “native language” refers to linguistic dominance, and not to the chronology of acquisition or the speaker’s cultural or ethnic heritage. An A language is also known as L1 and a B language as L2.

Quality assessment in A to B language interpreting is a debate that has occupied the attention of scholars and practitioners in the translating and interpreting industry for years. However, no consensus has been reached or research evidence has been produced to show whether quality in A to B language interpreting is superior or inferior to quality in B to A language interpreting. Some hold the opinion that B language interpreting should be avoided whenever possible. For example, the AIIC is well known to favour native production and advise against the practice of B language interpreting. Others hold that native comprehension of the SL is more important than native production and that fidelity to content is more important than the elegance of speech (Gorton, 2012).

Therefore, it is hard to decide whether native comprehension or native production results in a better rendition of the source text. A study conducted by Lee (2003) found that interpreting students tended to make more comprehension errors when working from the B into the A language and more linguistic and presentation errors when working from the A language into the B language. Donovan (2011) mentioned in an AIIC conference speech that interpreters have a “comprehension bonus” when interpreting into a B language. Their comprehension of the SL test is more intuitive and natural. However, the quality of interpreting trainees’ production is always going to require greater cognitive effort in interpreting into their B language. Interpreters must deal with an “expression deficit” due to lacking means of expressing themselves in the B language. In A language interpreting, the situation is simply reversed. The interpreter enjoys an “expression bonus” but can potentially suffer a “comprehension deficit”. As such, in this research project it may be more

important to conduct spoken-voice intervention for interpreters who work into their B language, especially in relation to the aspects of accent reduction, improving intonation and reducing cognitive effort and excessive tension.

Regardless of the debate, interpreting from A to B languages is an unavoidable trend with accelerated economic globalisation. As Gile (2009) pointed out, in some Asian and European countries, working in both language directions is unavoidable because there are not enough interpreters who are native speakers of these languages, such as native English-speaking interpreters. However, Gile (2009) recognised that the advantage of working into the B language is that the interpreters will have good comprehension of the source speech, although they may lack the high quality of interpreting delivery and expression which is possible only in their native language. Looking at the Mandarin–English language pair as an example, due to the huge demand for interpreting work from Mandarin to English in the market, most interpreting assignments from Mandarin to English are done by Mandarin A interpreters. Therefore, interpreting into the B language has been accepted as part of the interpretation curriculum (Lim, 2008), at least in China and other East Asian countries, to enable students to acquire strong skills when working into their B language.

In summary, Subsection 2.1.4 has discussed the interpreting quality standard from different perspectives, which has established the foundation for this research project to investigate interpreting students' spoken-voice quality from different perspectives. Also, it has shown that investigating interpreters working in their B language may be more relevant for this research project, particularly in relation to accent reduction, improving intonation and reducing cognitive strain and excessive stress.

## **2.2 Spoken-voice pedagogy applied in interpreting settings**

In Section 2.2, the essential concepts of spoken-voice pedagogy and their application to interpreting teaching are examined in order to explore the spoken-voice training approaches that may be useful

for improving interpreters' vocal health and quality, as well as to investigate the spoken attributes related to interpreters' work. The following subsections discuss the relevant spoken-voice literature pertaining to interpreting training. In subsection 2.2.1, the terms "voice," "speech," and "speech behaviour" are defined, setting the scene for the discussion of the subsequent subsections and the spoken voice training research in this study. Subsection 2.2.2 discusses the development of a vocal pedagogical system within the interdisciplinary scope, potentially in the training of interpreters' vocal skills, in an effort to bridge the gap between spoken voice pedagogy and interpreting studies. The relationship between language directionality and spoken voice training approaches is examined in subsection 2.2.3, providing direction for the study methods of this project. Subsection 2.2.4 explores several spoken-voice training approaches and techniques pertinent to interpreters' vocal training in order to investigate the spoken-voice training approaches that may need to be adapted to fit the special needs of interpreting work. Subsection 2.2.5 discusses spoken-voice categories and attributes closely related to interpreters' work and performance to lay the groundwork for the methodology development of this study.

### **2.2.1 Definitions of voice and speech**

The term "voice" has many definitions, depending on context and perspective. According to Kreiman and Sidtis (2011), voice is defined as laryngeal-supraglottal phonation in a narrow sense, while in a broader sense voice pattern refers to all the auditory-acoustic parameters that listeners use to make judgments about a speaker's characteristics and intentions. These parameters may include fundamental frequency, intensity, temporal cues including pausing and rate of speech, voice quality such as sounding harsh or breathy, and articulatory features of pronunciation.

The term "voice" is often used interchangeably with "speech" by outsiders to voice professions, which causes much confusion. For example, people often comment that they do not like their own voice. However, often what they mean is that they do not like their speech. Thus, clarification is necessary. Voice is the principal tool for forming speech, but voice is not always produced as

speech. For example, infants babble, animals howl, people create sounds when they cough or scream, and singers use the voice as an instrument for creating music. Human voice production is generated by 3 parts: the lungs, the vocal folds within the larynx (also referred to as the vocal cords or voice box) and the articulators (the speech organs). Everyone's voice is unique due to the actual shape and size of an individual's vocal folds, vocal tract and body. The sound is produced when the vocal folds in the larynx, the primary source of human sound production, are vibrated through adequate airflow and air pressure produced by the lungs. The vocal folds interact with the articulators and produce a highly sophisticated sound range (Stevens, 2000).

Voice is an essential tool for paraverbal communication, which is the middle of the 3 levels of communication – the other two being verbal and non-verbal communication (Heyne, 2013). Paraverbal means “parallel with verbal” (Heimlich & Mark 1990). This type of communication refers to an interaction between people through voice modulation (Iglesias Fernández, 2007), including tone, timbre, volume, pitch and inflection. Williams and Stevens (1972) proposed the concept of “vocal images” which reflects the individual personality, emotion, attitude and auditory identity. The tone of voice may be modulated to suggest emotions such as anger, surprise or happiness (Smith et al., 1975). For example, when people are confident, their voice sounds clear and well enunciated, but when people are nervous or fearful, they tend to speak faster, louder and often in a high-pitched tone.

Speech is produced by precise muscle-coordination actions in various parts of the body such as the head, neck, chest and abdomen, and by a series of complex movements that alter and mould the fundamental tone created by the voice into specific decodable sounds (National Institutes of Health, 2002). Through speech, humans express thoughts, feelings and ideas verbally to one another. As such, Grant (2011) pointed out that “human speaking shares an affective basis with the soundings of other animals. This affective basis is the ground of communication. It exists to affect others of its kind, to make a like body resound with feeling” (p. 6). Thus, feeling, passion and effectiveness are



vital for human speech development. The embodied, performative, affective, expressive and creative aspects of human speech are worthy of investigation and should be considered distinct from language.

Speech development is a gradual process that requires years of practice. As Turner (2003) pointed out:

Voice is instinctive, and speech is an acquired habit ... The child does not have to learn how to cry and croon, but speech is the result of the much laborious experiment, which is forgotten as soon as the movements of the tongue and lips have been repeated a sufficient number of times to set up a habit (p. 1).

Speech is the production of sounds that form words and sentences. It involves the precise coordination of the articulators such as the jaw, lips, tongue, teeth and vocal apparatus including the vocal cords, vocal tract and respiration. Humans are the only species on the planet that can communicate by producing speech (Hockett, 1960). Humans use speech to express thoughts, feelings and ideas to each other. Other animals vocalise but do not create speech by modulating vocal and speech organs. The most notable difference between speech and voice is that speech, unlike such vocal functions as crying or laughing, is a learnt activity (Virgil, 1967). An individual would not speak unless they were taught to employ their speech organs in childhood. Thus, speech is a function acquired by human beings to fulfil their needs for social and psychosocial adaptation and commitment. As Hahn et al. (1957) pointed out, speech is an acquired skill and speech sounds are habitually formed and articulated under the influences of home, associates, school and location. If someone's speech is poorly learnt and formed with improper speaking and vocal habits, or their speaking skills lack constant attention and exercise, the person's speech capacity is likely to become weak and affect their personal confidence and communication ability.

Despite the difference between voice and speech, they have many interdependencies. Horváth (2017) defined the term "speech behaviour" as the vocal form of a message. Speech behaviour refers to the

way people speak and utilise their voice. Body language, such as eye contact, gestures and physical appearance, is a component of speech behaviour and serves as a means of communication since it delivers a message. An interpreter's speech behaviour fulfils somebody else's communicational needs and purposes, and transfers the content and the message of a speech somebody else creates. Therefore, interpreters' speaking motivation is different from that of the original speaker, although they share the same communication goals. Horváth (2017) pointed out that, since interpreters have at least two functional languages which they use during their work, their speech behaviour needs to be:

as seamless as their communication behaviour to convey not only the message of the source language speech but also their professionalism ... The fact that interpreters are secondary speech producers may have several implications on their speech behaviour in terms of voice production, paralinguistic features such as intonation, speech rate, the pitch of voice and voice identity (p. 228).

To sum up, this study does not intend to study voice and speech as individual and separate entities. Instead, the study endeavours to study voice training for speech and spoken-voice training. The latter is often referred to as the current voice and speech pedagogy.

### **2.2.2 Voice training as an area of interdisciplinary pedagogy and research**

Spoken-voice pedagogy has been developed by interdisciplinary teams of otolaryngologists, speech-language therapists, voice trainers and singing teachers collaborating in the prevention, treatment and remediation of voice and speech problems, and training in the production of "beautiful and appropriate speaking voices" (Melton, 2012, p. xi). In recent years, spoken-voice pedagogy has been employed by voice coaches, speech therapists and elocutionists who work with those professionals who rely on voice as their primary tool. Specialist voice training for the acquisition and correction of accents, dialects and languages other than the performer's native tongue has also been developed (Wilson, 2011).

Before the mid-1950s, spoken-voice pedagogy focused primarily on elocution and voice projection (Barton, 2015). From the 1960s, various spoken-voice training techniques and approaches started to stand out. As such, a multidisciplinary approach to voice pedagogy has enabled both voice teachers and students to gain better understanding of a broader spectrum of voice training approaches and helped them discover numerous methods and strategies by which to achieve successful voice transformation. The development of a vocal pedagogical system has manifested in the following interdisciplinary scopes.

Firstly, the development of spoken-voice pedagogy is both a scientific and an artistic exploration. Science and art in voice work are not only compatible but also indispensable to each other. Thus, both scientific and aesthetic aspects must be taken into consideration in voice pedagogical approaches (Appelman, 1974). Voice pedagogy approaches based on aesthetics alone would produce a limited perception of voice work. Providing scientific evidence for voice training is becoming increasingly necessary to achieve the best training result possible. As Shewell (2009) put it, “all voice practitioners, whatever their disciplines, must base their intervention on a sound understanding of the anatomy, physiology and biomechanics of phonation if they are to practice safely and with maximum effect” (p. viii). On the other hand, a solely scientific approach to voice training could be detrimental to the beauty and efficiency of sound produced by students because of a lack of expression and emotion.

Secondly, over the past several decades some voice educators and researchers have called for an integrated pedagogy of speaking and singing, and argued there are more similarities in the two areas than there are differences (Thurman & Welch, 2000). For example, Melton (2012) stated, “Singing and speaking are accomplished by the same instrument; we can move easily from one activity to the other in the same breath; and speaking and singing onstage have similar technical requirements” (p. 135). Because speakers and singers share the same vocal health and quality concerns, basic principles and training techniques of vocal development and management for singers have been

adopted by teachers of the speaking voice (Melton, 2012) for those who use the speaking voice professionally, including actors, professional public speakers and radio and television announcers. Furthermore, since singing is a performing art, voice instructors spend time preparing their students for performance, including the etiquette of behaviour on stage such as bowing, managing stage fright, addressing problems like nervous tics and learning the use of equipment such as microphones, acting techniques and audience communication. These may shed light on the best type of spoken-voice training for interpreting students.

Thirdly, the modern spoken-voice pedagogy system has adopted a holistic approach, studying both physiological and psychological aspects of the vocal mechanism. Voice is a “two-way psychosomatic phenomenon” (Shewell, 2009). As a bridge between our inner and outer worlds, our voice can affect listeners’ physical and emotional states. According to modern somatic education techniques, for example the Alexander technique, the whole human body is a great voice instrument and therefore the quality of the human voice is achieved by the coordination of all muscles and the skeleton that are used in breathing, phonation and articulation (Jain et al., 2004; Siple, 1997; Turner, 1993).

Apart from being a manifestation of the human physical structure, voice is also an indicator of people’s mental state and emotion. The whole body reacts during emotion – breathing is affected, muscles in the larynx and the vocal-fold region tighten and change occurs in the throat and mouth. Attitudes are emotional and will be apparent in physical changes and feeling states. If people are enjoying a situation, that enjoyment is evident in their muscle tonus, in what is called the “aliveness of the whole body” (Hahn et al., 1957, p. 20). If a situation is unpleasant, the muscles tighten, breathing and heart rates increase, and tension is felt within the body as if certain processes have been interrupted. This interconnected body and mind situation is described as “psychophysical processes” (Thurman & Welch, 2000).

Because such neuromuscular activity produces voice, the work of spoken-voice training is fundamentally physically based on the acquisition of psychomotor skills (Wilson, 2011). The intensity of voice work required by professional voice users can be considered an “athletic feat” (Rodenburg, 2015). Thus, some of the findings of sports science in relation to neuromuscular skill acquisition, performance rehearsal and rehabilitation protocols are employed to enhance the effectiveness of voice performance training, for example, pre-performance voice warm-ups and appropriate warming down (Novak et al., 1991) and protecting the laryngeal system from “unwanted changes related to vocally violent manoeuvres” (Roy, 2000, p. 226).

Some “mindfully movement-related voice work” such as the Alexander technique (Jain et al., 2004; Siple, 1997) are also added to voice-training programs following the proposal that mindful movement and voice need to be related and integrated for all voice users (Heirich, 2011; Melton, 2012). These techniques are practical and straightforward methods for improving ease and freedom of movement, balance, support and coordination. The techniques can be applied to sitting, lying down, standing, walking, lifting and other daily activities.

Like any other neuromuscular skill acquisition task that is scientifically proven to alter the function and structure of the adult human brain and behaviour, the integrated voice-training approach profoundly affects trainees mentally, psychologically and physically. Exploratory research (Mithen & Parsons, 2008) showed that a quantifiable number of changes in neural activity occur in the early stages of acquiring vocal (singing) skills in an adult student. Functional magnetic resonance imaging brain scans conducted before and after vocal training showed significant increases in brain activity in the right hemispheres of the participants’ brains. This result confirms many other scientific findings that demonstrate “hemispheric lateralization” of sung and spoken language (Jeffries et al., 2003, p. 752). As Mithen and Parsons (2008) pointed out, with continued practice “the higher levels of skill development may be associated with bilateral brain mechanisms” (p. 420). These studies demonstrate that adults’ behaviours and bodies can be trained and changed by

acquiring new neuromuscular skills. In other words, spoken-voice training may help interpreting students change their habitual voices by building new structures and pathways in their brains. This is a gap that this research project attempts to fill.

### **2.2.3 Language directionality and vocal features**

Over the past few decades, there have been several empirical bilingual voice studies undertaken to understand the effects of language directionality on the acoustic aspects of speech and how the use of someone's A language and B language may pattern with certain vocal attributes. The studies have demonstrated that language directionality can affect some of the vocal features of bilingual individuals when they speak different languages. For example, Keating and Kuo (2012) compared the fundamental frequency (F0) of Mandarin and English monolinguals and found that the Mandarin speakers had a higher and larger F0 range than that of the English speakers. F0 is the frequency at which vocal cords vibrate in voiced sounds (Li & Jain, 2009). F0 is used as an acoustic measurement that reflects the rate of vocal-fold vibration (Baker et al., 2008), intensity (volume), speech rate (Lee & Sidtis, 2017; Wagner & Braun, 2003) and timbre (vocal brightness and roughness) (Wagner & Braun, 2003) across various language and ethnic groups. Ng and Chen (2010) examined the vocal attributes of Cantonese and English produced by Cantonese–English bilingual speakers by using “the long-term average speech spectra” (Löfqvist & Mandersson, 1987). The long-term average speech spectrum has been utilised for research on the human voice source and provides information about the speech signal's spectral distribution over time. The study (Ng and Chen, 2010) concluded that the same vocal apparatus was used differently by Cantonese–English bilingual speakers when they spoke different languages. Chong (2012) investigated the effects of language and race (with bilingual English and Mandarin speakers) on voice quality using various acoustical parameters and found that the language used by an individual could affect voice quality although the same phonatory apparatus was used. Nevo et al. (2015) studied the impact of language on vocal features among adult bilingual Hebrew–English speakers and found that there

were changes in resonance, <sup>1</sup>glottal attack, F0 variation and speech rate when they talked in Hebrew as compared to English. A study of perceptual analyses by Järvinen et al. (2017) also concluded that speakers' voice quality can be different in their native and foreign languages.

The above studies demonstrate that bilingual speakers produce notably different voice patterns in their two spoken languages across different speech tasks and speaking a second/foreign language can cause subjective sensations of vocal fatigue and vocal overloading due to the fact that a foreign language requires more mental and physical effort. Järvinen et al. (2017) explained that one reason for vocal overloading could be that speaking a foreign language affects phonation, which becomes more compressed. Studies have also found that speakers' L2 proficiency may impact on their vocal quality modification and vice versa, and that speakers' spoken-voice quality can also affect their listeners' impression of their language proficiency (Järvinen et al., 2013, 2017; Lee & Sidtis, 2017). For example, a slow speech rate may give listeners an impression of hesitation, a heavy accent is regarded as less competent, less reliable and less interesting than the performance of native speakers (Tsurutani, 2012) and a high pitch (F0) may indicate psychological anxiety, contributing to tension in the laryngeal structures, while a low and soft speaking voice can signal uncertainty and insecurity (Järvinen et al., 2013).

Despite the fact that spoken-voice training is not language-independent, the focus of the training approach for a native speaker and a speaker of a second language may be distinct. As Hlavac (2022) noted, it is easier to demonstrate fluency, grammatical accuracy, and rhetorical skills in one's A language, whereas the interpreter's interpretation into the B language may contain production flaws. Although current studies on interpreters' voice quality focus primarily on interpreters working into their A language, results from studies on the phonological features of the B language indicate that L2 speakers have a narrower pitch range than A language speakers (Zimmerer et al., 2014, see Halvac 2022). This suggests that the training of interpreters' spoken production skills for those who

---

<sup>1</sup> A glottal attack is produced when the vocal folds come together before a pitch is produced and air pressure builds behind the vocal folds. Hard or abrupt glottal attack is regarded as one of the most serious types of vocal abuse. (Andrade et al., 2000)

work in their B language may require a greater emphasis on distinctive voice and speech features, such as resonance and pitch range, prosody, fluency, articulation, intonation, and volume.

#### **2.2.4 Spoken-voice training approaches relevant to interpreting training**

Over the years, many vocal training approaches and techniques have been designed for actors, singers and other professional voice users. These techniques and approaches may need to be re-evaluated and modified to meet the specific demands of interpreting work and, when learnt, to be adequate for interpreting situations. Following is discussion of some spoken-voice training approaches and techniques that may be relevant to interpreters' vocal training.

##### **2.2.4.1 Intensive voice training model**

An intensive mode of delivery, also known as block mode teaching, is where course materials are delivered over a shorter period of time compared to standard courses by means of compressed teaching formats (Karaksha et al., 2013). The intensive teaching mode is often used for teaching credit-bearing university units and professional development courses for postgraduate students and professionals in industry (Greer et al., 2016). Intensive teaching has been popular with tertiary education institutions over the past decades and the amount of academic research in the area is growing (Davies, 2006). Many studies have compared the efficacy of intensive and traditional modes and suggested the following advantages of the intensive mode (Kucsera & Zimmaro, 2010). Firstly, the high level of student satisfaction is the primary factor driving the implementation of this teaching mode within tertiary education institutions (Karaksha et al., 2013). Secondly, compared with traditional weekly lectures and workshops or daytime teaching practices, the intensive mode teaching is convenient for both teachers and students who require more flexible modes of delivery that fit with demands at work and home (Curtis, 2000). Thirdly, studies have reported improved learning performances and outcomes due to increased motivation, commitment, concentration level and strong interaction among students who selected intensive mode teaching (Davies, 2006;



Karaksha et al., 2013). Fourthly, the intensive teaching mode has brought diversity and flexibility of teaching methods from instructors (Davies, 2006).

Despite these well-recognised benefits of the intensive teaching mode, some researchers have warned that the greater intensity of workload and fatigue due to longer teaching sessions (Traub, 1997; Wolfe, 1998) should be taken into consideration by instructors when planning and designing for intensive teaching (Davies, 2006). Lee and Horsfall (2010) recommended that instructors should ensure all students can engage with each other and the instructor in interactive learning in their teaching design and delivery.

Many different models of intensive teaching mode exist, for example block mode and sporadic mode (Davies, 2006; Male et al., 2016). The former involves very large chunks of teaching time over short periods (e.g., more than 60 minutes per session over a few weeks), while the latter offers smaller time chunks over longer periods (e.g., 18 days or 5–10 weeks). In this research project, both block mode and sporadic mode are considered in the design.

#### **2.2.4.2 Perceptual learning and optimal training regimens**

The term “perceptual learning” was initially defined by the psychologist Gibson (1963) as “any relatively permanent and consistent change in the perception of a stimulus array, following practice or experience with this array” (p. 29). Gibson’s definition of perceptual learning involves 3 aspects: first, long-term changes in perception; second, changes in perception rather than belief; and third, permanent changes in practice or experience because of the learning process.

Based on Gibson’s (1963) study, Goldstone (1998) proposed that perceptual learning improves a learners’ ability to respond to a changing environment. Molloy et al. (2012) also conducted a study to understand the effects of different perceptual training regimens such as the length of each training session and the space between training sessions on perceptual learning outcomes. The findings suggested that within- and between-session learning was an independent learning process during training progress and therefore efficient regimens should use shorter and less intense training

sessions and optimised spacing (e.g., a few days) between sessions as these allow for more latent between-session and post-training learning to emerge, especially in the early learning stage.

In addition to Molloy et al.'s (2012) study, Dam and Diaz (2013) indicated from a cognitive neuroscience perspective that short but more frequent training sessions were more effective than long training. Van Dam (2013) also pointed out that a training session should ideally be no more than 20 minutes long because the brain finds it easier to retain the first and last parts of a training session rather than the information presented in the middle of the session. This rule would be applicable to all types of learning. This study attempts to test the above theories.

#### **2.2.4.3 The visual auditory kinaesthetic learning model**

According to Fleming's (2001) visual auditory kinaesthetic learning model, most people possess a dominant or preferred learning style, although some people have a mixed and evenly balanced blend of these 3 styles. A study by Barton (1997) showed that 80 per cent of acting students at the introductory levels were visual learners, while an increasing number became auditory learners and even kinaesthetic learners at the advanced levels. Barton (1997) explained that these students had not realised their potential until they had been encouraged to respond to other modes of learning because our educational system strongly favours visual learners.

According to Barton (1997), actors who are predominant visual learners have their drawbacks. Their bodies may be often stiff when they work, showing little movement. They present as raising their shoulders, having shallow breathing, a high and unvarying pitch, and a rapid speaking tempo. They pull their chins down into their throats to intensify their eye contact with the audience. Often they are not aware of their body when they work. Because they favour visual cues, they take many notes while listening and then deliver their message while reading from their notes. Their speech is highly organised and orderly, but lacks vitality and expression. The audience may experience frustration because of their shallow breathing, the appearance of tension, their narrow pitch and their literal and expressionless speech.

On the contrary, as Barton (1997) pointed out, “Auditories often excel at the full range of voice and speech work because they hear sound distinctions others miss” (p. 84). Auditory learners can easily repeat what others have just said, almost word for word. They generally speak easily and often talk silently or aloud to themselves. They think in words rather than pictures. They learn better by listening, discussing, reading aloud, and using and playing back audio recordings. They have high auditory skills demonstrated in listening, speaking, storytelling, explaining, understanding the syntax and meaning of words, remembering information, arguing their point of view and analysing language usage.

Kinaesthesia refers to the awareness of body movements. It is the perception of and the quality of our body movement. Kinaesthesia is often referred to as the 6th sense. Although it is as important as the other 5 senses, it is often forgotten or unconsciously turned off (Malde et al., 2020). For example, people may be sitting or standing in an unbalanced position for hours in a tense work situation until their legs are numb and their neck hurts. Training in kinaesthesia is a core component of the training of actors because they need this capacity to build different characterisations. Since interpreting work is regarded as highly intellectual and involves less physical movement than acting and singing, the relevance of kinaesthetic awareness is rarely seen in interpreting training literature. However, interpreting work requires both high physical and mental stamina.

There has been very little research done so far on the predominant learning styles that most interpreters possess. It would be helpful for an interpreting instructor to know what learning style each interpreting student in the class possesses to achieve a better learning result. However, like actors, it is highly likely that interpreters develop auditory and kinaesthetic sense through long-time language and interpreting training even if they are not endowed with a refined auditory sense.

Interpreters are often too busy with cross-language transfer and fulfilling various cognitive tasks to feel and notice their mental and physical sensations and tensions, which eventually leads to stage fright. Kinaesthetic awareness can be tuned through experimenting with different movements and

procedures such as the semi-supine position in the Alexander technique or certain movements in Pilates. These exercises are the first step towards “freeing the natural voice” and developing the ability of “psycho-physical awareness” (Linklater, 2006, p. 31) or a more refined kinaesthetic and proprioceptive awareness, also referred by some Alexander technique teachers as “inclusive awareness” (Malde et al., 2013, p. 5).

Inclusive awareness training is a powerful tool that helps learners produce effortless and quality vocal performance (Malde et al., 2020). According to Malde et al. (2020), inclusive awareness is “the skill of perceiving self and world simultaneously” (p. 11) and “the singer’s most effective mode of attention” (p. 13). By applying inclusive awareness to performers’ training, they may simultaneously use all their senses and kinaesthesia to be aware of themselves and their surroundings in an interpreting situation, for example, their body movement, their emotional state, their voice, their gestures, the text, their audience, other interpreters, the venue, the stage, the room temperature, lighting and so on. With regular practice, as their inclusive awareness strengthens they will respond effortlessly to the challenges of their physical and mental surroundings when required.

Inclusive awareness training is helpful for interpreters’ performance training because this awareness will transfer their language and interpreting skills into better interpreting vocal performance. As they perform routine interpreting activities, such interpreters become fully present to what is happening in their whole self and their surroundings, and overcome self-consciousness to become self-aware.

#### **2.2.4.4 Prosodic feature awareness training**

Apart from the training of interpreters’ performative aspects, prosody is an overlooked area that has drawn the attention of some interpreting scholars (Shlesinger, 1994; Williams, 1995; Ahrens et al., 2010). Prosody includes word and sentence stress, word tone, phrasing, pauses and intonation, all of which help listeners to process and understand the incoming message. Therefore, prosody plays an important role in speech communication (Cutler, 2012).

Bellon-Harn (2011) defined prosodic features as the parameters in the speech signal used to convey the communicative functions, including pitch, loudness, duration, intonation, stress and rhythm. As fundamental aspects of speech communication (Ahrens, 2004, 2005), prosodic features extend over short stretches of an utterance (i.e., syllables and words), over longer stretches (i.e., a phrase, clause or utterance) (Cruttenden, 1997) or across a conversational turn (Couper-Kuhlen & Selting, 1996). Prosodic features of speech are anything that cannot be understood through a direct linear sequence of segments and these features give melody and rhythm to utterances (van Heuven, 1994). Ahrens (2004) conducted computer-aided analysis of voice characteristics and prosody to investigate the impact of non-linguistic factors on the acquisition of prosody in a second language and called for greater cooperation between experts on voice and experts on signal processing in reaching conclusions about these aspects.

Prosodic feature awareness training is a specific training methodology involving drilling of word and sentence stress, word tone, phrasing, intonation and rhythm as in the correct pronunciation of sounds (Kayi, 2012). Prosodic feature awareness training can be targeted towards the needs of second language learners and those who want to gain comprehensible communicative abilities or near native-like language competence, which deals with the complexity of prosodic features.

Gut et al. (2008) claimed that instructors' goal in teaching prosody in different academic settings is to make L2 learners perceive and produce the prosodic features of the second language adequately. Gut et al. (2007) suggested that instructors should employ different methodologies depending on the native language and the TL. One way to improve prosody is for L2 learners to record their speech, listen to it, compare their imitation with the native model and repeat the exercise until they can no longer detect a difference between model and imitation.

Due to the advantages of prosodic awareness training, Yenkimaleki and van Heuven (2018, p. 256) recommended that prosodic features have “the utmost importance in the interpreter training curriculum” and “should be included in the curriculum of the training of future interpreters”. With

such a view, Yenkimaleki (2018) examined the effect of prosodic feature awareness training on interpreting quality. Their findings showed that a prosody training program could be successful in enhancing the quality of the speech output in CI from native Farsi into non-native English (Yenkimaleki & van Heuven, 2018), especially in relation to accentedness, pace and voice quality. Their results showed that prosodic feature awareness training had a statistically significant effect on the voice and speech quality measures. Yenkimaleki (2018) concluded that these results have implications for curriculum designers for interpreter training, interpreting instructors and all involved in foreign-language study and pedagogy.

#### **2.2.4.5 Alexander technique training**

The Alexander technique was named after F. M. Alexander (1869–1955), an Australian actor who developed the technique by identifying his pattern of misusing his voice and discovered a way of curing his hoarseness by “use of the self” (Alexander, 1932). The Alexander technique helps improve the professional voice for performance, dealing with specific voice and speech problems and disorders, and gaining insight into one’s own awareness (Gelb, 2013). In Alexander lessons, the unified body–mind or the psycho-physical process of voice production is taught. This includes dealing with the automatic reaction to using one’s voice and the underlying physical restrictions throughout the body contributing to less than optimal performance. Students are taught to be mindful of excessive muscular and mental tension levels during their professional and everyday activities through a series of practical processes that awaken their awareness of how they use their bodies and develop their innate potential. The technique also guides students to give up maladaptive physical habits and return to a balanced state of body alignment (Gray, 1990, pp. 14–20).

Teachers employ the Alexander technique to demonstrate, explain and analyse an individual’s responses using mirrors, video feedback or classmate observations and to give advice for improvements according to the individual’s specific issue. In early lessons, the emphasis is on postural states and basic movements, which are fundamental to the state of being. As students

progress, they become interested in applying the Alexander technique to their emotional or intellectual reactions and with practice these can be the most rewarding changes.

In this research project, the “semi-supine procedure” of the Alexander technique is considered in the selection of spoken-voice activities for interpreting student participants. The semi-supine procedure, also known as the “lying down procure,” involves specific vocal exercises in the Alexander technique as a means of releasing cumulative muscular and mental tension. According to Soar (1999), the semi-supine position is the most fundamental and essential approach to working on oneself; it provides a systematic activity through which to release excessive tension and become more conscious of the relationship between one’s head, neck, and back, which forms the “core” of body structure. In the semi-supine position, the students are instructed to lie face up with the raised position of their legs, heads, and hands and to learn to use the firm contact of the floor to inform them about the firm texture - the relative states of stiff tension, floppy relaxation, and springy expansion of their back.

The semi-supine position is an excellent position in which to practise Alexander’s principle of working on oneself by “kinaesthetic awareness”, “inhibition” of automatic reaction and “direction” of conscious intention (Alexander, 1932; Clare-Newman, n.d.). Through practice, one learns to be sensitive to one’s body motion, finding out what causes one to feel tense and responding kinaesthetically by bring oneself back to balance and becoming free of tension.

#### **2.2.4.6 Berry’s approach – integration of voice and text**

Cicely Berry (1926–2018) taught voice and speech to acting students and worked with professional actors to help them improve their spoken voice and find confidence and enjoyment in using their vocal apparatus. Her approach involved guiding students to feel their breath, ground the tone of their voice and release resonance without unnecessary effort. She devoted much attention to articulation, or muscularity as she termed it (Berry, 1973). Berry believed in the physical power of words and taking time to unlock the meaning of a text as fully as possible; as she stated, “through

the physical exercises they were able objectively to find the sound and force of those words and gain confidence in speaking to them” (Berry, 2001, p. 42).

Berry’s work often began with a thorough vocal warm-up that involved some movement and physicality. Following the physical vocal exercises, Berry would guide the acting students to explore texts on a deeper level through work on the body and the voice. The students were led to try out various exercises linked to breathing, support, tongue and lip action, and the release of sound through vowels and consonants (Berry, 1973). Berry often asked the students to practise a specific text in terms of structure and meaning to inspire them to become clearer and more confident in interpretation of the text. The practice of a text, rather than being from an intellectual perspective, was to encourage the students’ sensitivity towards the rhythm, pace and vocal projection in the flow of the phrasing (Berry, 1987).

Berry’s approach offers itself as a reference model for the development of spoken-voice training for interpreting students. An issue affecting interpreting students who are not native English speakers is that they may often be concerned about their non-standard English accent. In her book, Berry (1987) emphasised that actors need not restrict their accent as used in the text to a so-called Standard English accent or Received Pronunciation. She said, “it is the vigour in the language that is crucial, and not an accepted set of vowel sounds ... we have to be open to different speech patterns” (Berry, 1987, p. 33). She maintained that the actors should discover a sense of vigour in the language and confidence in expressing the ideas and feelings clearly, rather than adhering to a certain accent that is considered “better” than other English accents.

Berry’s voice training methodology can also be used to improve interpreters’ spoken-voice training. For example, Berry preferred free responses from working in a group by structuring all the vocal exercises for small groups sitting in a half-circle or individual work within the support and presence of the group (Berry, 1987). Most of the time interpreters work independently and lack support and



feedback from their colleagues. This method may help take interpreting students away from over-concentration on the self and transfer their attention to something external.

#### **2.2.4.7 Linklater’s psycho-physical approach**

As one of the most influential voice teachers in the world of actor and voice training, Kristin Linklater (1936–2020) developed a psycho-physical approach. Linklater’s spoken-voice training technique (2006, 2016) uses a series of body exercises, vocal warming up and vocal work to create a release in all the muscles of the body that produce voice and speech through standing, walking, bending and even lying on the back, which is particularly useful as a basic vocal warm-up exercise for interpreting students to reduce stress before going on stage. Further, Linklater’s techniques include improving students’ breathing awareness, feeling of the vibrations of sound and physical awareness of the spine, jaw, tongue, throat and soft palate.

A research experiment (Wessendarp, 2012) confirmed the Linklater voice-training technique allows students to explore their voices in natural body positions such as sitting and standing. This is in contrast to artificially creating sound and improving the airflow control and sustained phonation that may assist them to cope with long hours of speaking while maintaining vocal quality and energy level. Although Linklater’s approach and techniques were mainly developed for professional actors and acting teachers, her psycho-physical approach opens a new exploration of better vocal performance in the interpreting profession.

The voice-training techniques she taught offer a reference for interpreter trainers to develop what she called an “essential vocal craft” (Linklater, 2006, p. 5). Linklater believed that voice and speech training is an art and a craft. She pointed out that this craft cannot be acquired overnight by most people but through self-awareness, consistent self-regulation, and proper and effective training methods. This is particularly useful for interpreters’ spoken-voice training, which requires not only physical awareness but also emotional recognition in the process of releasing the natural voice.

#### **2.2.4.8 Shewell's approach – integration of science and art**

In her book, *Voice work: Art and science in changing voices* (2009), Shewell strongly advocated that both scientific and aesthetic aspects must be considered in voice pedagogical approaches because science and art in voice work are not only compatible but also indispensable to each other. Shewell encourages voice teachers and students to be curious scientists eager to explore and discover the voice in the human body and mind, and find beauty in it by using both evidence-based techniques and experience-based techniques. Shewell (2009) pointed out that providing the scientific evidence for voice training is becoming increasingly important to achieve the best training result possible because voice pedagogy approaches based on aesthetics alone produce a limited perception of voice work.

For this purpose, Shewell (2009) provided a detailed explanation of eight core parameters for voice skills (p. 71), including bodywork, breath, “channel” (p. 114), phonation, resonance, pitch, loudness, articulation, with the hope that voice practitioners can base their voice practice on an understanding of anatomy and physiology, and practise safely and with maximum effect on these types of voice work.

Simultaneously, a pragmatic perceptual voice scheme was developed with the intention that any voice practitioner could find it useful. Two options are available in the scheme: the Voice Skills Framework and the Voice Skills Perceptual Profile (VSPP). The Voice Skills Framework permits the practitioner to record key observations, explorations, and client feedback under each of the eight core voice skills headings. Also included is a section for action planning at the end of the form. The VSPP is designed for practitioners who wish to follow a systematic step-by-step approach through the eight core voice skills parameters with specific tasks and questions, the majority of which generate numerical ratings.

The scheme serves as a tool that enables both qualitative description and numerical evaluation and satisfies “both intuitive need and reason-based criteria” (2009, p.82). A voice practitioner could

note personal observations under specific voice skills parameter headings, using their own choice of vocabulary and comments. This qualitative information could help guide the direction of appropriate voice works for the one assessed. In addition, these parameters could be rated to generate quantitative data for pre- and post-training results in record-keeping, research, and comparison. The perceptual voice skills scheme illuminates the research design, specifically the Voice Skills Evaluation Sheet for this study.

#### **2.2.4.9 Implicit and explicit instruction**

Explicit instruction is defined as teaching in which the instructor explicitly identifies the student's learning objectives and provides transparent, unambiguous explanations of the skills and information structures being presented. On the contrary, implicit instruction is when the instructor does not overtly state goals or provide explanations, but instead presents the information or problem to the student and allows them to draw their own conclusions, construct their own conceptual structures and assimilate the information in the way that makes the most sense to them (Smith, 2017).

Implicit learning has been defined as learning that occurs without conscious effort and frequently with the lack of clear understanding of what was learnt (Reber, 1993; Wolf & Müller, 2012). From learners' perspective, implicit learning occurs when learners are unaware of the learning that has taken place without the use of central attentional resources. As a result, learners cannot express what they have learnt verbally, although they make clear behavioural responses. In contrast, learners who engage in explicit learning are aware that they have learnt and can articulate what they have learnt. Explicit learning often places a high demand on working memory (Ellis, 2009).

Although instruction and learning are deeply connected, Schmidt (1994) argued that learning must be distinguished from instruction. According to Schmidt, no evidence shows that implicit instruction will necessarily result in implicit learning or that explicit instruction always leads to explicit learning. Even though instructors may wish for a direct correlation between the instruction

and the learning, students may follow their own desire regardless of the type of training they receive (Allwright, 1984).

The importance of the implicit/explicit distinction, especially for language acquisition (both first and second language learning), has been tested and affirmed by research evidence (Ellis, 1993; Schmidt, 1990). As Ellis (2009, p. 6) put it:

Explicit learning and implicit/explicit knowledge are related but distinct concepts that need to be separated. Whereas the former refers to the processes involved in learning, the latter concerns the products of learning. It is possible, for example, that learners will reflect on knowledge that they have acquired implicitly (i.e., without metalinguistic awareness) and thus, subsequently develop an explicit representation of it.

Several studies have been done to compare the effectiveness of explicit and implicit instruction. The general finding is that explicit instruction is more effective than implicit instruction (Ellis, 2009).

For example, Yenkimaleki (2018) investigated the effects of explicit and implicit prosody instruction on the development of listening comprehension skills by Farsi–English interpreting trainees and found explicit instruction in the use of prosody led to a more significant improvement in listening comprehension skills for the trainees than implicit instruction in their speech perception and production. Case et al. (2018) investigated whether listeners demonstrated improvement in spoken-language processing due to implicit learning of a speaker’s voice after short-term exposure for two days. The results showed no significant effect on implicit voice learning. However, Kemper (2008) conducted two experiments on second-language learning and claimed that the effectiveness of explicit and implicit instruction depended on the type of the learner and what learning rules were applied. This research project is an exploration of these differences as they apply to a spoken-voice intervention for interpreters.

In summary, the aforementioned vocal training approaches and techniques were chosen to review because they are pertinent to interpreters’ vocal training and will serve as a guide for the design of spoken-voice training for interpreting students in this research study.

### **2.2.5 Spoken-voice categories and attributes**

A person's spoken voice can be influenced by many factors that cause their voice to be different to that of anyone else and different from situation to situation. These factors may include the structure of their vocal apparatus, general vocal health, emotional state when speaking, vocalisation and speaking habits and techniques. This section discusses 3 general spoken-voice categories: spoken-voice production, quality and dynamics, which are closely related to interpreters' work and performance. Each category contains 5 features or attributes. Some attributes may seem not to be directly related to voice and speech, for example, "posture and alignment", "body movement", "audience connection", but do play important roles in the general spoken-voice function (Arboleda & Frederick, 2008). The 3 spoken voice categories and 15 attributes laid the foundation for developing the research methodologies of this study.

#### **2.2.5.1 Spoken-voice production**

Professional actors must learn the anatomy and physiology of voice production to create a sound that is appropriate to the performance they are giving. Interpreters may not need to learn about voice production comprehensively in the way that actors do. However, it is possible that it would be beneficial for interpreters to learn about basic voice mechanisms, apply this knowledge to their interpreting work and gain awareness of how sound is produced. As Houseman (2015, p. 182) put it: "The reason for this work, as for all the other work on the voice, is to allow your (i.e., the trainee's) thoughts and feelings to be expressed fully and subtly." Following are several spoken-voice attributes identified as being essential or instructive in spoken-voice production for interpreters: posture and alignment, body movement, physical tension balance, breathing technique, release of speech muscles and vocal care.

### **2.2.5.1.1 Posture and alignment**

Posture is defined as “the attitude assumed by the body either with support during the course of muscular activity, or as a result of the coordinated action performed by a group of muscles working to maintain the stability” (Physiopedia, 2021). Gardiner (1957) referred to two different kinds of posture: the first is dynamic posture, which is how people hold themselves while moving. The other type of posture is static posture, which describes how people hold themselves while not moving, such as when sitting or standing.

According to the Merriam-Webster.com online dictionary, alignment means “the proper positioning or state of adjustment of body parts in relation to each other”, for example, how the head, shoulders, spine, hips, knees and ankles line up. Alignment is an essential part of body posture (National Osteoporosis Foundation, n.d). Both posture and alignment are concerned with positioning.

However, alignment contains the word “proper” in its definition. It is this concept that distinguishes the two. Proper alignment of the body puts less stress on the spine and helps good posture (Nutritious Movement, n.d.). As Melton (2012, p. 1) pointed out, “Alignment is the first step in voice training. The body is designed to move, so alignment is not about being in any position; rather, it is about moving with ease and using the body with maximum efficiency”.

Postural alignment is “the composite of the positions of all the joints and limbs of the body at any given moment. Optimal postural alignment is a prerequisite for optimal movement function” (Graeme, 2016). A person’s postural alignment, particularly their cervical spine alignment, affects the quality of their voice and their ability to pitch. Many studies (Kooijman et al., 2005) have confirmed that poor body posture can reduce respiratory support for the voice, which in turn reduces the openness of the vocal tract, resulting in problems with vocal production and vocal quality (Davis, n.d.; Schneider et al., 1997).

There is an overall understanding that postural alignment training is essential in optimising voice function (Arboleda & Frederick, 2008). As Mills (2004) pointed out, a concave chest restricts the

movement of the diaphragm and prevents the breathing muscles from working efficiently to project the voice. A proper postural alignment enables one to breathe correctly, control the flow of breath and produce a resonant tone (Ferris, 2018).

When applied to interpreting studies, proper postural alignment is vital for interpreters who work either in sitting or standing positions, taking notes and speaking for an extended period of time without rest (Freeman, 2010, p. 2). As such, there is a growing need for physical ergonomics training (including training in postural alignment) and ergonomic awareness for the high cognitive load of interpreting students to help them to “be able to cope with pressure, stress and competition” and to optimise “safety, physical health, comfort and performance” (van Egdom et al., 2020, pp. 364–365).

#### **2.2.5.1.2 Body movement**

Body movement, including gestures and hand movements, is a critical part of non-verbal communication in performance. According to Davidson’s (1993, 2001) work, body movement allows the performer to express their inner feelings and communicate with the audience, and assists in their comprehension. A study done by Nafisi (2013) confirmed that gesture and body movement are valuable tools for enhancing expressiveness in the performing arts and can improve tonal quality and musical phrasing, not simply artistic expression. Benge’s (1996) study confirmed that movement training can bring benefits, including better vocal dynamism, vocal quality and articulation.

Unlike sign-language interpreters, who are required to use amplified facial expressions and body movements as part of their grammar, spoken-language interpreters remain reasonably still on the stage, standing or sitting beside the speaker. Unfortunately, literature regarding the training of spoken interpreters’ body movement is scarce. However, interpreters’ gestures and body movements are an extension of their posture during interpretation and play vital parts in

communication. Using particular hand gestures and body movements, interpreters may attempt to reach and involve their audience. Thus, body movement is considered in the design of this study.

### **2.2.5.1.3 Physical tension balance**

Performers need to balance their physical and mental tension. A certain level of “healthy tension” (Shewell, 2009, p. 21) is beneficial as it will tone, make the relevant muscles ready for action and make the mind alert and be ready for communication. However, excessive habitual tension will hinder voice production (Linklater, 2006).

Similarly, in interpreting practice excessive tension can restrict interpreters’ breathing and suppress their natural voice. When interpreters carry tension inside their bodies but are not clearly aware of it, they naturally strain their necks and shoulders, and push their voices purposefully so as to be heard by their audience. As the physical tension becomes habitual, they even do this with a microphone in front of them. When they push vocally, their voice is not supported by or connected to their body. Thus, their voice may sound high pitched, rough and tight. This type of voice makes them, as well as their audience, tired and nervous, and weakens their capacity to communicate and perform confidently. In the worst scenario of habitual tension, interpreters can face “voice loss and even formation of nodules” (Gudgon, 2009, pp. 1–2). The first way to balance this tension is to arouse awareness of it. As Houseman (2015, p. 16) pointed out:

Many of our habits are unconscious. We are often not aware of excessive tension building up; we are just aware of the result – fatigue or pain. At other times we are aware of the tension but cannot work out how to release it because we are not clear enough as to how exactly it is building up.

Releasing excessive tension also requires kinaesthetic bodily awareness and muscular control (Gates, 2011). Voice work such as vocal warm-up activities can reduce excessive tension and constriction in the breathing and the larynx. Posture and alignment correction can help release and isolate muscles that are not involved in a specific physical activity so as not to interfere with the muscles that do perform the activity. As Brown (1996) stated, when excessive tension is released,



the body and mind are more likely to be in a balanced state and the voice will operate in a natural way rather than being forced.

#### **2.2.5.1.4 Breathing technique**

People without spoken-voice training sometimes speak through chest breathing (also known as shallow or clavicular breathing) because chest breathing is the most efficient way to quickly replenish the body's oxygen supply (Gates, 2011). Chest breathing while speaking can also be caused by tension or incorrect breathing habits (Department of Health & Human Services, 2015). Chest breathing does not provide vocal support for speakers, who need extra energy and stamina to sustain strenuous work, and creates physical tension, anxiety and reduced focus (Hahn et al., 1957). When someone produces sound through chest breathing, their shoulders and chest are lifted and their neck and shoulder muscles are tense, which does not sustain their speaking voice.

Chest breathing happens more frequently when interpreters face challenges. Some gasp for breath in the middle of a sentence or develop a habit of gasping while talking. As a result, natural breathing for speaking breaks down and the natural rhythm and flow of the breath are lost (Gudgon, 2009). Hence, it is beneficial to interpreters' careers if they practise their breathing skills until these become automatic and reliable (Gates, 2011). Diaphragmatic breathing is well-known as the proper method of breathing for the speaking voice. The diaphragm is the primary muscle used in the process of breathing. Flerov and Jacobs (2016, p. 27) described the diaphragm as a "dome-shaped muscular membrane [that] is located below the rib cage that separates the abdomen from the chest". Interpreters who use diaphragmatic breathing can improve their resonance and stamina to talk for longer periods of time, allowing them to take more breaths and increase their volume without straining their voice (Flerov & Jacobs, 2016).

#### **2.2.5.1.5 Release of speech muscles**

According to Weiler (2018), when people speak they engage nearly 100 speech muscles to shape the breath into the fluent sequences of sounds that form words and sentences. This reveals how the

complex articulatory movements are coordinated in the brain. These movable structures of the speech muscles include the face, jaw, lips, tongue, soft palate, pharynx and larynx, and the vocal folds form a vocal tract or “vocal channel” (Shewell, 2009, p. 144). The conscious use of speech muscles may change the vocal quality. On the contrary, inappropriate muscle tension in the jaw, tongue root and pharynx can cause “the soft palate to be pulled down” (Chapman, 2006. p. 123), which affects articulation. For example, according to Shewell (2009) vocal warm-ups including facial massage and stretching can improve the tone of the voice and change the long-term facial structure, which is highly relevant for voice muscle work. Also Shewell (2009) stated that the lips are part of the face and their action has a powerful effect on spoken-voice quality:

The changes in lip spreading and rounding have a strong effect on the overall vocal sound ... Lively movement of the lips can contribute to the avoidance of throat constriction, by focusing the energy to the front of the vocal tract, creating a forward resonant quality (p. 148).

The above sheds light on how interpreters can work on their speech muscles to improve their spoken-voice quality. For example, it may be the case that interpreters focus too much on delivering the interpretation and tend to speak very fast, which can lead to a situation in which they are unaware that they are speaking with closed or clenched jaws. The “jaw-dropping” exercises advocated by Shewell (2009, p. 150) involve relaxing muscles to be released with gravity so that the lower jaw drops into a slightly open position, at the same time slightly widening the pharynx. Boone and McFarlane (1993) found that “performing the yawn-sigh demonstrated retracted elevation of the tongue, a lower position of the larynx, and a widened pharynx”. However, Shewell (2009) suggested that trainees should focus on opening the back of throat and flattening the tongue when performing yawning-sigh exercise.

#### **2.2.5.1.6 Vocal care**

The vocal apparatus is part of the physical body and is subject to and reflective of the same strains as the rest (Gates, 2011). A cold, fatigue and/or stress can harm vocal quality. In addition, the condition of the work environment, for example, background noise, dry or overheated conditions, can also affect vocal quality. Like any other professional voice user, interpreters are likely to experience vocal difficulties sometime during their careers. For example, interpreters sometimes travel to another city or overseas for work. The effects of travelling in an air-compressed, dehydrating aircraft can cause severe vocal discomfort. Interpreters may often work in noisy environments, such as in a crowded room where a microphone is not available or not working properly. Straining the voice to be heard is a common cause of vocal misuse.

Some voice trainers, for example Gates (2011) and Barton (2015), have suggested a few methods to prevent vocal damage for professional voice users. The priority is to prevent vocal problems by keeping physically healthy and keeping the voice in shape through proper diet, exercise and rest, and refraining from smoking, alcohol and harmful substances. It also means maintaining a sound vocal technique and avoiding vocal misuse.

Regarding voice volume augmentation, Gates (2011) recommended not to push the volume up by tightening the jaw and tensing the neck muscles, and instead use proper vocal support. One of the techniques for increasing volume naturally advocated by Gates (2011) is to breathe in before speaking, support the breath with the abdominal muscles and let the breath produce the volume; Gates advised not exhaling on the lines of speech and keeping the breath behind the tone.

Research has revealed that habitual use of vocal fry can lead to an increased risk of vocal fatigue and vocal-cord injury and damages (Kelly, 2015; Wolk et al., 2012). Vocal fry is a way of speaking in which the voice is low-pitched and has a characteristic rough or creaking sound (Stevenson & Lindberg, 2010). It happens when there is not enough breath being pushed through the vocal cords, which then cannot rub together. Thus it creates a creaky sound. Speakers who have a practice of

speaking with vocal fry are often viewed harshly by listeners (Klofstad et al., 2012). Their listeners frequently consider them incompetent, uneducated and untrustworthy (Anderson et al., 2014). Gates (2011) pointed out that one way of tackling vocal fry is to place the voice at an optimal pitch that will result in a comfortable and healthy voice for the speaker, as well as being pleasant to the listening ear. This is because when the voice is produced below the optimal pitch, it is harder to hear and causes the ends of sentences to trail off and become inarticulate.

### **2.2.5.2 Spoken-voice quality**

Spoken-voice quality is multidimensional and includes many aspects. Podesva (2007) defined spoken-voice quality as “the extra grammatical suprasegmental properties of speech resulting from the configuration of the vocal apparatus” (p. 478). Spoken-voice quality is influenced by several factors that cause one’s voice quality to be different from that of anyone else and different from situation to situation. These factors include one’s emotional state when speaking, the structure of one’s resonators, general health, long-term habits of tension and relaxation in the body, and habitual emotional responses (Hahn et al., 1957). Therefore, voice quality can create an impression of the speaker’s persona. The following sections discuss some attributes that can affect the vocal quality of interpreters during the interpreting delivery stage. These attributes are voice projection, vocal resonance and timbre, pace and pauses, articulation, pronunciation and enunciation, and modulation of the pitch range.

#### **2.2.5.2.1 Voice projection**

Voice projection is a technique that involves clear diction, sound placement in the frontal plane and the desire to be heard. The amplitude (physical strength) of vocal-fold vibrations determines voice projection (Lumen Learning, n.d.). On the other hand, volume or loudness is what the audience actually perceives (Glenn et al., 1998). According to Rosenthal (2008), voice projection is more than speaking loudly and is one of the most effective vocal presentation skills. Correctly projecting voice can help speakers keep their audience’s attention and increase speakers’ confidence.

Flerov and Jacobs (2016) provided several observations and advice regarding the training of interpreters' voice projection. Some of their perspectives are considered in the design of this research project. For example, Flerov and Jacobs (2016) rightly pointed out that the primary reason for the soft voice presentation of some interpreters is "insecurity or natural timidity or simply a lack of professional voice training, or a combination of the above" (p. 38). Flerov and Jacobs (2016) believed that breathing and voice projection are closely related and argued that shallow non-diaphragmatic breathing results in a soft voice. Therefore, they advocated that those interpreters should unlearn poor breathing habits and breathe properly using the diaphragm. However, Flerov and Jacobs's (2016) argument that volume can be culturally specific and that people from certain cultures might be habituated to speak softly in public needs to be further verified. Although culture can be a factor in habitual low volume, it is arguable that the lack of proper voice training in these cultures may be a more prominent reason.

#### **2.2.5.2.2 Vocal resonance and timbre**

Vocal resonance is the intensity of the tone people can hear caused by the reverberation of sound waves from the vibrating vocal cords. Resonance is an essential quality of the human voice that forms the impression of someone's voice being pleasant to listen to (Flerov & Jacobs, 2016). The structure of the resonators influences the quality of resonance, for example, the body size and mass (Shewell, 2009). However, people can learn to develop greater resonance, enhancing the voice qualities they were born with (Gates, 2011). Flerov and Jacobs (2016) recommended vocal warm-up and relaxation exercises for interpreters because quality resonance is not produced from a tense body. Shewell (2009) stated that the quality of a voice's resonance is essential to both voice practitioners and ordinary listeners. Shewell elaborated on her point by giving examples: if a speaker is said to have "a very resonant voice", this means that their voice has natural power and is easy to listen to. If a speaker is said to have a "warm resonance", this means the listener feels that the voice has "a pleasant quality of breadth and depth" (p. 175).

Vocal timbre, also referred to as sound waves, is a unique sound that gives colour and characteristics to one's voice (Reina, 2018). Resonance exercises such as humming exercises are recommended to increase the width of the sound waves (Flerov & Jacobs, 2016, p. 65) and maintain a balance between head resonance and chest resonance (Gates, 2011).

#### **2.2.5.2.3 Pace and pauses**

Speech delivery pace (or rate) refers to how quickly or slowly a person speaks. A pause is a period of silence during speech. Pause and speech rate are interlinked prosodic features (Cecot, 2001).

Speech rates depend on pause frequency and duration and, in turn, affect pause frequency. SL speakers' speech delivery pace and pauses are two factors that affect interpreters' TL performance (Lec & Khaleel, 2009). A varying speaking delivery pace is a powerful tool for SL speakers to keep the speech vibrant, engage the audience and add verbal colour to the speech (Evans, 2016).

However, studies have found that speakers' rapid speaking rates can pose a severe challenge to interpreters' output by affecting their ability to juggle other interpreting skills (Gile, 2009; Lec & Khaleel, 2009).

Pauses also play a significant role as one of the non-linguistic aspects of verbal production in TL text delivery (Viezzi, 1996). As Viezzi (1996) stated, "The interpreter, as a public speaker, has to use pauses correctly and to adopt a proper speech rhythm" (p. 96). Appropriately placed pauses can serve as verbal punctuation marks that give the audience the same clues as punctuation marks in a written text (Dlugan, 2012). Pauses are also beneficial for allowing the audience to catch up with the speech and the points that a speaker has addressed (Evans, 2016).

Pauses can be categorised into functional and non-functional pauses, as well as grammatical and non-grammatical pauses (Shlesinger, 1994). Functional pauses are desirable because they divide discourse into tone groups and organise it into information units (El-Menoufy, 1988; Halliday, 1985; Shlesinger, 1994). As such, interpreters are suggested to deliberately add a pause (when it is appropriate) between words or syntagma because this will allow them to listen to the ongoing

incoming original speech and to maintain rhythm in their interpretation, allowing each word of the TL to be clearly heard and taken in (Flerov & Jacobs, 2016). Non-functional pauses include long silent pauses, filler words such as “um” and “ah” and the “filled pauses” that are caused by hesitation, interruptions, repeats, restructuring and false starts, which cause disfluency (Tissi, 1999). According to Shlesinger (1994), non-functional pauses “tend to lower the congruence between chunking and syntax, since the ensuing junctures are nongrammatical ... interpreters are prone to introduce a disproportionate number of pauses in ‘unnatural’ positions, which are liable to impede understanding” (p. 229). Research shows that filler words are more common among conscientious people who are afraid to make long silent pauses and can sometimes be caused by tight throat and vocal folds (Flerov & Jacobs, 2016).

Non-grammatical pauses are breaks in a speech that occur at places that do not coincide with grammatical constituent boundaries, for example, in the middle of a phrase, between repeated words, in the middle of a verbal compound and where the structure of a sentence is disrupted by a reconsideration or a false start (Lec & Khaleel, 2009). Non-grammatical periods of silence are hesitation pauses and thus are inadequate use of pauses that can cause problems and impair speech fluency for inexperienced and nervous speakers and interpreters. Research (Yin, 2011) showed that word fillers and long pause problems are prevalent among CI beginner learners and that students used more pauses and filler words in tasks working into their B language than into their A language.

#### **2.2.5.2.4 Articulation, pronunciation, accent, enunciation**

Articulation, pronunciation, accent and enunciation are closely related terms regarding the production of speech sounds. They combine to provide speaking clarity. It is important to understand the distinctions and similarities between these terminologies. Hahn et al. (1957) defined articulation as “a process of forming clear, recognizable sounds by the proper manipulation of the articulators, including the tongue, lips, lower jaw, and soft palate” (p. 125). Hahn et al. (1957) explained that good articulation is when the sounds are distinct, accurately pronounced and

harmoniously blended into one another, while poor articulation is either when the sound lacks clarity and accuracy or when it is “overly precise and pedantic” (p. 125).

The way we make speaking sounds is referred to as pronunciation. Hahn et al. (1957) defined pronunciation as good patterns of sounds and stress by referencing carefully edited dictionaries and formal speeches. The accent is the pronunciation particular to a speaker or group of speakers. Everyone has an accent and the so-called standard pronunciation in one place will be regarded as an accent by an audience living in another place. Interpreters’ accents combine pronunciation, word choice, interference of their native language and rhythm and intonation, rather than pronunciation alone (Flerov & Jacobs, 2016).

Accent reduction, also known as accent neutralisation, is a systematic approach to training for speech clarity and fine-tuning a specific accent for non-native speakers. Flerov and Jacobs (2016) proposed that accent neutralisation or reduction is a requirement for professional interpreters, especially those who work into their B language, for interpreters’ credibility. As Flerov and Jacobs (2016) observed, a native-speaking audience will sometimes perceive the interpretation as less credible if there is a negative impact of a foreign accent on their comprehension of the original speech. On the other hand, Flerov and Jacobs (2016) maintained that while accent reduction is essential, there is no need for an accent to go away entirely because the accent is what, to some extent, gives interpreters their personality in the foreign language.

Pronunciation and enunciation are different speaking attributes. Pronunciation is the act of saying a word correctly, while enunciation is how words are spoken clearly, concisely and in a way that is easy to understand. Houseman (2015) recommended vocal exercises that exaggerate reading aloud of the two components of the sound, vowels and consonants, to be practised to improve trainees’ enunciation.



### **2.2.5.2.5 Modulation of pitch range**

Pitch in speech is defined as “the relative highness or lowness of a tone as perceived by the ear, which depends on the number of vibrations per second produced by the vocal cords”

(Encyclopaedia Britannica, 1998). Gates (2011) noted, “Optimal pitch is the pitch or frequency at which the voice resonates with the greatest ease and clarity” (p. 43). Reiman (2019) defined the pitch of voice as the frequency of vibration of the vocal cords. The frequency of vocal vibration, which is one’s natural average pitch level, is determined by the thickness and length of the vocal cords and the tension of the muscles surrounding them. Apart from the size of the vocal cords, the pitch of one’s voice can also be affected by one’s emotions and moods (Reiman, 2019).

The pitch range is the span of pitch suitable for a speaker. Usually people use a pitch range that is comfortable for their natural conversation. Cosmides (1983) defined the pitch range as the difference between the minimum and maximum F0, which is how the vocal cords vibrate in voiced sounds. According to Patterson and Ladd (1999), the pitch range has significant communicative effects which reveal speakers’ characteristics (e.g., monotonous speakers are perceived as boring) and speakers’ transient emotions and mental states (e.g., raising their voice to signal anger or surprise). Thus, adjustment of the pitch range is necessary because it allows speakers to use pitch variation to achieve meaningful effects in speech (Esen, 2018) and enable their voice to accurately reflect the speaker’s thoughts and feelings (Houseman, 2015).

Pitch range modulation is also important for the audience. Bolinger (1986) pointed out that the audience is sensitive to change in pitch. Wrong placement of pitch can temporarily disrupt the listeners’ perceptions and expectations (Cutler, 1987). Shlesinger (1994) stated that a significant amount of “low-rise nonfinal pitch movement (i.e., the rising tone at the end of a sentence), conveying the tentative, provisional attitude of an afterthought in positions where a final (falling tone) one would be expected, is likely to impede comprehension” (p. 228).

As discussed above, the pitch is a valuable tool because it changes and reflects intellectual and emotional involvement (Rodgers, 2002). Christodoulides (2013) studied convergence phenomena between the speaker and a simultaneous interpreter in relation to prosodic features. The findings indicated that interpreters had a narrower pitch range than speakers and their pitch did not mirror the speakers' pitch. This result suggested that interpreters adopted a less expressive but "more continuous" (Christodoulides, 2013, p. 36) speaking style than the original speaker due to cognitive constraints.

Pitch range modulation can be realised by physically altering the tension of the vocal cords (Ladefoged, 2001, p. 226). This means that a person can lower their vocal pitch by loosening their vocal cords and increasing their voice pitch by stretching their vocal cords. To produce a variety of pitch ranges, physical and vocal exercises are recommended for toning the laryngeal muscles to bring about changes in the length and thickness of vocal folds and to practise the breath support muscles – the abdominal and pelvic floor muscles – making them work together effectively (Houseman, 2015). Melton (2012) proposed extending the pitch range by exploring high and low pitches safely and playfully, for example, through yoga-based postures and movement work in Pilates, to release trainees from habitual and inefficient breathing patterns. Vocal exercises such as "humming" and "roller coaster" are also commonly recommended (Houseman, 2015).

### **2.2.5.3 Spoken-voice dynamism**

Spoken-voice dynamism, also referred to as vocal variety, is how a speaker uses their voice to create interest, excitement and emotional involvement for an audience. People vary their spoken voice to create a social quality and an individual quality, depending on the setting and context (Sapir, 1927). A good speaker uses different notes to convey variety and meaning. A one-note speaker is tedious to an audience and promotes inattention and boredom.

The training of spoken-voice dynamism for interpreters involves variations in intonation, rhythm, tone, body language, hand gestures, facial expressions and other features to maintain the

engagement of the interpreting audience with the spoken text being rendered. Spoken-voice dynamism creates the first impression that an interpreter gives their audience about their overall interpreting delivery skills and professionalism. Good management of spoken-voice dynamism can help produce effective communication between the original speaker and the audience. This section introduces 5 spoken voice attributes deemed essential to interpreters. These attributes are intonation and tone patterns, rhythmic and fluent delivery, expressiveness, vocal confidence and audience connection.

### **2.2.5.3.1 Intonation and tone**

“Intonation” and “tone” are synonyms in their most general meanings. Both refer to the patterned variation in voice pitch that contrasts and organises words and larger utterances (Venditti, 2006). However, the two are distinguished in the matter of linguistic function. Intonation is referred to as the “melody of speech” (Flerov & Jacobs, 2016, p. 8) produced by the rise and fall of the voice when speaking, which has the role of the “creation of meaning” (Shlesinger, 1994, p. 225). The intonation patterns can indicate a speaker’s “evaluation of the importance and emotional significance of ideas, words and phrases” (Hahn et al., 1957, p. 77). As a pragmatic part of communication, the tone of voice refers to the emotion and attitude behind the speaker’s words (Lynch, 2021). As people’s emotions can significantly affect their tone, changing one’s tone can change the entire meaning.

Intonation and tone of voice are important prosodic features underestimated in interpreting training (Holub, 2010). Mossop (1987) observed that sometimes when the audience listens to an interpretation, they are constantly aware that they are listening to a speech through translation, not because the source speaker is visible but because the interpreter’s intonation and tone, and the stop-and-start pattern of their voice, remind the audience. As Hahn et al. (1957) pointed out, if the interpreter’s intonational patterns and tone cause their listeners to make wrong judgements about

the speaker's ideas, this is an issue as serious as misleading the audience through a wrong choice of words.

Shlesinger (1994) pointed out that this phenomenon is due mainly to the intonational effect of the interpreting on listeners but has rarely been discussed in interpreting training. Thus, Shlesinger (1994) advised interpreting students to maintain good intonational patterns primarily because this is “conducive to information retrieval” rather than for its direct communicative effect (p. 227). An experiment conducted by Shlesinger (1994) also showed that the audience gave a much higher score to interpreters delivering with standard and natural intonation than with interpretational intonation, which sounds slightly awkward, non-natural, monitored or not congruent with the overall tone or cadence of the source speaker. This demonstrated that intonation and vocal tone can affect the ultimate speeches conveyed to the audience and play more critical roles in interpreting assessment than is recognised in surveys of user expectations, as evidenced by other experimental research (Garzone, 2003, Collados Aís et al., 2007; Iglesias Fernández, 2007, 2013).

In language, tone refers to “the linguistic use of pitch in marking lexical items, and intonation refers to non-lexical use of pitch to indicate utterance level pragmatic distinctions and phrase boundaries” (Ota, 2016, p. 18). All languages are known to have intonation, but only about 60–70 per cent of the world's languages, including Mandarin, have tonal marking of lexical items (Yip, 2002). This poses a great challenge to interpreting students working between a tone-centred language such as Chinese and an intonation-centred language such as English. Schack (2000) stated:

Chinese is a tone language, while English is an intonation language. That is to say, if Chinese has both tone and intonation, then Chinese assigns tonal targets on a lexical as well as phrasal level, while English only assigns an intonation tune on a phrasal level. Thus, in Chinese, the dual usage of tone leads to a more complicated picture than is found in English, making it more difficult to separate lexical tone from intonation. Moreover, it is not clear how a native speaker of a tone language would deal with tone in English since English uses intonation but makes no specifications for lexical tone (p. 25).

A Chinese native speaker, when making an English speech, will often make use of an intonation system that more closely resembles that of their native language. As a result, they may sound monotonous to their listeners and may be perceived as less professional when speaking in a monotonous voice regardless of whether they correctly convey the content of the original speech (Holub, 2010). This negatively affects their overall performance. As such, training of intonation and tone pattern deserves greater attention in interpreters' spoken-voice training, especially for those working in English as their B language. These interpreters need to be trained in their intonation to reduce a monotonous voice and bring more interesting expressions and dynamism into their work of message transfer.

### **2.2.5.3.2 Rhythm and fluency**

In language, rhythm (also known as sound patterns) refers to how words and sounds flow in a person's voice or speech, marked by the stress, timing and quantity of syllables. Rhythm is essential to a listener's comprehension of a speech. As Fujii and Wan (2014, p. 2) explained, "in order to understand spoken language, listeners are required to perceive temporal organization of phonemes, syllables, words, and phrases from an ongoing speech stream". As one of the most important prosodic features, a smooth rhythm is an attribute of what is considered a pleasant-sounding voice. Flerov and Jacobs (2016) described rhythm as the "backbone of interpretation; everything else is connected to it. Good rhythm creates the impression of coherence, predictability, stability, comfort and safety" (p. 42).

English and Mandarin are very different in rhythm. According to Mackay (1985), the Mandarin language is syllable-timed with nearly equal weight and time in all syllables. Because the same sounds are pronounced with different tones in Mandarin, each Mandarin word is pronounced distinctively and individually. English, in contrast, is a rhythmical language which is stress-timed with alternation of stressed and unstressed syllables. This means that if an English learner can maintain the rhythm of the language, they are more likely to sound both natural and fluent. The

striking difference in rhythm between the Mandarin and English languages causes rhythm to be one of the most significant issues facing Mandarin native speakers who speak English as a second language (Chen et al., 1996).

In speech pathology, speech fluency means the smoothness or flow of delivery when speaking in which the sounds, syllables, words and phrases are connected well (Chambers, 1997; Koponen & Riegenbach, 2000). Fluency in second-language studies refers to effortlessness or automaticity in a second-language learner's speech, demonstrated in the flow, continuity and smoothness of the speech that they deliver (Segalowitz, 2010; Skehan, 2014; Tavakoli et al., 2016).

As essential prosodic elements in speech, rhythm and fluency are two distinctive but correlated aspects. A study (Salomaa, 2019) found that rhythm and fluency are intertwined and that the core components of fluency, such as pausing and speech rate, also contribute to speech rhythm. It is impossible to produce a rhythmic speech when speaking without speech fluency (Hirvonen, 1973). Studies (Adams, 1979; Paananen-Porkka, 2007) have indicated that L2 users are more likely to pause frequently at inappropriate places, which negatively affects their fluency, causes incorrect phrasing and disturbs the rhythmical pattern. Grant (2015) suggested some ways of improving the rhythm, for example, by extending the vowel sounds and sliding the words together so the rhythm does not appear choppy or staccato and by connecting speaking with the flow of gentle body movement. This rhythm-based training could also be beneficial to improving trainees' speech fluency (Fujii & Wan, 2014).

#### **2.2.5.3.3 Expressiveness**

Expressiveness refers to “the transmission of emotions and intentions in the message” (Marquezin et al., 2015, p. 161). Expressiveness involves revealing a speaker's emotions in their voice and speech through vocal and non-verbal communication. Table 2.2 shows the relationships between vocal, non-vocal, verbal and non-verbal aspects of communication (Hargie, 2011).

Table 2.2 Vocal and non-vocal elements of communication

	<b>Verbal Communication</b>	<b>Nonverbal Communication</b>
Vocal	Spoken words	Paralanguage (pitch, volume, speaking rate, etc.)
Nonvocal	Writing, sign language	Body language (gestures, facial expressions, eye contact, etc.)

Source: Adapted from Hargie (2011, p. 45).

According to Table 2.2, verbal communication consists of exchanging information using words. Non-verbal communication consists of all the messages other than words that are used in communication. Both features must be operating side by side as part of the same system and present themselves in a manner consistent with the subject to be spoken about (Neiva et al., 2016). Verbal and non-verbal communication also includes vocal and non-vocal features or behaviours. Several studies (Hargie, 2016; Joseph, 2013; Mehrabian, 1971) have confirmed that a large majority of interactions are not taken up by spoken words but by non-verbal communication and that people may put more weight on non-vocal features when determining a speaker's credibility because non-verbal communication is mostly an unconscious part of everyday behaviour (Guerrero & Floyd, 2006; Hargie, 2016).

As Table 2.1 shows, moving from lexical content, a vocal element of non-verbal communication is paralanguage, including speaking rate, pauses, tone of voice, intonation, volume and pitch. Hence, vocal expressiveness is a comprehensive skill that a speaker employs in a vocal variety or paralanguage to help their audience comprehend their message. In contrast to vocal expressiveness, a monotonous output signals to the audience that the speaker does not care much about their message or the response of their audience. Anderson (1977) pointed out that the cause of vocal monotony is when proper vocal variety in speech is lacking. Thus Anderson (1977) proposed that students should pay attention to vocal expressiveness drills involving wide variations on pitch range, rhythm, loudness, vocal tone and emphasis.

#### **2.2.5.3.4 Vocal confidence**

Confidence can be defined as “an acquired attribute that provides individuals with the ability to maintain a positive and realistic perception of self and abilities” (Evans et al., 2010, p. 335).

Research has demonstrated that a high confidence level positively impacts on one’s performance (Feltz, 2007). Beyond physical appearance, voice is the most persuasive way to get the message across in many personal and professional situations. While a speaker carefully monitors their words to achieve the desired effect, their voice may reveal their emotion and confidence level despite their best efforts. For example, a submissive voice reveals shyness (Ilardo, 1981), preventing effective communication. A shy voice may be caused by a few reasons such as cultural background, fear (e.g., stage fright), emotional suppression, excessive self-consciousness or feelings of inadequacy and inferiority due to the suppression of a voice or speech defect (Anderson, 1977).

A speaker’s confidence can be transmitted through various spoken-voice techniques, including tone of voice, speaking clarity, volume, delivery pace, pitch modulation, pauses and emphasis to create an effect, intonation and body movements such as posture and gesture. Hargie (2016) recommended using vocal tonality, one of the paralanguage features that creates a strong presence and a confident voice that commands attention. Notas (2012) pointed out that there are two types of vocal tonality: dominant and submissive. Dominant voices express assertiveness and security, whereas submissive voices express uncertainty and passiveness. Dominant voices stay the same or lower in pitch when finishing a sentence or phrase, giving the audience the impression that the speaker is speaking without hesitation. On the other hand, submissive voices rise in pitch.

When discussing interpreters’ confidence, Flerov and Jacobs (2016) also asserted that one of the objectives of interpreter voice training should be to change the intonational pattern and use falling tones in a sentence instead of rising tones. They believed that the pattern of using a falling tone at the end of a syntagm must be learnt to the degree when it becomes automatic for interpreters. Flerov and Jacobs (2016) explained that sometimes interpreters cannot finish a sentence for various



reasons such as information density or cognitive constraint. If interpreters' intonation goes up at the end of syntagma, listeners will expect the sentence to be continued. Also, a rising intonation sounds awkward, unsure and less credible. On the other hand, interpreters' voices will sound more confident, professional and collected with a falling intonation. However, Flerov and Jacobs (2016) warned that overusing a low falling intonation may create a monotone and suggested that interpreters use a combination of high falls and low falls.

Other recommendations by Hargie (2016) also shed light on the training of interpreters' vocal confidence, for example, good use of hand and arm movements to maintain listeners' attention, indicating enthusiasm and assertiveness; projecting the voice from the lower chest rather than the head, which will aid in building a dominant and authoritative voice; controlling the rate of speech; and enunciating words clearly and fully using pauses effectively.

#### **2.2.5.3.5 Audience connection**

Interpretation is a bridge between the meaning of the speech and the interest of the audience.

Interpreters help a speaker render a speech into another language because the speaker has a message to share or, more accurately, because the audience wants to understand a message that is being given in another language (Bacher et al., 2007). Through interpreters' spoken-voice and non-vocal communication skills, the audience is given an opportunity to meaningfully connect with the speaker.

“Audience analysis” (Miller, 2017, p. 52) can be crucial for interpreters to connect with their audiences through vocal modulation that is appropriate to the speech being interpreted. Bacher et al. (2007) suggested that interpreters should possess general knowledge of their target audience including their interests, expectations and current attitudes towards the topic of the meeting:

It is the role of the interpreter to ensure that those connections are built on the interests of the audience. And it is the role of the audience to determine which bridges will be crossed ... Truly meaningful interpretation relates what is being interpreted to the hearts and minds of the audience and

answers the question ‘Why should I care?’ ... connections involve moments of intellectual and emotional revelation, perception, insight, or discovery (p. 5).

By understanding the audience, interpreters can respond to and interact with the audience by modulating their volume, pitch, tone, eye contact and other vocal and non-verbal behaviour.

In conclusion, the 3 spoken-voice categories and 15 attributes addressed in this subsection are closely related to the vocal work of interpreters and will serve as the foundation for constructing the methodology for this study.

### **2.3 Metacognition, self-regulation and their application to interpreters’ vocal performance**

This section examines metacognitive strategies and self-regulation, and how they are potentially related to spoken-voice intervention for interpreting students. The section starts by discussing definitions of each of the key terms, relevant theories and the interconnections between these terms. Then it discusses the importance of integrating metacognition theory into interpreting teaching practice and curricular design to utilise metacognitive knowledge and strategies in improving the education and training of future interpreters. Finally, it gives a few examples of metacognitive strategies and self-regulation that could be utilised in training interpreting students’ vocal delivery.

#### **2.3.1 Metacognition**

The term “metacognition” was first formally used by Flavell, who defined it as “knowledge and cognition about cognitive phenomena” and “thinking about own thinking” (1979, p. 906). Since Flavell, the definition of metacognition has been further developed and refined. Livingston (2003) stated that metacognition is “higher-order thinking which involves active control over the cognitive processes engaged in learning” (p. 2). Livingston (2003) explained that anyone who plans a learning task or monitors and evaluates progress towards completing a task is engaged in a metacognitive activity, either intentionally or unintentionally.

Metacognitive knowledge is one's awareness and beliefs about the factors that affect one's cognitive activities (Wenden, 1999). Learners who are metacognitively aware "know what to do when they do not know what to do" and find motivation and strategies to figure out and achieve what they need to do (Anderson, 2002, p. 1). Learners' metacognitive awareness helps identify the problem encountered, analyse the nature of the problem and determine what knowledge and skills are needed to approach the task (Wenden, 1999). Research (Garner & Alexander, 1989; Henter & Indreica, 2014; Schraw, 1994) has shown that learners who have metacognitive awareness have better problem-solving strategies and a higher level of performance than unaware learners. For example, Schraw (1994) explained that metacognitive awareness allows individuals to plan, arrange and self-monitor their learning in a way that directly improves performance and that "differences in strategy use and performance were related to differences in metacognitive awareness rather than differences in intellectual aptitude" (p. 461).

Metacognition also covers the ability to self-monitor the learning process. Metacognitive regulation refers to how people control their thinking to facilitate learning (Stanton et al., 2015). From a cognitive psychology point of view, learning is a shift from conscious attention to automatic processing (Pammu et al., 2011). Metacognitive monitoring (also known as self-monitoring) is an important learning strategy for practising various tasks leading to automaticity when performing a skill (Henter & Indreica, 2014).

Because metacognition is vital to successful learning, over the past 4 decades or so researchers have studied across different disciplines (Hacker, 1998) to discover how to enable successful learners to regulate their cognitive processes and learning activities actively. Several studies have been done on understanding the role of metacognition in first- and second-language teaching and learning, and have found that metacognitive knowledge plays an essential role in many cognitive activities related to language use, such as the skills of spoken communication, spoken persuasion, listening

comprehension, memory, attention, problem-solving and various types of self-instruction (Anderson, 2002; Flavell, 1979; Wenden, 1999).

### **2.3.2 Self-regulation**

Self-regulation is the capacity to exert control over one's cognitive process and behaviour during the learning and performance processes (Flavell, 1987; Schraw, 1998). Self-regulation plays a significant role in learning and mastering new skills and requires conscious effort. According to Universal Design for Learning (UDL) guidelines (UDL, 2018), while it is critical to construct an extrinsic learning environment to promote motivation and engagement, it is also critical to enhance learners' intrinsic capacities for self-regulation of their cognition, emotions and motivations.

Guideline 9 in the UDL states:

The ability to self-regulate – to strategically modulate one's emotional reactions or states in order to be more effective at coping and engaging with the environment – is a critical aspect of human development. While many individuals develop self-regulatory skills on their own, either by trial and error or by observing successful adults, many others have significant difficulties in developing these skills. Unfortunately, some classrooms do not address these skills explicitly, leaving them as part of the “implicit” curriculum that is often inaccessible or invisible to many. Those teachers and settings that address self-regulation explicitly will be most successful in applying the UDL principles through modeling and prompting in a variety of methods. As in other kinds of learning, individual differences are more likely than uniformity. A successful approach requires providing sufficient alternatives to support learners with very different aptitudes and prior experience to effectively manage their own engagement and affect (UDL, 2018).

There are 3 checkpoints suggested by the UDL guidelines (UDL, 2018), namely, to promote expectations and beliefs that optimise motivation, to facilitate personal coping skills and strategies, and to develop self-assessment and reflection.

Self-regulated learning was developed based on self-regulation. Self-regulated learning is defined as “self-generated thoughts, feelings and actions for attaining one’s learning goals” (Zimmerman & Moylan, 2009, p. 363). With heightened self-regulation capacity, learners take proactive steps to evaluate their learning, identify their strengths and limitations, and effectively employ appropriate strategies. Self-regulated learning involves developing key processes such as “goal setting, time management, learning strategies, self-evaluation, self-attributions, seeking help or information, and important self-motivational beliefs, such as self-efficacy and intrinsic task interest” (Zimmerman, 2002, p. 64).

### **2.3.3 Intersections of metacognition with self-regulated learning and motivation**

Metacognitive knowledge and self-regulation are two essential components useful in improving students’ learning behaviours. As Papaleontiou-Louca (2008) pointed out, metacognition is learning and development on the one hand, and the process of regulating learning and development on the other hand. Metacognition and self-regulation are not identical (Dinsmore et al., 2008; Lajoie, 2008; Zimmerman, 1990). Metacognition emphasises that an individual’s self-awareness triggers subsequent decisions or evaluations. Self-regulation focuses on the interaction between the environment and the individual’s awareness and regulatory response (Dinsmore et al., 2008) in academic learning and professional learning contexts (Lajoie & Azevedo, 2006; Pintrich, 2004).

Despite the difference, metacognition and self-regulation are closely related. Metacognition is a significant part of self-regulated learning (Schraw et al., 2006; Zimmerman, 1986) that also has cognitive and motivational components. Shepard (2000) suggested that learning is an active process of mental construction and sense-making, or an interaction between the mind and the environment, which means an effective learning intervention will not solely focus on one of these skills but look at all aspects. The development of self-regulated learning requires awareness and conscious thinking about the learning process itself and what it involves (Cañada & Arumí, 2012). An accurate self-assessment of what is known and not known is the key to effective self-regulating

(Schoenfeld, 1987). As Hacker (1998) stated, “only when students know the state of their knowledge can they effectively self-direct learning to the unknown” (p. 13).

One of the most challenging issues confronting educational researchers is explaining how students learn in self-regulated contexts such as studying or practising independently (Zimmerman & Moylan, 2009). Research evidence shows that students’ ability to maintain the effects of strategy training and transfer the technique to non-experimental contexts needs to be enhanced (Pressley & McCormick, 1995). In this goal, concentration, self-awareness, retention and motivation have been identified as crucial characteristics of self-regulated learners (Zimmerman & Moylan, 2009). A concerted effort has been made to incorporate motivational factors and metacognitive processes in models of self-regulated learning (Schunk & Zimmerman, 2007).

Motivation is a driving force that energises and guides one’s behaviour towards a specific goal (Eggen & Kauchak, 1994). Motivation in learning is defined by Gopalan et al. (2017) as a persuasive sensation or desire that encourages students to persist in learning no matter how difficult the task is. Motivation is one of the fundamental factors in academic achievement apart from cognition, intelligence and preference (Kong, 2009). It is influenced by several factors, including belief in one’s effort in fulfilling a task and awareness of the value and attributes of the academic tasks assigned (Legault et al., 2006). Chang and Chang (2012) claimed that learning is a complicated psychological process and motivation is the only factor that stimulates learning effectively among factors such as individuals’ goals, beliefs, expectations and environments. The amount of motivation reflects students’ engagement and contribution in a learning environment.

There are many different categories of motivation. The broad types used commonly in education research include intrinsic motivation, extrinsic motivation and amotivation (Brief & Aldag, 1977; Deci & Ryan, 2013; Gopalan et al., 2017). Intrinsic learning motivation arises from the learner’s internal needs such as curiosity, the desire to learn and feelings of competence or growth (Eggen & Kauchak, 1994) instead of expecting external rewards or working under external pressure (Deci &

Ryan, 2016). On the other hand, extrinsic motivation exists when individuals are driven by external incentives, recognition or compulsion (Tohidi & Jabbari, 2012). Amotivation is a state in which both intrinsic and extrinsic motivation are absent (Ryan & Deci, 2000).

Several studies have concluded that intrinsic motivation and academic achievement have significant and positive correlations (Lepper et al., 2005; Pérez-López & Contero, 2013), indicating that intrinsic motivation plays a vital role in influencing students' academic achievement. Nevertheless, Ozcelik et al. (2013) and Li and Lynch (2016) have claimed that both intrinsic and extrinsic motivation are needed in the learning process. Huang (1992) suggested that adults tend to select education activities based on intrinsic psychological factors and the external environment.

According to Tohidi and Jabbari (2012), learning may start based on a mixture of extrinsic theories but transfer to intrinsic theory.

Much research has been done to understand the factors that affect learners' motivation and their relationships. Studies have found that a person's motivation can be affected by many factors, including cognition, behaviour, environmental and personal factors that work together to form their motivation (Gomel, 2017; Wood & Bandura, 1989). Kong (2009) claimed that personal factors play a vital role in individual learning, including a learner's arousal, needs, beliefs and goals. Research found that using appropriate metacognitive learning strategies and self-regulation helps increase motivation and beliefs (Schunk, 1995). Self-regulation is the ability to overcome difficulties, refocus and avoid being distracted by appealing stimuli (Baumeister et al., 1994). Self-regulation plays a large role in converting motivation into action. This suggests that including metacognitive strategies and motivational beliefs may help in understanding how students self-regulate and self-sustain their learning (Zimmerman & Moylan, 2009).

#### **2.3.4 Metacognition and self-regulation in interpreting pedagogy and spoken-voice training**

Research shows that metacognitive strategies are associated with successful learning (Borkowski et al., 1987) and that these strategies can be taught (Halpern, 1996). Several studies have confirmed

that successful learners have various metacognitive strategies to select from and can transfer these strategies to new settings (Pressley et al., 1987). While metacognition theory and self-regulated learning have been applied to learning foreign languages for some time, their application to the activity of interpreting is much more recent. Several interpreting educators have argued for the importance of integrating metacognitive theory and self-regulation into interpreting teaching practice and curricular design. Luis and Morell (2011) proposed that metacognition can be applied “for all the variants and professional modalities of this specialized communicative activity, be it consecutive or simultaneous interpreting” (p. 116). They called for the early introduction of metacognitive knowledge and strategies to interpreting students under the guidance of instructors to let students personally experience and learn to use self-monitoring and self-reflecting strategies in order to cope with the complex interpreting process of every new and specific interpreting setting, rather than imposing a one-size-fits-all teaching method and procedure. Cañada and Arumí (2012) conducted a study on the self-regulating process in students’ CI from German into Spanish using a “metacognitive guide” to understand the importance of self-regulating the interpreting teaching and learning processes. Hild (2014) compared self-regulatory competence between skilled interpreters and novice interpreters, and called for setting self-regulation as one objective of interpreter training and interpreting professional development. These studies may align with the focus of this study and ultimately the further research of applying metacognition and self-regulation to the teaching of interpretation and improving professional interpreters’ performance.

Acquiring metacognitive awareness is the first step for students in beginning a journey that may be known by the term “metathinking” (Flavell, 1976, p. 107). Metacognitive awareness can lead individuals to engage in or abandon a particular cognitive activity based on their interests, abilities and goals. In other words, what learners know and believe influences their behaviour, strategies and attitudes. As Wenden (1999) pointed out:

Metacognitive knowledge may be acquired unconsciously through observation and imitation or consciously as learners listen to teachers or



peers providing them with advice about how to learn. As they gain in cognitive maturity, learners may reflect on their learning process and revise earlier assumptions or develop new ones. They are capable of bringing this knowledge to consciousness and talking about it (p. 436).

When applied to interpreting training, some interpreters prefer to work from their A language into their B language as a result of an awareness that anticipation skills (a critical cognitive activity in SI) are more effortless when working from their native language than from an acquired language due to their better knowledge of “cultural patterns and transitional probabilities” (Gile, 2009, p. 253).

Other interpreters refuse to work into B in SI mode, but are willing to do so in CI as they believe that they have more processing capacity available for speech production in the latter (Gile, 2009, p. 177). These interpreters, either intentionally or unintentionally, are engaged in metacognitive activities. They are doing what Wenden (1999) described as “deliberately calling upon their metacognitive knowledge when the learning task requires conscious thinking and accuracy or if it is new and when previous learning has not been correct or complete” (p. 437).

Regulation of one’s cognition is the second step. Cognition is a mental process that includes thinking, knowing, remembering, judging and problem-solving. These mental processes are essential to interpreting tasks and are discussed extensively in Gile’s (2009) effort model. Gile’s “competition-between-efforts principle” and “cognitive coping strategies” (2009, p. 189) are well recognised as being relevant to the current view of cognitive psychology and have been applied extensively in the interpreting classroom. Different from cognition, metacognition is the capacity to know, reflect upon and regulate our cognition processes as individuals during the development of various learning activities (Anderson, 1995). It is essential in order to help learners become aware of their own learning style, learning process and the demands of learning tasks so as to know how to choose appropriate learning strategies in different contexts (Schraw, 1998; Vandergrift & Goh, 2006).

Instructors play a vital role in modelling both cognitive skills (i.e., how to perform a task) and metacognitive skills (i.e., thinking about and monitoring the thought processes of performing a task) for their students to make them understand the distinction between cognition and metacognition, and hence become self-regulated (Schraw, 1998). Interpreting instructors may endeavour to identify problems in the cognitive process and offer diagnoses and advice relating to the process of interpreting, rather than commenting on performance as being “right” or “wrong” (Gile, 2009, p. 14). This approach can further expand to encourage “higher-level processes responsible for self-regulated activities” (Hild, 2014, p. 129).

By using a set of interactive teaching activities and instructional practices, instructors help learners develop metacognitive knowledge and skills rapidly, achieving transformation from awareness to self-regulation (which is not an automatic process) (Cañada & Arumí, 2012). When students are given regular opportunities to reflect on their successes and failures, they will be able to develop metacognitive knowledge and regulatory skills on their own and have “a greater sense of self-efficacy such as effort, strategy and perseverance when faced with challenging circumstances” (Schraw, 1998, p. 122). Apart from the role of instructors, learners, as the agents of their own learning, play an essential part in the process of self-regulation. They are required to devote effort to modifying or controlling thoughts and emotions, and changing behaviours that were not previously under conscious regulation, which is discussed in relation to interpreters’ spoken-voice training below.

When applied to spoken-voice training, learners are required to make a conscious effort to use, in a different way, their vocal organs and muscles to achieve changes in vocal quality or, in some cases, to correct inefficient vocalisation behaviours that may be produced with little awareness (Colton & Casper, 1996). Based on experimental evidence, the self-regulatory capacity is a limited set of resources that can be depleted (Schmeichel, 2007) in two concurrent or consecutive tasks that demand self-regulatory effort. Self-regulation on one task may result in inadequate self-regulation

on other tasks. This is well related to interpreters' work. Interpreters are often forced to stretch their cognition and working memory processing capacity to cope with high-level cognitive activities concurrently or consecutively with only seconds apart. As Gile (2009) pointed out, effortless complete speech is a myth as "speech production may be making quite some very substantial demands on our linguistic system" (p. 163). Speech production in interpreting is a "non-automatic operation" (Gile, 2009, p. 165) and maintaining quality spoken-voice production under interpreting conditions may increase the cognitive load and require greater self-regulation capacity.

These multiple demands on cognitive effort compete with limited self-regulatory capacity and cognitive processing capacity, causing "self-regulatory fatigue" or "mental fatigue" (Schmeichel & Baumeister, 2004, pp. 84–98), which leads to "quality deterioration" (Gile, 2009, p. 171) in the interpreting performance, including mind-wandering, decreased ability to control emotion and undesirable responses (Vinney & Turkstra, 2013) and poor voice and speech production. If these difficulties and quality deterioration are not detected, monitored and controlled by the interpreter at the outset, they may lead to non-optimal consequences for the interpreting assignment.

Metacognitive awareness and self-regulatory capacity can be strengthened by making a conscious effort with repeated physical practice until reaching an automatic state. Physiological voice training typically involves self-directed behaviour change using specific exercises. These exercises focus on changing motor behaviours related to respiration, muscular effort and resonance via specific exercises and motor adjustments (Vinney & Turkstra, 2013). For example, exercises like humming and chanting as used in Lessac–Madsen resonant voice therapy were shown to enhance oral resonance, decrease muscular effort and reduce tension in the interpreting production phase (KV, 2008).

Apart from physical voice-training activities, some metacognitive strategies can be borrowed to develop interpreting students' metacognitive awareness and self-regulation, for example, repeated practice and training based on a problem-based learning (PBL) approach (Dunlap, 2005; Paris &

Paris, 2001), metacognitive prompts such as metacognition and reflective inquiry (MRI) (Anderson et al., 2009); reflective journal writing (Henter & Indreica, 2014); trainer- or student-led metacognitive discussions based on the learner-centred principles (McCombs, 2004); prosodic feature awareness activity (Yenkimaleki & van Heuven, 2018) and other metacognition modelling by trainers (Zohar & Barzilai, 2013). These metacognitive and self-regulated learning strategies are detailed in the following paragraphs and tested in this research project.

PBL is identified as one of the educational methodologies (Barrows & Tamblyn, 1980) that helps develop effective self-regulation. PBL has gained popularity among curriculum developers and trainers because it is based on interdisciplinary learning and emphasises metacognitive skills and real-life perspectives (Putnam, 2001). In PBL, each practice and training activity focuses on resolving a specific or practical problem (Wood, 2003).

The learner-centred principles were developed by the American Psychological Association in 1993 and were expanded in 1997 (McDonough, 2012). The 14 learner-centred principles are defined as follows:

The perspective that couples a focus on individual learners – their heredity, experience, perspectives, backgrounds, talents, interests, capacities, and needs – with a focus on learning – the best available knowledge about learning and how it occurs and about teaching practices that are most effective in promoting the highest levels of motivation, learning, and achievement for all learners. This dual focus then forms and drives educational decision making. Learner centered is a reflection in practice of the Learner-Centered Psychological Principles – in the programs, practices, policies, and people that support learning for all (APA, 1997).

The learner-centred approach allows learners to make decisions about how and what they learn and how that learning is assessed, values each learner's unique perspectives, respects and accommodates individual differences in learners' backgrounds, interests, abilities and experiences, and treat learners as co-creators and partners in the teaching and learning process (McCombs, 2004).

The MRI approach (Anderson et al., 2009) is a potential tool to prompt students to become aware of their metacognition and its effectiveness (Donkersgoed, 2016). Several MRI studies have confirmed that metacognition is the active monitoring, conscious control and regulation of learning processes (Baird, 1986; Flavell, 1987; White, 1993). According to Anderson et al. (2009), the general methodological procedure of the MRI consists of 3 stages. In the first stage, the participants are provided with an initial metacognitive interview where they engage in metacognition and discuss their knowledge, self-awareness, monitoring, evaluation and control of their knowledge construction and cognitive processes. In MRI studies, semi-structured interview protocols are typically adopted for the participants to engage in. In the second stage, as Anderson et al. (2009) explained, the participants are given opportunities to reflect on their metacognition and learning processes through predesigned activities. During post-interviews at the third stage, the change in students' understanding of metacognition is investigated by comparing students' answers in pre- and post-interviews, and questioning them about their understanding.

A reflective journal is another metacognitive technique to raise learners' metacognitive awareness so as to improve their learning performance. The effectiveness of the reflective journal was recognised as early as the 1980s (Sanford, 1988; Voss, 1988). Reflective journal writing combines metacognitive processing with the learning process. Regular reflective journal writing allows learners to make new connections between what they have been learning and make sense of the importance of why they are learning it (Olson & Johnson, 2012). Reflective journal writing can be applied to monitor both metacognitive dimensions, including knowledge of cognition and regulation of cognition, as learners are encouraged to reflect on their strengths and weaknesses while performing an activity. It can also be used for other self-management or resource-management skills, such as time management, managing efficient learning environments and peer learning (Henter & Indreica, 2014; McCrindle & Christensen, 1995; Roelle et al., 2012).

Reflective journals have the potential to make significant contributions to learning as learners are encouraged to express and reflect upon their feelings, beliefs, knowledge and skills by recording their thoughts on their awareness involved in the cognitive effort, commenting on their learning decisions and exploring their ways of using learning strategies consciously and adapting the learning process to real situations (Henter & Indreica, 2014). Such a learning journal can offer students a stimulus to start thinking about their cognitive processes and improve their motivation. In some studies, instructors employed the practice of analysing students' reflective journals (with students' permission) to identify the extent to which metacognitive training can improve learning and to gain insight into learning new strategies for academic study which could consequently improve teaching (Olson & Johnson, 2012).

To conclude, this chapter has analysed the essential literature on interpreting training in relation to spoken-voice delivery, interpreting performance quality and interpreting performance evaluation criteria. Then, a review of the core concepts of spoken-voice pedagogy and their application to interpreting teaching has been undertaken. Finally, theories of metacognition and self-regulation and their application to interpretation and spoken voice training have been examined. The literature review served as the foundation for this research investigation.

## Chapter 3 Research Methodology

From the overview of the research literature provided in Chapter 2, it can be seen that a number of themes relating to interpreter performance and pedagogy are relevant to the topic of spoken-voice training for trainee interpreters. Based on the overview, this thesis study has formulated the following research questions, related to the 4 experimental groups, in large part based on gaps or under-researched areas of interpreter training:

1. Does any group of students record a significantly higher level of vocal performance at the post-intervention stage compared to the pre-intervention stage?
2. Do the students who receive the spoken-voice intervention attain a significantly higher level of vocal performance than the students who do not receive spoken-voice related intervention? If yes, which group attains the best result?
3. Are there correlations between the students' spoken-voice performance and their general interpreting performance? If so, to what degree does those students' spoken-voice performance inform their interpreting performance?
4. Where there is a significant improvement in the post-intervention assessment compared to the pre-intervention assessment, what are the spoken-voice attributes that record high levels of improvement?
5. Where there is a higher level of performance recorded, which factors are reported to contribute to the changes?

This study has adopted a mixed research method by integrating quantitative and qualitative data collection and analysis to address these research questions. The quantitative results have been triangulated with qualitative data for cross-validation and explanation of the general research design, and the methodology adopted for this research project is explained in detail in this chapter.

Quantitative data was collected and analysed through controlled experiments to obtain the statistical results. The students' pre-and post-intervention videos were scored by 3 external assessors based on

predesigned evaluation sheets in order to understand any within-group improvements and between-group differences. The results from the experimental studies are presented through descriptive statistics. Qualitative data including the assessors' comments, the students' reflective journals and the interview responses were collected concurrently with the quantitative data collection (during and after the intervention) in order to interpret the quantitative results and draw conclusions to every individual research question. In this way, findings from the quantitative data about students' improvement were explored further with the qualitative data to understand better how the personal experiences of individuals matched the quantitative instrument results (Creswell & Poth, 2018).

This chapter has 3 parts. Section 3.1 describes the data preparation stage, including setting up the intervention groups and the control group, recruiting the research participants and how the researcher organised the intervention activities and assessments. Section 3.2 discusses how the qualitative and quantitative data were collected in different stages. Section 3.3 presents details of the quantitative and qualitative data analysis methods. Following is an outline of the 3 stages undertaken for this research project from data collection through to data analysis. Each stage contains several steps. Detailed illustration of the various steps relating to data collection is presented in the following sections of this chapter.



## **Stage one: Preparation phase**

- Step 1: Recruiting research participants (students and external assessors)
- Step 2: Setting up experimental groups and control groups
- Step 3: Preparing Chinese source speeches
- Step 4: Preparing interview materials and self-evaluation activities for spoken-voice awareness-raising group
- Step 5: Preparing materials for voice-training groups
- Step 6: Preparing assessment materials for external assessors

## **Stage two: Data collection phase**

- Step 1: Collecting video recordings of pre- and post-intervention interpreting assessments
- Step 2: Collecting interview responses from spoken-voice awareness-raising group
- Step 3: Collecting reflective journals from spoken-voice workshops
- Step 4: Collecting rating scores and comments from external assessors

## **Stage three: Data analysis phase**

- Step 1: Conducting quantitative data analysis using SPSS, statistical analysis software
- Step 2: Conducting qualitative data analysis using qualitative content analysis (Dörnyei, 2007; Schreier, 2012) and a magnitude coding method (Saldana, 2018)

### **3.1 Material preparation stage**

#### **3.1.1 Recruiting research participants**

Two groups of research participants were recruited in this study, including student participants and 3 external assessors.<sup>2</sup> The student participants are the focus of this study. The assessors are not the research objects of the study, although their role is important. Data was elicited from their assessments of the student participants' vocal performance on the video recordings. Following are the details of how the student participants and external assessors were selected and contacted.

---

<sup>2</sup> The researcher received ethics permission from MUHREC on 21 March 2013. Project Number: CF13/1404 – 2013000728. Project Title: Investigation into voice and speech training for interpreters.

### **3.1.1.1 Selecting student participants**

Student participants were selected to ensure homogeneity at baselines such as language combination and language background, residing in predominantly Anglophone countries or not, degree track and previous formal or informal interpreting experience. The possibility that a student participant had already received spoken-voice training before the study also needed to be considered and any such students were excluded from the study. The selection of student participants in this study was based on the following criteria:

1. All student participants must not have resided in an Anglophone country for more than 6 months before they were recruited for this study.
2. All student participants must be beginner interpreting students in their first semester of interpreter training, with no previous interpreting training or work experience before the study, be it formal or informal.
3. All student participants must not have undertaken voice-training courses before the study.
4. All student participants must share the same A language – Mandarin.

Based on the above criteria, altogether 20 interpreting students of Chinese ethnic background were selected and recruited on a voluntary basis as research participants over 3 years. They were all beginner students who had just embarked on a Monash University Master of Translating and Interpreting Studies (hereafter referred to as MITS) program at either the Melbourne (Australia) campus or the Suzhou (China) campus; 15 were from the Melbourne campus and 5 from the Suzhou campus. The student participants' background information is presented in Table 3.1.

As Table 3.1 shows, all recruited students were from a Chinese background, Mandarin was their A language and they had studied English formally as their second language for more than 10 years. None of these students had had any previous interpreting training or work experience, or had received voice training before this study. There was a clear predominance of females in the study population, with only one male student in the control group. However, gender is neither a variable nor a focus of this study. The students recruited from the Suzhou campus had not been abroad. They

had studied and had contact with the English language exclusively via formal instruction in China. Students in the other 3 groups had arrived in Australia shortly before enrolling in the MITS and had not travelled to other Western countries before arriving in Australia. They had also received English language education in China, similar to the students of the short daily voice training group. Therefore, the difference regarding years of English language education outside of China can be considered relatively insignificant.

Table 3.1 Summary of student participants' background information.

<b>Student code</b>	<b>Gender</b>	<b>Years of learning English</b>	<b>Any formal or informal interpreting-related experience before entering the MITS</b>
A	F	14	No
B	F	14	No
C	F	14	No
D	F	12	No
E	F	14	No
F	F	13	No
G	F	15	No
H	F	13	No
I	F	14	No
J	F	15	No
K	F	12	No
L	F	15	No
M	F	15	No
N	F	12	No
O	F	15	No
P	F	16	No
Q	F	12	No
R	F	13	No
S	F	14	No
T	M	12	No

\* To protect the privacy and identity of informants, the researcher used single letters to identify the students that do not relate to the first letter of either their first or last names.

Monash University is a member of the Conférence Internationale Permanente d'Instituts Universitaires de Traducteurs et Interprètes, an international association of university institutions

with translation and interpretation programs. The MITS program is an approved course by the National Accreditation Authority for Translators and Interpreters (NAATI).

When collecting data from the student participants from 2013 to 2015, the MITS was a one-and-a-half-year (3-semester) translation and interpretation postgraduate program that had an intake once a year, but was upgraded to a two-year program from 2017. The student numbers offered a place in the Mandarin–English stream in each calendar year was modest, between 5 and 12 students, which resulted in a small intake of student participants in this study. This is one of the limitations of this research project.

The researcher started to contact potential student participants in early 2013 when she was involved in teaching at Monash University Melbourne campus, where she has been teaching ever since. In 2015, she recruited a group of student participants from the Suzhou campus in China when she was sent there to teach for two months. The selection of these two sites ensured that a certain number of Mandarin–English pairs could be contacted and selected, with inclusion criteria restricted to those whose A language was Mandarin.

### **3.1.1.2 Contacting student participants**

In early March 2013, the researcher first contacted potential student participants who had signalled interest to confirm their willingness to participate in the study to form a control group. Using the same method, the researcher contacted the other 3 cohorts' potential student participants in early March 2014 and 2015. All prospective student participants were informed of the general purpose of the study, the nature and duration of the intervention, the confidential treatment of each student, which means all the information about participants would be de-identified to ensure their anonymity, and secure storage of data. They were also informed that their participation was entirely voluntary and that they could withdraw at any time. All the student participants were informed that their participation or non-participation would not affect their course grades before the first video

recordings were collected. The Explanatory Statement and Consent Form (see Appendix A-1) were provided to potential student participants to read and sign.

### **3.1.1.3 Selecting external assessors**

This study has utilised “multiple raters’ appraisal”, which is also called “360-degree evaluation” input (Cleveland et al., 1989; Fleenor & Prince, 1997), to assess the student participants’ vocal performance. Multiple inputs provide a broader and more in-depth understanding of how participants perform and involve many sources in measuring results. The purpose of employing multiple assessors in this study is to understand the divergent views held by various interpreting users on the quality of spoken-voice intervention for interpreting trainees. Guided by this rule, 3 external assessors from different professional backgrounds were recruited:

1. A voice trainer with extensive knowledge of voice pedagogy, voice-quality assessing and diagnosis, and applied voice skills; the voice trainer was a specialised spoken-voice user
2. An interpreting instructor who was also a professional interpreter, had comprehensive linguistic knowledge in the language pair of this study, namely English and Mandarin, with first-hand interpreting work experience and interpreting teaching and assessment experience
3. A lay assessor (an interpreting audience) who was an individual interpreting user or audience member who had nothing more than cursory knowledge of interpreting and voice study

The voice trainer assessor was an expert in the analysis and diagnosis of spoken-voice quality. At the same time, the voice trainer assessor was also a specialised audience who was an interpreting end-user and a voice-training professional who had “clinically well-trained ears” (Hammarberg et al., 1980, p. 90). The interpreting instructor assessed the interpreting students’ vocal performance based on a set of interpreting quality criteria recognised in the interpreting industry. The lay assessor, who was the end user of interpreting services and an outsider to the interpreting profession, judged the students’ spoken-voice presentation from an audience perspective. Each assessor provided an independent and unique perspective for understanding how much effect the spoken-

voice intervention had on students' vocal performance. It is worth noting that these assessors were not involved in any of the interpreter training and intervention activities in this study. Thus, their assessment was not influenced by knowledge of any of the intervention activities conducted in this study. Table 3.2 shows the 3 assessors' profile. Again, their names are removed to protect their privacy.

Table 3.2 Assessors' profiles.

<b>Assessors</b>	<b>Years of working in own professional field</b>	<b>Professional background</b>
Voice trainer	20	She was an actor. She had worked as a voice lecturer at the University of Melbourne.
Lay assessor	25	She had worked as a marriage and family counsellor and therapist for around 25 years and had used interpreting services on many occasions.
Interpreting instructor	Around 30	She had worked as a professional interpreter in Singapore, New Zealand and Australia over the past 20 years. She later settled in Melbourne, working as a freelance interpreter and interpreting instructor.

#### **3.1.1.4 Contacting external assessors**

The researcher contacted the potential assessors to explain the research project's intentions and seek their consent to participate in the study. After receiving their consent to participate in the research project, the assessors were provided with an Explanatory Statement and Consent Form (see Appendix A-1). To encourage the assessors to provide candid responses rather than the responses that they may have thought the researcher wanted to receive from them, the researcher made it clear that their evaluation of students' performance was entirely for research purposes and that their scores or assessments would have no impact on the students' academic results.

#### **3.1.1.5 Researcher's positionality and qualification**

Apart from serving as the primary investigator for this study, the researcher has also been involved in interpreting academic curriculum instruction with 4 cohorts of student participants since 2013. Hence, the researcher was both an interpreting instructor and a data collector. Her contact with the

student participants included conducting interpreting workshops 2 to 4 hours per week and collecting data samples for the research project. However, the researcher was not involved in any assessment of students' vocal performance in this study. Instead, student participants' vocal performance was assessed entirely by the external evaluators to guard against a possible conflict of interest in data integrity. This sought to minimise the effect of bias that the assessors may have had towards the research project. Further, the same interpreting training curriculum was applied to all 4 cohorts – this ensured uniformity and consistency in the content of the interpreting training across all 4 groups.

Moreover, to gain an in-depth understanding of the technical aspects of the spoken-voice training, the researcher completed a Graduate Diploma in Voice Studies at the Victorian College of the Arts (VCA) at the University of Melbourne. There she met some experienced voice teachers and professional actors, and had a chance to work with them in individual and group settings; more importantly, the researcher received good advice on compiling the Interpreter Voice Skill Evaluation Sheet from her lecturers and classmates for her research project.

The researcher has personally been practising spoken-voice skills since then, reading voice-related journals and books to search for a spoken-voice training method that suits interpreting students and attempting to find connections between spoken-voice training and interpreting pedagogy. In conjunction with this research, she presented a pilot poster at the 2016 Conférence Internationale d'Instituts Universitaires de Traducteurs et Interprètes (CIUTI) annual forum in Geneva, where she made a brief speech introducing this research project and received positive feedback from the conference attendees.

The 8-hour spoken-voice training workshops were co-developed with Ms P, a professional actor attending the VCA program. Ms P regularly works as an actor in conjunction with her private practice as a speech coach to clients who speak English as a second language. Ms P and the researcher modified the voice-training materials originally designed for actors to suit interpreting

students and resemble a short spoken-voice training course in progress and content for interpreting student participants at the Monash Melbourne campus. Because Ms P was unavailable to travel to China at the time, the researcher delivered the spoken-voice training workshops at the Suzhou campus without Ms P's presence. This is another limitation of the study. However, neither Ms P nor the researcher assessed the student participants' vocal performance in this study.

### **3.1.2 Setting up experimental groups and control group**

As mentioned in Section 1.2, different views and attitudes may be held by interpreters towards spoken-voice training. Hence, 3 intervention groups and one control group were set up to mirror these interpreters' views and attitudes. It was expected that data would be gained from the students in each intervention group who had different experiences from each other and that these experiences would serve as a variable for comparison between the groups, enabling the research questions to be addressed.

The non-intervention group (control group) consisted of interpreters who had a "get by" mentality and did not seek to develop their spoken-voice skills or believe that interpreting training could improve their spoken-voice skills. Setting up a non-training group enabled comparison of the groups of students who received different spoken-voice interventions and those who did not. The spoken-voice awareness-raising group represented the interpreters who believed that simply monitoring their vocal performance would make them perform well instead of needing to practise spoken-voice skills. The spoken-voice training group corresponded to interpreters seeking to improve their vocal performance. The group was formed to understand whether and how much progress the student participants would make by participating in spoken-voice training.

The spoken-voice training group was further split into two groups that were set apart by different voice-training instruction modes so as to compare the effectiveness of two different voice-training regimens. The two groups shared the same training content, activities and exercises designed for them, but had different lengths of each training session and spaces between training sessions. One



group was designed to have a longer training time per session and large intervals between the training sessions, while the other had a much shorter training time per session with more regular but shorter breaks between the training sessions. Below is an overview of the 4 groups:

*The control group* (abbreviated as the CoG hereafter): No intervention was provided to this group.

*The spoken-voice awareness-raising group* (abbreviated as the ARG hereafter): This group was formed to observe the outcome of the spoken-voice awareness-raising intervention. Two interview sessions and self-evaluative activities were conducted at an office at Monash University Melbourne Caulfield Campus between the pre- and post-intervention interpreting assessments. The students of the group were interviewed individually. Each session lasted two hours. No voice training was provided to the students of the ARG.

*The long-interval voice-training group* (abbreviated as the LTG hereafter): These students received spoken-voice training in early April 2015. The workshops were conducted fortnightly from mid-April to early June 2015 in a pre-booked classroom at Monash University Melbourne Caulfield Campus. The workshops were delivered every other week in 4 separate sessions over 8 weeks. Each session lasted 120 minutes.

*The short daily voice-training group* (abbreviated as the STG hereafter): From mid-June to early August 2015, voice-training sessions for the STG students were held at Monash University's Suzhou campus in China. The 8-hour voice training sessions were divided into 24 smaller sessions with each lasting about 20 minutes and conducted on successive weekdays.

### **3.1.3 Preparing Mandarin source speeches for pre- and post-intervention assessment**

Two different Mandarin source speeches were used for the pre- and post-intervention interpreting assessments. The source speeches were 3 minutes long and were divided into two parts, with 150 words in each part, and were pre-recorded to ensure recording quality. Both pre- and post-intervention Mandarin source speeches were on the topic of "Cooperation in the field of education between China and Australia", a familiar topic to the students. Both were of low to medium level of

difficulty and interpreting technicality to require the conscious use of spoken-voice skills. The texts were adapted and modified from authentic speeches to enable the students to apply a variety of interpreting and vocal strategies. The pre-intervention speech was originally delivered by Mr Chen Yuming, former Chinese Ambassador to Australia, at the Australian Chamber of Commerce Perth Education Forum in 2012. The post-intervention speech was based on the remarks of Mr Chen Jianjun, head of the International Cooperation and Exchange Division of China's Ministry of Education, at the China–Australia Cross-border Education Cooperation Forum 2012. The speeches were written texts intended for verbal delivery at a conference. Because no video or audio recordings of the Mandarin speeches were available, the researcher recorded their speeches using a digital recorder at an average speech rate of 130 Chinese characters per minute.

The original texts were modified by deleting some remarks that were not relevant to the education theme, to make them both of similar length. Some lexical and syntactic features, references to numbers and quoted remarks, which are a frequent source of problems for novice interpreters, were retained, to test the students' spoken-voice presentation strategies. After the modifications to the texts, the pre-speech had 1138 characters and the post-speech had 1102 characters, equal to 3 minutes of speaking and interpreting (see Appendix B-1 and B-2 for full transcripts and reference translation for the pre- and post-intervention speeches, respectively).

Employing Mandarin source speeches in the pre- and post-intervention tests was done for the following reasons. First, as mentioned in Subsection 3.1.1.1, all student participants were from Chinese ethnic backgrounds speaking Mandarin as their A language and English as their B language. Thus, English listening comprehension may have challenged their interpreting performance. However, cognitive capability, including listening comprehension, is not the research focus of this study. Thus, assessing the students' interpreting performance from the A language into the B language would eliminate the potential barriers to English listening comprehension.

Conversely, interpreting from the A language to the B language brought a challenge to the students in terms of their verbal expression and interpreting delivery. According to Gile's (2009) efforts model, more monitoring and delivery efforts are required from interpreting students working into their B language. This effort may cause them to overlook their rhetorical and vocal presentation, which they may otherwise focus on if interpreting into their A language.

Thus, the Mandarin to English interpreting directionality would make all students subject to the same disadvantage or difficulty of interpreting into their B language, which accords with the objective of this study. Also, Mandarin into English directionality was chosen to facilitate the inclusion of assessors who were English native speakers. It would have been much more challenging to locate and secure the involvement of Mandarin-speaking assessors with similar profiles to those described in Table 3.2.

### **3.1.4 Preparing intervention for spoken-voice awareness-raising group**

In-depth semi-structured interviews and self-evaluation activities were undertaken with the purpose of raising the ARG students' spoken-voice awareness, based on the MRI method, which would potentially facilitate their spoken-voice skills (refer to Subsection 2.3.4). The implicit instruction method was adopted in MRI sessions. Instead of explicitly explaining detailed intervention objectives and procedures to the students, the researcher asked them a list of predesigned interview questions and asked them to complete some self-reflection activities. The students were aware of the interview and self-reflective activity, but were ignorant of the learning that would occur and could not have described the learning process.

The interview questions and self-evaluation sheets were designed to provide opportunities for the students to increase their awareness of the essential spoken-voice attributes and to apply these in their daily interpreting delivery practice, either consciously or subconsciously. Subsection 3.1.4.1 reports on pilot interviews conducted prior to the main study to determine the validity of the data

collection tools. For the main study, Subsection 3.1.4.2 presents the initial and follow-up interview questionnaires, as well as the self-evaluation sheets.

#### **3.1.4.1 Pilot testing interview questionnaires**

A pilot interview was conducted in early March 2013 with two interpreting students recruited from the Mandarin–English language stream in MITS. The interview procedures and questionnaires were administered the same way as they were to be conducted in the main study. The two students were not included in the main study because they had already been exposed to the pilot interviews. The purpose of conducting pilot interviews was to validate the interview questions, including the wording and order of the questions, and to try out the interview process. Notes were taken during the pilot interviews to initiate a revision of questions where necessary. After the pilot testing was done, the student participants were asked to identify what questions might sound ambiguous or unclear, which were to be revised for the main study. Through the pilot study, the researcher found the interview questions and self-evaluation activity elicited rich responses from the student participants. Based on the interviewees’ feedback, interview questions were revised as follows:

1. “Yes” and “no” questions were added to the open-ended questions, which allowed the student participants to elaborate on their responses.
2. Questions that did not stick to the research questions or elicit relevant information were revised.
3. Long questions that contained too much information were shortened or divided into separate smaller questions to prevent ambiguity.
4. Probes and prompts were added to the open-ended questions to encourage the interviewees to talk more.

After completing the pilot testing of the questionnaires, the researcher finalised the interview questions for the main study.

### **3.1.4.2 Devising interview questions and self-evaluation sheet for main study**

This section contains two parts. Part 1 is about the setting up of the interview questionnaire and Part 2 is about the planning of the self-evaluation sheet.

#### **Part 1 – Interview questions**

Following comments from the pilot interviewees, 14 questions were finalised for the initial and follow-up interviews in the main study. Prompts and probes were included in the questions to elicit more information from the interviewees – the student participants. The probes and prompts served as a reminder for the researcher to remember each question’s objective and keep the interview process on the right track. As deMarrais and Lapan (2003) pointed out, since participants can react to the interview experience differently, the guide and questions should not necessarily be designed as a one-size-fits-all protocol for each interview. As such, probes and prompts are suggested to be designed along with the interview questions to elicit responses from participants who are more reticent or hesitant in their responses.

The initial and follow-up questions were listed next to one another to allow for comparison of the differences in response. Questions 1 to 5 aimed to elicit information about the interviewees’ attitudes and experiences with vocal management and voice care. The purpose of Questions 6 to 11 was to elicit interviewees’ knowledge and self-reflection regarding various spoken-voice skills pertinent to the interpreting sector. Questions 12 to 14 were devised to prompt the interviewees to consider the importance of spoken voice in interpreting training and practice (details of the interview questions can be found in Appendix C-1).

Follow-up questions were designed to ascertain if the student participants’ responses to the topics discussed had changed since the initial interview. For example, Question 4 was: “What have you done to your spoken voice after the last interview, such as undertaking spoken-voice-related training or lessons or initiating self-practice exercises?”. This question helped the researcher understand the students’ actions after the initial interview session.

## **Part 2 – Self-evaluation sheet**

After completing the interview questions, the students were provided with video recordings of themselves that were taken at the pre-intervention interpreting assessment and they were invited to provide a self-appraisal based on a list of spoken-voice attributes provided in the self-evaluation sheet (see Appendix C-2). As mentioned in Subsection 2.1.4.3, self-evaluation of their own interpreting performance is one of the metacognitive strategies that could help the students to be more conscious of their performance.

A 5-point rating scale was developed in the self-evaluation sheet. The students were required to grade their vocal performance by circling one of the 5 ratings on the evaluation sheet. The 5-point scale ranged from “Very good” at one end to “Very poor” at the other, with “Average” in the middle. The rating scale served only as a reference for comparing the degree of satisfaction between the first and second evaluations, rather than being used as a quantitative instrument to analyse the individual scores statistically. To elicit more comments from the students, additional questions that captured open-ended feedback were included to obtain more profound insight into the effects of the vocal awareness intervention.

### **3.1.5 Preparing spoken-voice training for voice-training groups**

Explicit instruction (see Subsection 2.2.3.9) was employed for the two spoken-voice training groups, the LTG and the STG. The learning objectives, training schedules and materials were all clearly outlined through explicit instruction (see Appendix D-1, D-2, D-3 and D-4). Several metacognitive instruction approaches (refer to Section 2.3.4) were also applied to enhance the students’ metacognitive awareness and self-regulation. For example, a PBL strategy (Dunlap, 2005; Paris & Paris, 2001; Turan et al., 2009) was employed to guide the students to self-monitor their own vocal performance, focus on their vocal habits, find their unique vocal issues and understand why they had to practise specific exercises before commencing each spoken-voice activity and exercise, applying the metacognitive concept of thinking about their own thinking process (Flavell, 1979). In

addition to the students' self-monitoring, the trainer took an individual approach – monitoring every student during the vocal exercises to give instant and specific feedback according to the students' personal vocal attributes and performance.

The student-led small group metacognitive discussion employed a learner-centred approach (McCombs & Whisler, 1997), focusing on individual vocal management problems. The trainer provided directions regarding the questions that the students raised during the group discussion, guiding the students towards the path of further vocal practice on their own. Thus, the students not only received information from the trainer, but also learnt how to apply it to lifelong and self-directed learning (Hoban et al., 2005).

The purpose of writing a reflective learning journal straight after each exercise or at the end of each training session was to stimulate the students to contemplate their own cognitive processes and regulation of cognition (two metacognitive dimensions) (Henter & Indreica, 2014).

To compare the effectiveness of the two distinct voice-training modes and to determine the optimal voice-training regimen for interpreting students, two separate spoken-voice training schedules for the two groups were devised which are presented in Subsections 3.1.5.1 and 3.1.5.2. Subsection 3.1.5.3 presents some examples of the spoken-voice training exercises that were developed to enable the students to acquire and develop vocal skills that contributed to their overall interpreting performance.

#### **3.1.5.1 Designing spoken-voice training schedule for LTG**

The researcher collaborated with Ms P to develop the training content and delivered the training to the LTG students (see Subsection 3.1.1.5). Ms P was responsible for delivering the general information sessions on voice and speech production, and guiding the students in various vocal exercises to improve their vocal skills and techniques. The researcher instructed the students in specific vocal exercises related to interpreting practice and provided on-site feedback. The training

sessions were video recorded with the consent of the student participants. Each session included a training schedule as follows:

- Opening (voice warm-up) (20 min)
- Introduction to a particular aspect of spoken-voice training (30 min)
- Spoken-voice exercises (50 min)
- Reflection on practice and group discussion (10 min)
- Closing, self-reflective notes (10 min)

The voice-training schedule for the LTG (see Appendix D-1) contained the date, training topics, training exercises and learning outcomes expected for each training session. As indicated in the schedule, the first workshop began with an introduction to spoken-voice production and some warm-up activities, followed by an invitation for students to give a 2- to 3-minute impromptu speech on a given topic using their B language, English. The students' impromptu speeches were videotaped for self-reflection. During the second workshop, the students were given an overview of vocal care and guided on appropriate posture, body alignment and movement, breathing techniques, voice projection, articulation, and physical and emotional tension balance. The third workshop focused on improving students' spoken-voice quality and vocal dynamism, including pitch, intonation and speaking tone, emphasis and speaking clearly. Physical tension in facing a large audience and maintaining audience interaction in CI situations were also discussed. The final session revisited the critical points discussed over the 8 weeks. The students played roles as both speakers and interpreters in turn and were asked to appraise their own and their peers' vocal performance. Their speech and interpreting performances were video recorded for their self-reflection. The students received feedback from Ms P and the researcher individually.

Each voice-training session was wrapped up with a group discussion focused on various questions that the students may have raised. During the discussion, Ms P and the researcher monitored the



students' questions and gave timely guidance and constructive feedback. At the end of each session, the students were given a few minutes to write down their reflections for the training session.

### **3.1.5.2 Designing spoken-voice training schedule for STG**

As indicated in the STG's voice-training schedule (see Appendix D-2), each session focused on a different topic, vocal technique or activity. In the first 14 sessions, the students were given an overview of vocal care and spoken-voice production. The discussions and exercises were focused on appropriate posture, body alignment and movement for interpreters, breathing techniques, voice projection, articulation, and physical and emotional tension balance. The students were also invited to make a 2- to 3-minute impromptu speech on a given topic in their B language (English). The impromptu speeches were video recorded for post-workshop feedback and the students' self-reflection.

The following 9 sessions focused on improving the students' spoken-voice quality and vocal dynamism, including pitch, intonation and speaking tone, emphasis and speaking clearly. The students were also guided to discuss physical tension in facing a large audience and maintaining audience interaction in CI situations.

The final session revisited the critical points discussed over the 8 weeks. The students played roles as both speakers and interpreters. After the speaking and interpreting activity, the students were asked to appraise their own and their peers' vocal performance. Their speech and interpreting performances were video recorded for the students' self-reflection and post-workshop feedback. Each session ended with a short group discussion. At the end of each session, the students spent 2 to 5 minutes writing down their reflections.

Because Ms P could not travel to China to provide voice training to the STG students as she did for the LTG in Melbourne, the researcher delivered the voice-training sessions on her own. However, Ms P kindly provided a few voice-training video recordings containing general information on voice and speech production and exercises. The researcher played these video recordings to the

STG students and also took some video recordings during the voice workshops and sent them back to Ms P for post-workshop feedback and comments. As mentioned, the STG and LTG students received the same amount of training (8 hours) with the same training content, including spoken-voice activities and exercises. These activities and exercises are presented in the next section.

### 3.1.5.3 Planning spoken-voice exercises for interpreting students

There were 23 spoken-voice exercises included in the training material. These exercises were chosen because they were related to the work of interpreters and were reasonably simple to master in comparison to other more complex spoken-voice training methodologies and techniques. Each exercise was chosen based on its primary purpose and in accordance with the grading criteria specified in the Interpreter Voice Skill Evaluation Sheet (see Appendix E-1 and E-2). The 23 exercises and their principal purposes are summarised in Table 3.3.

Table 3.3 List of spoken-voice exercises for interpreting students.

Item	Exercise	Primary purpose
1	Ice-breaking activity	To help students connect emotion, body and voice when interacting with people around
2	Drawing body image (otherwise referred as body mapping)	To help students increase their awareness of physical aspects of voice production
3	Posture and alignment and vocalisation exercises	To help students be aware of their posture while standing and sitting, correct their posture and experience the connection of posture and spoken voice
4	Kinaesthesia exercise	To help students be aware of posture while moving around and experience the connection of movement and spoken voice
5	Semi-supine position and vocalisation	To help students be aware of posture, experience the connection of body and spoken voice, and reduce tension and performance anxiety
6	Physical and vocal warm-up and vocalising exercises	To help students warm up physically and vocally, prepare for challenging interpreting tasks and reduce tension and performance anxiety
7	Breathing and vocalising exercises	To help students ground and deepen their breath, improve alignment and enhance their sense of empowerment
8	Storytelling	To help students improvise in speaking and understand the structure of speech
9	Resonance exercise	To help students develop chest resonance in speaking

10	Voice projection exercise	To help students project and sustain sound and clarity of speech
11	Articulation exercise	To help students practise using strong, clear articulation of consonant and vowel sounds, and improve speech clarity
12	English accent-reduction exercise	To help improve speech clarity and reduce English accent
13	Pitch exercise	To help students explore and extend their pitch range
14	Pace and pausing exercise	To help improve pace and pauses in speaking
15	Speech shadowing	To help students improve natural pronunciation, intonation and fluency and produce sounds accurately
16	Voice personas (also known as character traits)	To help students connect voice and emotion, speaking fluency and expressiveness, and emotion and words, and connect with the audience
17	Connecting voice and body	To help students connect voice and body, and understand better how speech is constructed
18	Intonation and tone exercises	To help students change their tone of voice to express emotion
19	Rhythm and fluency exercise	To help students improve fluency, intonation, emphasis and rhythm
20	“Me”: self-consciousness reduction exercise	To help students reduce tension and excessive self-consciousness
21	Role-playing of impromptu speaking and interpreting	To help students improvise in speech, understand the structure of speech and experience roles as a speaker and an interpreter
22	Role-playing of a prepared speech, interpreting and evaluating	To help students develop a pre-performance vocal warm-up routine and experience roles as a speaker and an interpreter
23	Role-playing: speaker and audience	To help build rapport and friendliness with the audience and improve confidence

Exercises 1–7 were used at the beginning stage of the voice training targeting the students’ spoken-voice production skills, including alignment, posture, movement, physical tension balance and breathing support (see Appendix D3 and D-4). Following is an example of the training instruction:

### **Exercise 7: Breath and vocalisation exercises**

1. Students sit comfortably in a chair, with hands in their lap and feet on the floor. The instructor guides them to relax back into the chair. The students become aware of their back resting comfortably against the back of the chair and experience the weight of their buttocks resting easily on the seat of the chair. Their lips are lightly closed – tongue flat. The tip of the tongue is resting behind the bottom front teeth.

2. The instructor guides the students to focus on their breath. The students breathe in through their nostrils and into their abdomen, and back out. They imagine the inhalation filling up their abdomen and ribcage. The instructor reminds the students to be aware of the full 360 degrees of their ribcage. The students imagine their diaphragm moving with each inhalation and exhalation. The students begin to feel increasingly grounded as they experience their breath going into a deeper part of their body. As they continue to breathe in through their nostrils, they experience an increasing sense of autonomy and strength.
3. The instructor guides the students to exhale on the voiceless fricative “f” or “s” while engaging their abdominal muscles to support exhalation. Repeated three times.
4. Students exhale on the voiced fricative “v” or “z”. Repeated three times.
5. Students breathe into their lower back.
6. Students repeat each line after the instructor. The students take a quick, deep and silent breath after each line to ensure enough breath to sustain the thought so that the volume will not fade. The instructor adds words to make the sentence longer and longer to build up the students’ breath control. Following is an exercise is adapted from Rodgers (2002, p. 33):

“Hello!”

“Hello, I am a feminist!”

“Hello, I am a feminist because it bothers me!”

“Hello, I am a feminist because it bothers me that a woman gets killed!”

“Hello, I am a feminist because it bothers me that a woman gets killed by her male partner!”

The above training content was chosen from a speech against family violence. The reason for choosing a topic such as family violence is that violence against women is frequently encountered in interpreters’ work. In such an interpreting context, interpreters are required to express the victim’s emotion through their voice and speech. Novice interpreters may find the topic challenging. They may be emotionally involved in the SL speech and dialogue, forgetting to take a breath and being unable to remain emotionally detached from the victim. Training using such challenging topics would provide an opportunity for interpreting students to practise breathing skills, release their tension and adapt better to a challenging interpreting situation.

The instructor had a responsibility to create a learning environment that encouraged students to participate. However, violence against women affects almost everyone in some way. It can be emotional and challenging for students to encounter complex topics that touch them personally. Some students may not be able to stay emotionally neutral while vocalising the victim's emotion. One way of dealing with this is to agree on group norms at the beginning of the activity. The voice instructor may convey a message through a group discussion about whether the students have experienced a violent relationship or not. If the group gets off-topic or problems arise, the instructor can refer to the list of ground rules and bring the group back on track.

Exercise 8–14 were used at the mid-stage of training, focusing on spoken-voice quality, including vocal volume, resonance, vocal tone, pace and pausing, speaking clarity and modulation of the pitch range. The rest of the exercises focused on spoken-voice dynamism such as interpreters' intonation, rhythmic and fluent delivery, expressiveness, vocal confidence and audience engagement.

Following is an example of an English accent-reduction exercise:

### **Exercise 12: English accent-reduction exercise**

1. Strengthening word stress and emphasis

e.g.: product /'prɒdʌkt/

colleague /'kɒli:g/

2. Improving English weak vowels (otherwise referred to as the schwa sound)

e.g.: carrot /'kærət/

support /sə'pɔ:t/

purchase /'pɜ:tʃəs/

onion /'ʌnjən/

3. Distinguishing English vowel sounds

\*Minimal pairs word exercises:

e.g. sleep /i:/ slip /ɪ/

gem /e/ jam /æ/

hut /ʌ/ heart /ɑ:/

cushion /ʊ/ fusion /u:/

\*Minimal pairs sentence exercises:

e.g.: I don't like these *spots*.

I don't like these *sports*

4. Commonly mispronounced words  
e.g.: recipe, salmon, says, suit, suite, chaos, women, nature
5. Learning speech sounds that are not used in Mandarin  
e.g.: “v” as in value or five  
“z” as in authorise  
“sh” as in share or punish  
“zh” as in vision or measure  
“ch” as in chairman or cheque  
“j” as in judge or gentleman  
unvoiced “th” as in thing; and  
voiced “th” as in this or that
6. Learning combinations or “clusters” of sounds found in English but not used in Mandarin  
e.g.: words containing sl, sn, sm, br, cr, fr, bl, cl, spr, sq, nt
7. Speech melody and intonation  
Learning English sentence stress patterns to distinguish content words (like nouns and verbs) from function words (like prepositions and articles) and rules of intonation (e.g. to signal statements or yes/no questions)  
e.g.: Mary ‘went.  
Mary would ‘go.  
May would have ‘gone.

The last 3 exercises aimed to integrate students’ vocal skills into their interpreting performance. For example, in exercise 22 the students were instructed to take turns to play roles as a speaker or an interpreter. Students were encouraged to apply vocal skills in presenting their speech and interpret for the speakers. A speech and interpreting self-evaluation sheet were provided for peer evaluation and self-evaluation (see Table 3.4). The two left columns of the sheet listed vocal attributes that were chosen in accordance with the spoken-voice categories and attributes described in Section 2.2.4. It is worth mentioning, however, that the students’ self-evaluation sheet was not identical to the external assessors’ evaluation sheet in terms of function, complexity and aim. The student assessment sheet was designed to assist students in self- and peer-reviewing of their vocal performance (see Subsection 2.1.4.3). It was much more straightforward than the Interpreter Voice Skill Evaluation Sheet created for the external assessors. There were no numerical points to be

added to the evaluation sheet used by the 3 external assessors (see Appendix D-3 for detailed instructions for each exercise).

Table 3.4 Student Speech and Interpreting Self-Evaluation Sheet.

	<b>Negative</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>Positive</b>
Volume	Too quiet							Good projection
	Inaudible							Easily heard
	Flat							Vibrant and dynamic
Pitch	High							Low
	Shrill (head resonance)							Full (chest resonance)
	Monotonous							Varied
Voice quality	Nasal							Open
	Breathy							Clear
	Harsh, raspy							Mellow
	Lifeless							Enthusiastic
Articulation	Mumbling							Clear
	Lazy lips							Crisp
	Tangled tongue							Controlled
	Tight jaw							Open mouth
	Heavy accent							Accentless
	Mispronunciation							Correct pronunciation
Rate and pause	Jerky							Smooth
	Slow, plodding							Fluent
	Unvaried							Varied, exciting
	Hesitant							Deliberate
Vocal variety	Emotionless							Conveys emotion
	Unfriendly							Genial
	Strained							Natural
	Dull							Vital

### 3.1.6 Preparing assessment materials for external assessors

The Interpreter Voice Skill Evaluation Sheets, including the pre-intervention evaluation sheet and the post-intervention evaluation sheet, were designed for the 3 external assessors to evaluate the student participants' vocal performance based on their video recordings. Two factors were considered when designing the evaluation sheets. Firstly, both perceptual and objective rating

measurements were adopted. In clinical voice and speech assessment, perceptual and instrumental evaluations are the two main measures. Perceptual evaluation involves “subjective human assessment.” Oates (2009) defined the perceptual nature of voice quality as follows:

[voice quality is] perceptual in nature and the perceptual characteristics of voice have greater intuitive meaning and shared reality among listeners than do many other instrumental measures which has limitations in the validity and reliability and lack of agreements as to the most sensitive and specific instrumental measures of voice quality (p. 1).

Instrumental examination conducts “acoustic, aerodynamic or imaging-based analysis” (Shipley & McAfee, 2016, p. 421) to provide objective data. The design of the assessor evaluation sheets in this study was based on existing perceptual voice assessment measures, schemes and spoken-voice terminology. Perceptual voice evaluation was relevant to this research project because more stringent acoustic analysis of student participants’ performance is not the focus of this study. However, Oates (2009) also pointed out that perceptual evaluation has the limitation of subjectivity as “listener reliability is not always adequate and perceptual ratings can be confounded by factors such as the listener’s shifting internal standards and listener experience” (p. 1). Thus, the decision was made to include both perceptual and objective evaluation measurement tools in this study.

Secondly, the parameters set up for the Interpreter Voice Skills Evaluation Sheet were determined by extrinsic vocal features, rather than intrinsic vocal features. As Laver (1991) explained, “intrinsic features lie outside the normal control of an individual, are universal rather than culturally specific, and are not normally learnable” (p. 179). On the other hand, extrinsic features cover a range of vocal activities and behaviours that a speaker chooses to learn and exercise volitional control over based on the speaker’s intrinsic vocal features. As Laver (1991) stated:

The segmental choices of one’s pronunciation, the choices of manipulation of voice production, all the controllable habitual muscular settings which characterise the manipulable component of the speaker’s voice quality and to the extent of all the learnable aspects of the speaker’s persona (p. 179).







### 3.2 Data collection method

This section explains how data was collected for each experimental group over the 4 years from 2013 to 2016. The researcher spent 3 consecutive years collecting students' pre- and post-intervention interpreting video recordings, one group a year. Table 3.5 presents the data collection timeline for each group. Detailed data collection methods are presented in Subsections 3.2.1 to 3.2.4.

Table 3.5 Data collection timelines.

<b>Experiment groups &amp; Participants</b>	<b>Month/year</b>	<b>Data collection activity</b>
<b>CoG</b>	Mid-March 2013	Recruiting student participants
	Mid-March 2013	Collecting pre-intervention video recordings
	Mid-July 2013	Collecting post-intervention video recordings
<b>ARG</b>	Mid-March 2014	Recruiting student participants
	Mid-March 2014	Collecting pre-intervention video recordings
	Early April 2014	Initial interviews and self-evaluation activity
	Mid-June 2014	Follow-up interviews and self-evaluation activity
	Mid July 2014	Collecting post-intervention video recordings
<b>LTG</b>	Mid-March 2015	Recruiting student participants
	Mid-March 2015	Collecting pre-training video recordings
	Early April to mid-June 2015	Four voice training workshops over 8 weeks (2 hours per session)
	Mid-July 2015	Collecting post-training video recordings
<b>STG</b>	Early June 2015	Recruiting student participants
	Early June 2015	Collecting pre-intervention video recordings
	Early June to early August 2015	24 sessions of spoken voice training (20 minutes per session)
	End August 2015	Collecting post-training video recordings
<b>External assessors</b>	Mid-April to mid-July 2016	Collecting rating scores and comments

### **3.2.1 Collecting video recordings for pre- and post-intervention assessments**

Video recordings are an important data source used in this research project to measure and compare the students' vocal performance at the pre- and post-intervention assessments. A pilot video-aided test conducted by Kellett Bidoli (1995) of students' delivery and presentation in CI proved that the video recording is a valuable instrument for evaluating the delivery aspect of interpreting, irrespective of linguistic proficiency and accuracy, enables a more objective evaluation of students' interpreting performance and is a didactic tool that can be applied to interpreter training. Altogether, 40 video recordings were collected, 20 from the pre-intervention assessment and another 20 from the post-intervention assessment.

### **3.2.2 Collecting interview responses and self-reflection sheets from ARG students**

As mentioned in Section 3.1.4, the MRI intervention was designed to provide students with opportunities to reflect on their own spoken-voice skills and improve their spoken-voice awareness. The interview responses of the ARG students were elicited during the MRI intervention sessions. The students were encouraged to express themselves freely in their favourite language, either Mandarin or English. With the students' consent, audio recordings were made to capture the details of verbal responses. Simultaneously, the researcher took written notes. Each file of the interview recordings was labelled immediately upon completion of the interview sessions to avoid data duplication or loss. After the interviews, the comments were transcribed into writing for qualitative analysis (see Appendix F for samples of interview responses).

The initial interviews lasted from one-and-a-half hours to two hours for each student. The lengthy discussion allowed the interviewer (the researcher) to establish rapport with the student participants (the interviewees) and to foster trust (Maykut & Morehouse, 1994). Upon completing the interview questions, the students were invited to view a video recording of themselves taken at the pre-intervention interpreting assessment and to complete the self-evaluation sheet.

Three months after the initial interviews, follow-up interviews were conducted. The students were asked to respond to 14 questions that were almost identical to those asked at the initial interview but with a few small differences. The researcher then analysed the responses from the initial and subsequent interview sessions to ascertain the extent to which the students' responses had changed over the course of 3 months. During the sessions, each student was shown a video recording of themselves performing an interpreting task in an interpreting workshop. These recordings were retrieved from an interpreting workshop one-and-a-half months after the initial interview. These video recordings were used merely for the students to conduct a self-evaluation. The students performed a self-appraisal, either written or verbal, based on the self-evaluation sheet. The post-intervention assessment was completed one week after the follow-up interview sessions.

### **3.2.3 Collecting reflective journals from voice-training groups**

Ten reflective journals were collected from the two spoken-voice training groups, LTG and STG. As mentioned in Section 2.3.4, writing a self-reflective journal was one of the metacognitive strategies to enhance the students' metacognitive awareness and self-regulation (Henter & Indreica, 2014). The students were instructed to jot down a few thoughts immediately after each voice-training session when their memories were still fresh, such as their physical sensations in a particular exercise or any queries or views on specific activities or vocal techniques. There were no detailed instructions regarding the written style for the reflective journals. The students were also advised to write the journal simply for self-reflection, rather than as an assignment. The journal entries were collected and analysed to understand the effectiveness of the spoken-voice training, the level of improvement in spoken-voice awareness and what challenges may have prevented the students from utilising their spoken-voice skills effectively in their interpreting performance (see Appendix G for samples of reflective journals).

### **3.2.4 Collecting rating scores and comments from assessors**

From mid-April to mid-July 2016, the researcher contacted the 3 external assessors individually to collect their rating scores. These evaluations took place at either a library or another public place convenient to both the assessors and the researcher. A pre-assessment briefing was conducted to harmonise the evaluation standards among the 3 assessors. Prior to assessing the actual data, the researcher spent 15–30 minutes discussing the research project’s objectives and the rating scale, and presenting sample evaluation sheets and sample data to the assessors. The Evaluation Instructions (refer to Appendix A-2) and Glossary of Interpreter Voice Skills Evaluation Sheet (refer to Appendix A-3) were devised to help the assessors understand the evaluation sheets’ marking criteria.

Through the briefing, the assessors were made aware that the evaluation was not to check the students’ transferred interpreting skills or level of accuracy, nor was it the type of acoustic voice analysis that is commonly conducted by a speech pathologist, but was to assess the students’ spoken-voice skills from the point of view of an interpreting audience. At the conclusion of the briefing, the assessors were given the opportunity to clarify any remaining questions or concerns.

To avoid preconceived impressions and maintain objectivity and reliability, the assessors were not told which video recordings were taken before or after the intervention. As Mackey and Gass (2016) pointed out, “It may be desirable to keep coders selectively blind about what part of the data (e.g., pre-test or post-test) or for which group (experimental or control) they are coding in order to reduce the possibility of inadvertent coder biases” (p. 139).

Following the briefing, the assessors examined the pre- and post-intervention video recordings and offered rating scores and written or verbal comments on the evaluation sheet that had been predesigned. The researcher aided with the recording but made no personal comments.

### **3.3 Data analysis method**

This section discusses how the data was analysed through a mixture of quantitative and qualitative data analysis methods. Section 3.3.1 presents how a quantitative method was used to analyse the rating scores collected from the external assessors. Additionally, a quantitative method was used to analyse the correlation between spoken-voice performance and overall interpreting performance, as described in Subsection 3.3.1.3. Sections 3.3.2 and 3.3.3 discuss the qualitative method that was utilised to analyse the assessors' comments, the students' reflective journals and the students' interview responses.

#### **3.3.1 Quantitative analysis**

The quantitative results were generated via the SPSS system to address Research Questions 1, 2 and 3. To address the first two research questions, the numerical data including the pre- and post-rating scores collected from the 3 external assessors were processed separately to check their reliability. The results showed good internal consistency with a Cronbach's alpha coefficient value of 0.826, which is above 0.7 (DeVellis, 2012), suggesting internal consistency and reliability for the scale (see Appendix H for reliability statistics). Then, both within-group and cross-group comparisons were conducted using SPSS software.

##### **3.3.1.1 Within-group improvement**

Paired-sample t-tests were processed for the different groups to compare the data collected from pre- and post-interventions within the groups. The sig(2-tailed) item in the output (see Appendix I-1, I-2 and I-3) is the probability ( $p$ ) value. The  $p$ -value is the evidence against the null hypothesis. In a normal situation, the smaller the  $p$ -value, the stronger the evidence that the null hypothesis should be rejected. However, in this research project the  $p$ -value was not taken as the significance indicator because the sample size of each group was small ( $N=5$ ), which is one limitation of the research project. Instead, the  $\text{Eta}^2$  (otherwise referred to as Eta squared), which is one of the most used effect-size statistics, was employed to examine the intervention program's effect size. The  $\text{Eta}^2$  of

paired t-tests can be obtained using the formula:  $\text{Eta}^2 = t^2 / t^2 + (N-1)$  (Pallant, 2016, p. 253). The guidelines proposed by Cohen (1988) for interpreting the effect size are  $>0.01$  = small effect,  $>0.06$  = moderate effect,  $>0.14$  = large effect (pp. 284–7).

Further paired-sample t-tests were conducted to evaluate whether and to what degree each spoken-voice category (spoken-voice production, spoken-voice quality and spoken-voice dynamism) improved in the post-test scores (see Appendix J-1, J-2 and J-3).

### **3.3.1.2 Cross-group comparison**

The researcher undertook a one-way between-groups analysis of covariance test (ANCOVA) (see Appendix K-1) to compare the effectiveness of the different intervention methods designed for the 4 groups. The dependent variable was the post-intervention scores for student participants' vocal performance. The student participants' pre-total intervention rating scores were used as covariates in the analysis to remove any pre-intervention differences between groups. The estimated marginal means provided the adjusted means of the dependent variable for each of the 4 groups, which means that the covariate effect was statistically removed. If  $\text{Eta}^2$  was more significant than 0.14, the between-group difference was considerable. Further analysis was processed via the post-hoc LSD test to examine the detailed differences between every two means groups, group (i) and group (j). Any significance level that was less than 0.05 was taken as statistically significant.

To find out whether and to what degree there were significant differences between the 3 spoken-voice categories and which categories made the most significant difference to participants' interpreting performance, another ANCOVA analysis was conducted taking the post-intervention scores for each spoken-voice category as the dependent variable and the pre-intervention scores as covariates. The analysis details can be found in Appendix K-2.

It is worth noting that the study aimed to analyse and report the results for the scores of the within-group and cross-group comparisons collected from the 3 assessors separately, rather than the average scores from the 3 assessors, for two reasons. Firstly, the assessors examined the students'



vocal performance based on their professional standard, personal observation and judgement. Inter-rater reliability could not be ensured if the study results were based on the average scores. And secondly, it was desirable to obtain individual results from the 3 different angles, that is, from the voice trainer assessor, the lay assessor and the interpreting instructor assessor.

### **3.3.1.3 Pearson correlation coefficient analysis**

To explore whether there was a significant correlation between students' spoken-voice performance and their interpreting performance regarding the delivery, a Pearson correlation coefficient analysis was conducted via SPSS software, based on the total scores (N=60) collected from the 3 external assessors for 20 students' spoken-voice performance and interpreting delivery performance (See Appendix L-1). Another correlation analysis was conducted to understand which spoken-voice attributes contributed more to the overall interpreting performance in both pre- and post-assessments (See Appendix L-2).

### **3.3.2 Qualitative content analysis of interview responses and reflective journals**

To select the most appropriate approach to qualitative analysis, after carefully considering the goal of the research enquiry the researcher adopted qualitative content analysis (QCA) (Dörnyei, 2007; Schreier, 2012) from a range of qualitative methods. QCA is referred to by Hsieh and Shannon (2005) as "a research method for subjective interpretation of the content of text data through the systematic classification process of coding and identifying themes or patterns" (p. 1278). Schreier (2012) described QCA as "a method for systematically describing the meaning of qualitative material" (p. 1).

QCA was chosen for this study for several reasons. Firstly, QCA enables the researcher to extract categories from the collected data, understand both surface and underlying meanings from close study of texts and engage in data analysis unobtrusively. As noted in Subsection 3.1.1.1, the qualitative data was collected over 3 consecutive years because the student participants were recruited from 3 semester years. Consequently, some of the data had to be stored safely and wait to

be assessed by the external assessors and then analysed further. QCA allowed the researcher to conduct the data analysis based on the materials collected without the need to have further direct contact with the students, who may already have graduated from the MIT program by the time of data analysis (Cho & Lee, 2014). Furthermore, QCA can be used to analyse almost any kind of qualitative data, including narrative responses, open-end survey questions, interviews, focus groups, observations and printed media such as articles, books and manuals (Abrahamson, 1983; Hsieh & Shannon, 2005). In this study, QCA has been used to analyse the assessors' comments and the students' interview responses and journal entries.

More specifically, the inductive coding approach of QCA (Thomas, 2003) and thematic data analysis (Braun & Clarke, 2006) were used to analysis the qualitative materials including reflective journals and interviewees' responses. According to Braun and Clarke (2006), thematic data analysis tends to be driven by the researcher's conceptual or analytic interest in the area and therefore is "more explicitly analyst" (p.84). Figure 3.1 shows the data analysis procedure used to analyse the students' reflective journals (see samples of students' journals in Appendix G). The researcher started by focusing on the relevant aspects of the research questions and stayed close to the surface of words and sentences. Then the researcher carefully read the journal texts line by line and determined the preliminary codes that emerged from the text before coding the remaining text.

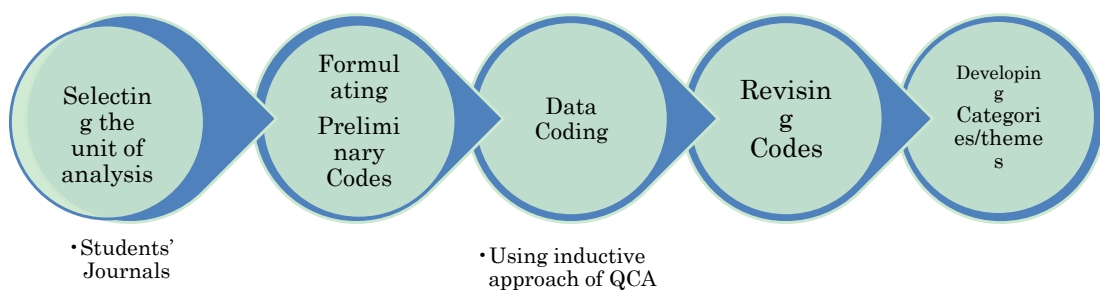


Figure 3.1. Procedure for analysing students' journals (STG & LTG)

After completing the first coding cycle, the researcher grouped similar codes and placed them into different categories (themes). In the second cycle of coding, the researcher revised and refined the categories and formed the final themes. While coding, the researcher constantly questioned the coded data to identify new meanings and link them to possible concepts. She also looked for evidence of relationships between the data. The categories/themes gave the researcher clues as to what factors may have led to student participants' improvement.

The same method was applied to analyse the data collected from the interviews. However, one step before starting the analysis process was transcribing and translating the entire interview content from audio recordings into written form. Transcribing is a time-consuming process, especially when it involves translating oral accounts. As Ochs (1979) asserted, "transcription is a selective process reflecting theoretical goals and definitions" (p. 44). Because the qualitative data in this study was mainly intended to complement the quantitative results, the researcher was more interested in transcribing the content (what the assessors said) than the form (how they spoke). Thus, the researcher did not use the type of verbatim transcript required for a discourse or conversation analysis. Nor did the researcher include imperfect speech features such as hesitation markers (er, uh). Since some interviewees preferred to speak Mandarin during the interviews, their discourse was also translated from Mandarin into English. During transcribing and translating of the data, the researcher went through the data several times, took notes and marked ideas for coding before the more formal coding process.

### **3.3.3 Magnitude coding method for assessors' comments**

The external assessors' comments were analysed by adopting a deductive coding approach (Elo & Kyngäs, 2008) to provide detailed information about the spoken-voice attributes that recorded high levels of progress in the post-intervention interpreting assessment. The 3 predetermined spoken-voice categories and 15 attributes (refer to Section 2.2.4) formed the list of codes before the analysis.

Creating content-related categories is a means of data reduction. Categories refer to matters “with similar meaning and connotations” (Weber, 1990, p. 37). They must be “mutually exclusive and exhaustive” (Crowley & Delfico, 1996, p. 20), which means no coded data should “fall between two categories or be placed in more than one category” (Cho & Lee, 2014, p. 10). Some categories may be refined into subcategories that contain clusters of coded data. By reducing a large amount of data to a few labelled categories, the researcher obtains an impression of what is inside the data and prepares for a more in-depth thematic or conceptual analysis. Figure 3.2 illustrates the data analysis procedure for the assessors’ comments.

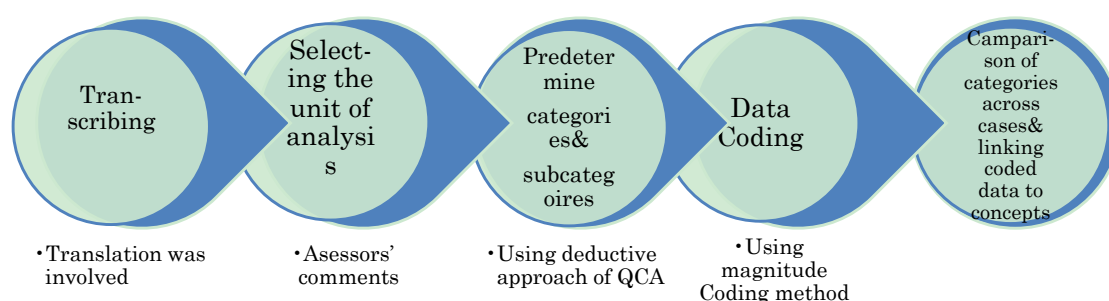


Figure 3.2. Procedure for analysing assessors’ comments using QCA

As Figure 3.2 illustrates, the researcher started to analyse the transcripts of the assessors’ comments upon completing the transcribing process. After selecting the units of analysis from the transcript, the researcher analysed the texts chunk by chunk and line by line, looking for words and phrases that appeared to describe the 15 spoken-voice attributes and 3 main vocal categories. The researcher employed the magnitude coding method (Saldana, 2018) consisting of descriptive coding and numerical ratings. The descriptive codes, including “Very positive”, “Positive”, “Neutral”, “Negative” and “No comments,” indicated each of the attributes that the assessors addressed. The numerical rating system, ranging from 0 (No comments) and 1 (Negative comment) to 4 (Very positive statement), was added to the descriptive codes, showing the researcher’s understanding of the evaluative content. This allowed the researcher to interpret the assessors’ comments on each

spoken-voice attribute's quality and compare the pre- and post-performance for each student. Table 3.4 shows the rating guidance and examples of the descriptive words and phrases in the assessors' comment text that the numerical ratings were based on.

Table 3.6 Rating scale descriptions for assessors' comment text.

Rating scale descriptions		
Indicators	Rating guidance	Examples of comment text from assessors
Excellent = 4	<ul style="list-style-type: none"> <li>❖ Knowledge, skills and abilities of a spoken-voice attribute have been demonstrated at an exceptional level highly satisfactory or exceeding expectations</li> <li>❖ Shows significant progress in an attribute</li> </ul>	<p>“Wonderful”, “Excellent”, “Impressive”, “Great/significant improvement”, “Dramatic” “Much better”</p> <p>“Had a really great emphasis”</p> <p>“I think her improvement has surprised me.”</p> <p>“Very significant improvement in her performance.”</p>
Good = 3	<ul style="list-style-type: none"> <li>❖ Knowledge, skills and abilities of a spoken-voice attribute have been demonstrated at high levels or satisfactory or meeting expectations</li> <li>❖ Shows consistent progress in an attribute</li> </ul>	<p>“Good”, “Nice” “Reasonable” “Better”</p> <p>“Her delivery is very smooth and very clear.”</p> <p>“She has marked improvement in her pronunciation”.</p>
Average = 2	<ul style="list-style-type: none"> <li>❖ Knowledge, skills and abilities of a spoken-voice attribute have been demonstrated at appropriate levels or meeting expectations</li> <li>❖ Shows some progress but may be inconsistent or inadequate</li> <li>❖ Improvement is needed in an attribute</li> </ul>	<p>“OK”, “Not too bad”</p> <p>“About average”</p> <p>“I think she made some improvement. But she can do better.”</p>
Below average = 1	<ul style="list-style-type: none"> <li>❖ Knowledge, skills and abilities of a spoken-voice attribute have not been demonstrated at appropriate levels or unsatisfactory or below expectations</li> <li>❖ Shows little or no progress</li> <li>❖ Significant improvement is needed in one or more areas</li> </ul>	<p>“Weak”, “Tedious”, “Not much improvement”, “Need to work on the area,” “Non-expressive”, “Not confident”</p> <p>“Her voice is tedious and low pitch.”</p> <p>“She was not looking at audience. She was mainly looking at her notes.”</p> <p>“I couldn’t understand what she was saying.”</p>
No evaluative comment = (blank)	No comments made for the attribute	

Following is an example showing how each step of the analysis was conducted for a transcript collected from the voice trainer assessor for an STG student. In the left column of the textbox, a superscript number precedes the part of the voice trainer assessor’s response that is the focus of classification. In the right column, the researcher stated what spoken-voice attributes this response relates to and then allocated a numerical grading to that response. In this way, the relevant parts of

the assessors' comments were identified and subjected to a classification process to both locate and evaluatively describe them. A similar method of analysis was applied for the other 3 groups. (See Appendix M)

*Transcript analysis (pre-voice training video):*

<p><sup>1</sup>She has a clear accent, I should say. I could clearly understand her English. Good consonant and vowel clarity. <sup>2</sup>But apart from that, she was very non-expressive, and <sup>3</sup>not confident and <sup>4</sup>not connected at all with the audience, <sup>5</sup> with low volume and <sup>6</sup> quite a low modulation of pitch range, very low expressiveness. <sup>7</sup>The pace was a little fast. <sup>2</sup>She was sort of skimming over, therefore just lacked emphasis. <sup>8</sup> And her alignment and posture were very bent forward. so she got her hips juttet forward. Seems sort of deflated. <sup>9</sup>Physical tension balance and <sup>10</sup>breath ... I'd say average. But she had really good accent or clear accent, I should say.</p>	<p><sup>1</sup> Articulation: 3  <sup>2</sup> Expressiveness: 1  <sup>3</sup> Confidence:1  <sup>4</sup> Audience connection:1  <sup>5</sup> Voice projection:1  <sup>6</sup> Modulation of pitch range:1  <sup>7</sup> Pace and pause:1  <sup>8</sup> Posture and alignment:1  <sup>9</sup> Physical tension balance:2  <sup>10</sup> Breath: 2</p>
--	--

\*4= Very positive 3= Positive 2= Neutral 1= Negative (blank) = No evaluative comment

The coded data was placed in a matrix of numerical ratings for the presentation of the result. The researcher went through the data for all groups and applied the same procedure to subsequent transcripts in the coding. In the analysis process, the researcher continued noting down her thoughts and observations to identify concepts through the QCA procedure. Table 3.5 presents an example of the matrix showing the data coded based on individual assessors' verbal or written comments on every STG student's 15 spoken-voice attributes. The matrixes compiled for all groups are shown in Appendix N-1, N-2, N-3 and N-4.

Table 3.7 Matrix of numerical ratings of evaluative content (STG).

Spoken voice categories	Spoken voice attributes	Voice Trainer Assessor										Lay Assessor										Interpreting Instructor Assessor									
		student A		Student B		Student C		student D		student E		student A		student B		student C		student D		student E		student A		student B		student C		student D		student E	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Production	Posture & alignment	1	4	1	3	1	3			3	3	1	3	1	3	1	3	1	3	1	3			1	3						
	Physical tension balance	2	4					3	3					1	1	1	1	3	3	2	4					3	3				
	Movement			1	1											1	3	1	1												
	Breathing support	2	3	1	1	3	3			3	3							2	2												
	Release of speech muscles			1	1	3	3							1	1																
Quality	Voice projection	1	3					1	3	1	3					3	3	4	4			1	3	1	3			3	3	1	3
	Resonance & timbre							1	3	1	1					3	3	2	2	3	3			1	3	3	3				
	Pace & pauses	1	2	1	1	3	3	1	3					1	3																
	Speaking clarity	2	4	2	4	3	3			1	3	3	3	2	3	1	1	1	3	1	3			1	3			2	4	1	3
	Modulation of pitch range	1	3	3	3											2	2														
Dynamics	Intonation and tone patterns					2	2					1	3	1	3			2	4	2	4										
	Rhythmic & fluent rendition			1	1			1	4	1	3	1	3	2	1	2	2	2	4	3	4			2	4	2	4	1	4		
	Expressiveness	1	4	1	3	3	4	1	3	1	3	1	3	1	1	1	1	1	3	1	3	1	3	1	3					1	3
	Vocal confidence	1	4	1	3	1	3	3	4	1	4	1	3	1	4	1	3	3	3	3	4	1	4	1	3	3	4	3	4	1	4
	Audience connection	1	4	1	3	1	4	3	3	1	3	1	4	1	1	1	3	1	4	1	4	1	4	1	3	2	4	2	4	1	3

\*4= Very positive 3= Positive 2= Neutral 1= Negative (blank) = No evaluative comment



In the matrix, the 3 spoken-voice categories are presented in the first column of the table and the 15 attributes are listed in the second column beside the categories that they belong to. Each numerical rating placed in a cell indicates an evaluated quality assessed by one of the 3 individual assessors for the 15 spoken-voice attributes in both pre- and post-interpreting assessments: 1 refers to “Negative”, 2 refers to “Neutral”, 3 refers to “Positive” and 4 refers to “Very positive”. Some of the attributes have a blank entry because the assessors did not provide any comment. A description of what has been used as a yardstick for high improvement follows. If a student received 1 in the pre-training test and 2 in the post-training test, it suggests that the student made some progress, although the progress may not have been substantial. If a rating increased from 1 to 3 or 4, it demonstrates that the student made a noteworthy improvement in the spoken-voice attribute evaluated. If an assessor mentioned the same attribute twice or more, the second and subsequent mentions with the same evaluative response are not represented by a double or additional point grading. In a case where the same feature was positively described in a pre- or post-intervention assessment but later the same feature was negatively described in the same assessment, or vice versa, the point gradings were allocated according to the second response. If a student received the same rating in both the pre- and post-training assessments, the student would not be considered to have made any progress. If a student received 3 (Positive) or 4 (Very positive) in both pre- and post-intervention assessments, this indicates that the student remained unchanged in the post-intervention assessment.

Note that the researcher only looked at the substantial positive or negative changes between the two assessments in this study. Some attributes may have received positive comments on both occasions but are not further discussed since there was no change between the pre- and post-assessment. Due to the small number of subjects in each group (N=5), the number of students who improved in an attribute had to score at least 3 (i.e., 60 per cent or more) to be considered to meet the criteria for a high-performance attribute. High improvement is signalled in the grey-shaded cells in Table 3.5. The red-shaded cells, in contrast, show the ratings that were lower in the post-training assessment,

indicating these students regressed in the post-training assessment in the spoken-voice attribute assessed.

In summary, Chapter 3 has explained the research design, data collection and data analysis procedures. As explained, this study has adopted a mixed research method. Quantitative data was collected from assessors' rating scores for the students' pre- and post-intervention interpreting assessments. Qualitative data was collected from multiples sources including assessors' comments and students' interview responses and self-reflective journals. The use of triangulation of the quantitative and qualitative methods has increased the credibility and validity of the results.

The quantitative analysis was produced via the SPSS system. Paired-sample t-tests were used to compare the effects of the interventions within groups and ANCOVA tests were conducted to compare the effects of the interventions between groups. Correlational analysis was employed to explore the relationships between the scores for spoken-voice-related performance and overall interpreting performance, as well as the relationships between the spoken-voice categories and attributes and the overall interpreting performance.

QCA was selected to analyse the qualitative data collected from multiple sources including the students' journals and metacognitive interview and self-evaluation responses. The external assessors' comments were analysed using a deductive coding approach and the magnitude coding method for identifying the spoken-voice attributes that recorded high levels of progress.

## Chapter 4 Presentation of Results

The previous chapter has described the mixed-method approach used to investigate the research propositions empirically and also described the data preparation, collection and analysis processes. This chapter presents the quantitative results in Section 4.1, corresponding to Research Questions 1, 2 and 3. The qualitative results are presented from Section 4.2 to answer Research Questions 4 and 5. Each section concludes with a summary.

### 4.1 Quantitative results

#### 4.1.1 Results for comparison of vocal performance before and after intervention

The first research question relates to the students' vocal performance skills at the post-intervention stage compared to the pre-intervention stage. A paired-sample t-test was conducted to evaluate the impact of the intervention on students' scores on the post-intervention interpreting assessment (see Subsection 3.3.1.1 for the data analysis method). Descriptive data (Table 4.1) was generated to help understand the results (see Appendix I-1, I-2 and I-3 for the complete output). In Table 4.1, "N" represents the number of student participants in each group. The mean scores represent the mean performance for each group of students' pre- and post-intervention assessments, and the standard deviation (SD) is a measure of how dispersed the data is in relation to the mean. The *t*-value represents the size of the difference relative to the variation in the sample data. The *p*-value is used to show the statistically significant differences between the pre- and post-intervention scores within each group. Given the small sample sizes in the project, Eta<sup>2</sup> (Eta squared) is employed to demonstrate the magnitude of the intervention effect when a difference was obtained in the pre- and post-intervention scores. As mentioned in Subsection 3.3.1.1, a difference was significant if the *p*-value was less than 0.05 and Eta<sup>2</sup> >0.14 was taken as a large effect size (Cohen, 1988).

Table 4.1 Descriptive data and within-group comparison (pre- and post-intervention scores from 3 assessors).

Assessor	Group	Pre-intervention		Post-intervention		<i>t</i>	<i>p</i>	Eta <sup>2</sup>
		Mean	SD	Mean	SD			
Voice trainer assessor	STG (N=5)	58.4	9.15	81.6	12.67	3.367	.028	<b>.74</b>
	LTG (N=5)	67.1	12.81	83.6	13.03	2.867	0.046	<b>0.67</b>
	ARG (N=5)	65.4	14.47	68.7	19.02	713	0.515	0.11
	CoG (N=5)	69.3	13.06	72.5	12.90	596	0.583	0.08
Lay assessor	STG (N=5)	68.2	3.01	88.3	5.32	4.449	0.011	<b>0.83</b>
	LTG (N=5)	74.8	27.21	77.9	20.25	0.375	0.726	0.03
	ARG (N=5)	59.2	9.57	81.2	16.57	2.385	0.076	<b>0.59</b>
	CoG (N=5)	82	11.77	71	22.72	-1.033	0.360	<b>0.18</b>
Interpreting instructor assessor	STG (N=5)	60.2	15.58	86.2	10.51	5.556	0.005	<b>0.89</b>
	LTG (N=5)	68.8	16.82	69.2	17.84	0.074	0.944	0.00
	ARG (N=5)	58.4	10.72	71.6	21.05	2.029	0.122	<b>0.50</b>
	CoG (N=5)	59.6	12.54	58.4	16.82	-0.319	0.766	0.02

As shown in Table 4.1, a large magnitude of intervention effect was found within the STG (Eta<sup>2</sup> =0.74, Eta<sup>2</sup>=0.83 and Eta<sup>2</sup>=0.89, respectively, as shown in bold), indicating that the STG students achieved a significantly higher level of vocal performance at the post-intervention stage than the pre-intervention stage based on the 3 assessors' rating scores.

A significant increase for the LTG was found only in the voice trainer assessor's rating scores. The Eta<sup>2</sup> (0.67) indicated a large effect size. In contrast to the voice trainer assessor's scores, the effect size calculated from the lay assessor's (Eta<sup>2</sup>=0.03) and the interpreting instructor assessor's (Eta<sup>2</sup>=0.00) rating scales were small (<14), which indicated that the LTG had a slight improvement

between the pre- and post-interventions from the lay assessor's and the interpreting instructor assessor's perspectives.

A substantial magnitude of the intervention effect within the ARG was found from the lay assessor's and interpreting instructor assessor's scores ( $\text{Eta}^2 = 0.59$  and  $\text{Eta}^2 = 0.50$ , respectively). However, the  $\text{Eta}^2$  statistics from the voice trainer assessor was at the level of ( $\text{Eta}^2 = 0.11$ ), indicating a moderate effect size. These results show that the ARG students' improvement was significant from the lay assessor's and the interpreting instructor assessor's viewpoints, but moderate judged from the voice trainer assessor's rating scores.

In contrast, there was a significant decrease in the CoG according to the lay assessor's scores ( $\text{Eta}^2 = 0.18$ ), a moderate decline ( $\text{Eta}^2 = 0.02$ ) based on the interpreting instructor assessor's scores and an insignificant increase ( $\text{Eta}^2 = 0.08$ ) according to the voice trainer assessor's scores. This indicates that the CoG students did not improve in the post-intervention stage as determined by the ratings of all 3 external assessors.

In summary, the STG was the only intervention group that recorded a significant improvement in the post-intervention stage based on all 3 assessors' scores. The LTG recorded a significant improvement based on the voice trainer assessors' scores, while the ARG noted a significant improvement according to the lay assessor's and the interpreting instructor assessor's scores. On the contrary, no results showed any improvement within the CoG.

#### **4.1.2 Results for multiple group comparison**

The second research question relates to between-group comparison and, in particular, whether those students who received spoken-voice intervention achieved a significantly higher level of vocal performance than those who did not. An ANCOVA was conducted to understand whether there were significant differences in the mean scores on the post-test scores across the 4 groups (refer to Subsection 3.3.1.2 for the analysis method). Table 4.2 displays the descriptive data that summarises the samples and the measures generated from the between-subjects comparison effects (see

Appendix K-1 for the complete output). The  $\text{Eta}^2$  in Table 4.2 shows there were statistically significant differences ( $\text{Eta}^2 > .14$ ) post-intervention (the dependent variable) between the groups (the independent variable), which means an overall significant difference between the groups according to each assessor. The pre-intervention scores were taken as the covariates. In doing so, the differences in pre-intervention scores were statistically removed before the ANCOVA test.

Table 4.2 Tests of between-subjects effects (using scores from 3 assessors).

<b>Assessor</b>	<b>Source</b>	<b>df</b>	<b>F</b>	<b>Sig</b>	<b>Eta<sup>2</sup></b>
Voice trainer assessor	Group	3	2.382	0.110	<b>0.323</b>
Lay assessor	Group	3	1.527	0.248	<b>0.234</b>
Interpreting instructor assessor	Group	3	5.448	0.010	<b>0.521</b>

Table 4.3 shows where the differences among the 4 groups were and how they were different from the 3 assessors' points of view (see Appendix K-1 for detailed pairwise comparisons).

Table 4.3 Cross-group post-hoc LSC pairwise comparisons of post-test results.

Assessor	Group (i)	Group (j)	Mean difference (i - j)	Sig. <sup>a</sup>
Voice trainer assessor	STG	LTG	3.766	0.652
		ARG	17.539*	<b>0.046</b>
		CoG	16.324	<b>0.069</b>
	LTG	STG	-3.766	0.652
		ARG	13.773	0.102
		CoG	12.558	0.133
	ARG	STG	-17.539*	<b>0.046</b>
		LTG	-13.773	0.102
		CoG	-1.215	0.880
	CoG	STG	-16.324	<b>0.069</b>
		LTG	-12.558	0.133
		ARG	1.215	0.880
Lay assessor	STG	LTG	13.400	0.248
		ARG	3.010	0.793
		CoG	23.572	0.061
	LTG	STG	-13.400	0.248
		ARG	-10.390	0.392
		CoG	10.172	0.377
	ARG	STG	-3.010	0.793
		LTG	10.390	0.392
		CoG	20.562	0.124
	CoG	STG	-23.572	<b>0.061</b>
		LTG	-10.172	0.377
		ARG	-20.562	0.124
Interpreting instructor assessor	STG	LTG	24.637*	<b>0.006</b>
		ARG	13.002	0.104
		CoG	27.267*	<b>0.002</b>
	LTG	STG	-24.637*	<b>0.006</b>
		ARG	-11.636	0.157
		CoG	2.630	0.739
	ARG	STG	-13.002	0.104
		LTG	11.636	0.157
		CoG	14.266	0.077
	CoG	STG	-27.267*	<b>0.002</b>
		LTG	-2.630	0.739
		ARG	-14.266	0.077

Based on estimated marginal means

\* The mean difference is significant at round .05 level.

\*a. Adjustment for multiple comparisons: Least Significant Difference.

Based on the voice trainer assessor's rating scores, there was a significant difference between the STG and the ARG ( $p=0.46<0.05$ ) and between the STG and the CoG ( $p=0.069<0.1$ ), indicating that the STG significantly outperformed both the ARG and the CoG. However, the differences between the STG and the LTG, the LTG and the ARG, and the LTG and the CoG were insignificant ( $p=0.652$ ,  $p=0.102$  and  $p=0.133$ , respectively). This suggests that the difference between the STG and the LTG was insignificant from the voice trainer assessor's perspective. However, the STG significantly outperformed the two non-spoken-voice training groups, while the LTG did not.

The lay assessor's rating scores showed a significant difference between the STG and the CoG ( $p=0.061$ ), indicating that the STG significantly outperformed the CoG in the post-intervention test. However, the differences between the STG, the LTG and the ARG were insignificant, suggesting the ARG's post-intervention performance had no significant difference to those of the two spoken-voice training groups.

According to the interpreting instructor assessor's rating scores, there was a statistically significant difference between the STG and the LTG ( $p=0.006<0.05$ ) and between the STG and the CoG ( $p=0.002<0.05$ ) in the post-intervention test. There was also a significant difference between the ARG and the CoG ( $p=0.077>0.1$ ). However, the difference between the STG and the ARG was insignificant ( $p=0.104>0.05$ ) so was the difference between the ARG and the LTG ( $p=0.157>0.05$ ). This shows that the STG significantly outperformed the LTG and the CoG from the interpreting instructor's perspective, although both groups went through spoken-voice training. Even though the ARG students were not spoken-voice trained, their vocal performance did not differ significantly from those of the two voice-training groups and the ARG significantly outperformed the CoG.

In conclusion, from the 3 external assessors' perspectives, the STG attained a significantly higher level of vocal performance than the CoG. Apart from this, the 3 assessors held different views regarding the between-group differences among the 3 intervention groups. Although the STG performed better than the LTG from the standpoint of the voice trainer assessor, the statistical difference between the two was insignificant. However, the STG significantly outperformed the



ARG. While the LTG outperformed both the ARG and the CoG, the differences were statistically insignificant. This suggests that, from the voice trainer assessor's perspective, both voice-training groups made significant improvements. The awareness-raising and non-training groups did not demonstrate significant improvements. The STG achieved the optimal training outcome among the 4 groups.

According to the lay assessor, the STG substantially outperformed the CoG. However, the statistical differences between the STG, the LTG and the ARG were insignificant. This means that the ARG performed almost as well as the other two voice-training groups in the eyes of the lay assessor.

From the interpreting instructor assessor's point of view, the STG substantially outperformed the LTG and the CoG, showing that the interpreting instructor was unsatisfied with the LTG's post-intervention performance. This assessor determined that the ARG significantly outperformed the CoG and that the statistical differences between the STG and the ARG and the LTG were minor, indicating that the interpreting instructor had a favourable opinion of the ARG's improvement.

The results demonstrate that the STG outperformed the other groups and attained the best intervention results. Although the LTG students received spoken-voice training, according to the lay assessor's and the interpreting instructor assessor's rating scores they did not outperform the other groups significantly and did not achieve performance levels as high as those student participants in the STG. Interestingly, from the lay assessor's perspective the ARG performed almost as well as the two spoken-voice groups who received spoken-voice training.

#### **4.1.3 Results of correlational analysis**

The third research question relates to correlations between students' spoken-voice performance and interpreting performance regarding delivery. Pearson bivariate correlational analysis was conducted to examine the relationships between the scores of spoken-voice performances and general interpreting performance based on the total scores collected from the external assessors (see Subsection 3.3.1.3 for the correlational analysis method). Table 4.4 shows the correlation

coefficients between the two (see Appendix L-1 and L-2 for the complete correlational analysis). In the table, “*r*” represents the Pearson correlation coefficient, “*N*” represents the number of cases being analysed (in this scenario, a total number of 60 cases were studied as each external assessor submitted 20 rating scores for 20 students in total) and “*p*” signifies the significant level (given as Sig. (2 tailed)).

Table 4.4 Correlation coefficients between spoken-voice scores and interpreting performance scores (pre and post).

		Interpreting performance related pre-total scores	Interpreting performance related post-total scores
Voice-related pre-total scores	Pearson correlation ( <i>r</i> )	0.627**	
	Sig.(2-tailed) ( <i>p</i> )	0.000	
	N	60	
Voice-related post-total scores	Pearson correlation ( <i>r</i> )		0.776**
	Sig.(2-tailed) ( <i>p</i> )		0.000
	N		60

\*\* Correlation is significant at 0.01 level (2-tailed).

The results demonstrate a strong positive correlation ( $r=0.627$ ,  $N=60$ ,  $p<0.001$ ) between the pre-scores (i.e., spoken-voice-related pre-intervention total scores and interpreting-performance-related pre-intervention total scores). Strong significant correlations ( $r=0.776$ ,  $N=60$ ,  $p<0.001$ ) were also found in the post-scores (i.e., spoken-voice-related post-intervention total scores and interpreting-performance-related post-intervention total scores). The results indicate that the higher the spoken-voice-related scores the students received, the better scores they received in general interpreting performance. In other words, the overall spoken voice-related scores informed students’ general interpreting performance.

Another correlation analysis was conducted to understand which spoken-voice attributes contributed more to the overall interpreting performance in both pre- and post-tests. Table 4.5 shows the correlation coefficients between the scores for spoken-voice attributes and interpreting performance at the pre- and post-intervention stages. A moderate correlation was found in the pre-

scores for 4 spoken-voice attributes and interpreting performance as shown in the table: “speaking clarity” ( $r=0.288, p=0.026<0.05$ ), “pace and pauses” ( $r=0.353, p=0.006<0.05$ ), “modulation of pitch range” ( $r=0.310, p=0.016<0.05$ ) and “posture and alignment” ( $r=0.371, p=0.004<0.05$ ). This suggests a medium strength of the relationships between these attributes and the students’ overall interpreting performance in the pre-intervention test.

Table 4.5 Correlational coefficients between scores for spoken-voice attributes and interpreting performance (pre and post).

Spoken voice category	Spoken-voice attributes	N	Scores for spoken-voice attributes and interpreting performance (pre-intervention)		Scores for spoken-voice attributes and interpreting performance (post-intervention)	
			Pearson correlation	Sig.	Pearson correlation	Sig.
Production	Posture and alignment	60	0.371*	0.004	0.479**	0.000
	Physical tension balance	60	0.483**	0.000	0.646**	0.000
	Movement	60	0.506**	0.000	0.446**	0.000
	Breathing	60	0.485**	0.000	0.563**	0.000
	Release of speech muscles	60	0.414**	0.001	0.406**	0.001
Quality	Voice projection	60	0.572**	0.000	0.648**	0.000
	Resonance and timbre	60	0.476**	0.000	0.520**	0.000
	Pace and pauses	60	0.353*	0.006	0.646**	0.000
	Speaking clarity	60	0.288*	0.026	0.588**	0.000
	Modulation of pitch range	60	0.310*	0.016	0.685**	0.000
Dynamism	Intonation and tone patterns	60	0.329*	0.010	0.626**	0.000
	Rhythmic and fluent rendition	60	0.546**	0.000	0.572**	0.000
	Expressiveness	60	0.534**	0.000	0.712**	0.000
	Professional confidence	60	0.463**	0.000	0.683**	0.000
	Audience connection	60	0.530**	0.000	0.809**	0.000

\*\* . Correlation is a higher level of significance at 0.01 level (2-tailed). \* . Correlation is significant at 0.05 level (2-tailed).

A stronger correlation ( $p<0.01$ ) was noted between the rest of the spoken-voice attributes and interpreting performance in the pre-intervention test, suggesting a strong and positive relationship between these attributes and the overall interpreting performance in the pre-intervention test.

Among the 5 spoken-voice attributes from the category of “spoken voice dynamism”, 3 attributes had strong positive correlations: “expressiveness” ( $r=0.534, p=0.000$ ), “rhythmic and fluent

rendition" ( $r=0.546, p=0.000$ ) and "audience connection" ( $r=0.530, p=0.000$ ), This shows that the "spoken-voice dynamism" category contributed most to the overall interpreting performance at the pre-intervention stage.

Compared with the pre-scores, there was a higher level significant correlation ( $p<0.01$ ) between the scores for the spoken-voice attributes and the interpreting performance in the post-scores. Five out of 15 spoken-voice attributes had large correlation coefficients: "audience connection" ( $r=0.809, p=0.000<0.01$ ), "expressiveness" ( $r=0.712, p=0.000<0.01$ ), "modulation of pitch range" ( $r=0.685, p=0.000<0.01$ ), "vocal confidence" ( $r=0.683, p=0.000<0.01$ ) and "intonation and tone patterns" ( $r=0.626, p=0.000<0.01$ ), demonstrating the greatest correlation with overall interpreting performance among the variable pairs. The result suggests that these spoken-voice attributes may have contributed more to overall interpreting performance scores in the post-intervention test than the other attributes. Table 4.5 also shows that the "spoken-voice dynamism" category contributed the most to the overall interpreting performance compared to the other two categories at the post-intervention stage.

The results indicate more significant correlation coefficients between the spoken-voice attributes and the overall interpreting performance at the post-intervention stage than at the pre-intervention stage. This shows that although the spoken-voice attributes in both pre-and post-intervention assessments informed students' general interpreting performance, the spoken-voice attributes in the post-assessment could better predict the overall interpreting performance.

In conclusion, the results show a significantly positive correlation between the students' spoken-voice performance and their overall interpreting performance in both pre- and post-intervention assessments, indicating that the higher the spoken-voice-related scores the students received, the better scores they received in general interpreting performance. Furthermore, the result show, among the identified speech attributes, attributes in the "spoken voice dynamism" category contributed more to the students' overall interpreting performance than the other two categories (spoken-voice production and quality).

#### 4.1.4 Comparison of vocal categories before and after intervention

In order to provide an answer to Research question 4, which aims to understand what spoken-voice attributes recorded high levels of improvement, another within-group paired t-test was conducted to explore whether there were statistically significant increases for the 3 spoken-voice categories in the post-test scores compared to the pre-test scores (see Subsection 3.3.1.1 for the methodology). The 3 spoken-voice categories were “spoken-voice production”, “spoken-voice quality” and “spoken-voice dynamism” (refer to Section 2.2.4). Descriptive data was generated from the SPSS output to help understand the results based on the rating scores provided by the 3 assessors separately (see Appendix J-1, J-2 and J-3 for the complete output). The statistical results for each group are presented in Table 4.6, Table 4.7 and Table 4.8.

Table 4.6 Descriptive data and within-group comparison for 3 categories (pre- and post-category scores from 3 assessors for STG).

Assessor	Category	Pre-intervention		Post-intervention		t	p	Eta <sup>2</sup>
		Mean	SD	Mean	SD			
Voice trainer assessor	Production	14.3	3.80	19.9	3.89	-3.269	0.031	<b>0.73</b>
	Quality	20.2	4.66	25.8	3.70	-2.682	0.055	<b>0.64</b>
	Dynamism	21.0	4.94	32.2	5.71	-2.608	0.060	<b>0.62</b>
Lay assessor	Production	18.2	2.48	22.8	3.89	-2.674	0.056	<b>0.64</b>
	Quality	21.8	2.58	28.4	3.28	-5.462	0.005	<b>0.88</b>
	Dynamism	24.8	2.68	33.0	4.58	-3.982	0.016	<b>0.80</b>
Interpreting instructor assessor	Production	14.8	3.03	22.0	2.00	-7.060	0.002	<b>0.92</b>
	Quality	20.6	4.82	28.2	4.02	-5.729	0.005	<b>0.89</b>
	Dynamism	21.8	7.15	31.8	4.14	-4.767	0.009	<b>0.85</b>

Table 4.6 shows that within the STG, all 3 spoken-voice categories significantly increased in the post-training assessment according to the rating scores provided by the 3 external assessors. The Eta<sup>2</sup> values shown in bold indicate a large effect size (Eta<sup>2</sup>>0.14).

Table 4.7 Descriptive data and within-group comparison for 3 categories (pre- and post-category scores from voice trainer assessor for LTG).

Assessor	Category	Pre-intervention		Post-intervention		t	p	Eta <sup>2</sup>
		Mean	SD	Mean	SD			
Voice trainer assessor	Production	22.4	6.02	28.6	3.71	-2.258	0.087	<b>0.56</b>
	Quality	16.3	2.28	21.3	2.07	-4.016	0.016	<b>0.80</b>
	Dynamism	25.4	6.73	30.0	7.64	-1.210	0.293	<b>0.30</b>
Lay assessor	Production	29.0	7.89	21.2	7.29	1.258	0.277	<b>0.28</b>
	Quality	23.4	5.63	23.2	6.37	-0.875	0.431	0.16
	Dynamism	19.2	3.64	21.6	4.72	-3.720	0.020	<b>0.77</b>
Interpreting instructor assessor	Production	18.4	6.72	19.8	4.91	-0.269	0.801	0.01
	Quality	21.8	8.73	24.4	6.76	-0.434	0.686	0.04
	Dynamism	25.4	10.79	29.6	7.92	-0.189	0.859	0.00

Table 4.7 shows statistical results for the LTG. Based on the voice trainer assessor's scores, a large magnitude of increase was found in the 3 spoken-voice categories with large effect sizes (Eta<sup>2</sup>=0.56, Eta<sup>2</sup>=0.80 and Eta<sup>2</sup>=0.30, respectively). According to the lay assessor's scores, the "spoken voice quality" and "spoken voice dynamism" were significantly increased with large effect sizes (Eta<sup>2</sup>=0.16 and Eta<sup>2</sup>=0.77, respectively). However, "spoken voice production" had significantly lower mean scores in the post-assessment than the pre-test, indicated by a large effect size (Eta<sup>2</sup>=0.28), suggesting a significant decline in this category according to the lay assessor's scores. In relation to the interpreting instructor assessor's scores, there was no substantial increase in the 3 categories, with a small effect size (Eta<sup>2</sup><0.14).

Table 4.8 Descriptive data and within-group comparison for 3 categories (pre- and post-category scores from 3 assessors for ARG).

Assessor	Category	Pre-intervention		Post-intervention		t	p	Eta <sup>2</sup>
		Mean	SD	Mean	SD			
Voice trainer assessor	Production	19.0	2.91	19.2	4.68	-0.094	0.929	0.00
	Quality	21.0	6.12	22.4	6.84	-0.695	0.525	0.10
	Dynamism	22.2	7.85	23.6	10.31	-0.476	0.659	0.05
Lay assessor	Production	16.0	2.54	20.8	3.76	-2.160	0.097	<b>0.53</b>
	Quality	18.8	2.77	26.2	5.31	-2.785	0.050	<b>0.65</b>
	Dynamism	21.4	4.92	30.2	6.76	-2.011	0.115	<b>0.50</b>
Interpreting instructor assessor	Production	15.2	3.70	18.0	4.74	-1.376	0.241	<b>0.32</b>
	Quality	19.4	3.28	23.0	7.87	-1.462	0.217	<b>0.34</b>
	Dynamism	21.2	4.02	26.6	9.01	-1.654	0.173	<b>0.40</b>

According to Table 4.8, all 3 spoken-voice categories had significantly higher mean scores in the post-test, indicated by large effect sizes shown in bold, based on the lay assessor’s and the interpreting instructor assessor’s rating scores for the ARG. On the contrary, according to the voice trainer assessor no significant increases were found in the “spoken voice production” (Eta<sup>2</sup>=0.00), “spoken voice dynamism” (Eta<sup>2</sup>=0.05) or “spoken voice quality” (Eta<sup>2</sup>=0.10) categories with small or moderate effect sizes.

In summary, according to the 3 assessors’ ratings, the STG improved significantly in all 3 spoken voice categories. Based on the voice trainer assessor’s scores, the LTG was reported to improve substantially in all 3 categories. However, according to the lay assessor’s assessments, there was no significant improvement in any of the LTG categories. On the contrary, the lay assessor reported that the LTG regressed considerably in “spoken voice production”. In the ARG, all categories improved significantly based on the lay assessor and the interpreting instructor assessors’ rating scores. However, according to the voice trainer assessor’s evaluations, none of the spoken voice categories improved considerably. This result is in line with the results shown in Section 4.1.2. The qualitative findings in the following sections will help to explain the statistical findings.

## **4.2 Qualitative results**

This section answers Research question 4 and Research question 5. As shown in Section 4.1.1 and Section 4.1.2, all 3 intervention groups including the STG, the LTG and the ARG recorded significant improvements in the post-intervention assessment compared to the pre-intervention assessment and there were statistically significant increases for the 3 spoken-voice categories in the post-test scores compared to the pre-test scores. Section 4.2.1 presents the qualitative results in order to understand which spoken-voice attributes recorded high levels of improvement in these groups based on the 3 external assessors' comments. Section 4.2.2 explains what factors contributed to the positive changes. Each section concludes with a summary.

### **4.2.1 Spoken-voice attributes that recorded high improvement**

This section discusses the qualitative results for the spoken-voice attributes in the 3 categories that contributed to the significant improvements of the 3 intervention groups in the post-intervention assessment based on the 3 independent assessors' verbal and written comments, following the quantitative results that compared vocal categories before and after the intervention presented in Section 4.1.4. Table 4.9 summarises the highly improved attributes in the 3 groups and shows how recurrent the high ratings from the 3 assessors were (shown in grey in the table) based on the data collected from the matrix of numerical ratings for evaluative content (refer to Section 3.3.3 and Appendix N-1, N-2 and N-3). As is shown in Table 4.9, 8 spoken-voice attributes demonstrated high improvements in the STG. Five of the 8 attributes were in the category of "spoken voice dynamism": "expressiveness", "vocal confidence", "audience connection", "rhythm & fluency" and "intonation and tone patterns", while two attributes in the "spoken voice quality" category – "voice projection" and "speaking clarity" – were highly improved. "Posture and alignment" was the only attribute that improved significantly in the category of "speaking voice production".



Table 4.9 Summary of spoken-voice attributes that recorded high improvements.

Highly improved spoken-voice category	Highly improved spoken voice attributes	Intervention group		
		STG	LTG	ARG
Spoken-voice production	Posture & alignment	✓ Voice trainer ✓ Lay assessor		✓ Interpreting instructor
	Movement			
	Physical tension balance			
	Breathing support		✓ Voice trainer	
	Release of speech muscle tension			
Spoken-voice quality	Voice projection	✓ Voice trainer ✓ Interpreting instructor		
	Resonance & timbre			
	Pace & pauses			✓ Interpreting instructor
	Speaking clarity	✓ Voice trainer ✓ Lay assessor ✓ Interpreting instructor	✓ Voice trainer ✓ Lay assessor	
	Modulation of pitch range		✓ Voice trainer	
Spoken-voice dynamism	Voice confidence	✓ Voice trainer ✓ Lay assessor ✓ Interpreting instructor	✓ Voice trainer ✓ Lay assessor ✓ Interpreting instructor	✓ Voice trainer ✓ Lay assessor ✓ Interpreting instructor
	Intonation & tone	✓ Lay assessor		
	Expressiveness	✓ Voice trainer ✓ Lay assessor ✓ Interpreting instructor		✓ Lay assessor ✓ Interpreting instructor
	Rhythmic & fluent rendition	✓ Lay assessor ✓ Interpreting instructor		
	Audience connection	✓ Voice trainer ✓ Lay assessor ✓ Interpreting instructor		✓ Lay assessor

The LTG had 4 highly improved attributes: “breathing support”, “speaking clarity”, “modulation of pitch range” and “vocal confidence”, as reported mostly by the voice trainer assessor. Table 4.9 demonstrates that 5 attributes improved significantly in the ARG: “posture & alignment”, “pace & pauses”, “expressiveness”, “vocal confidence” and “audience connection”. Three of these 5 were in the category of “spoken voice dynamism”, as reported mostly by the lay assessor and the interpreting instructor assessor.

The following subsections present the 3 assessors’ comments on each of these highly improved spoken-voice attributes in the order of their occurrence in the assessors’ comments.

#### **4.2.1.1 Vocal confidence**

According to the external assessors’ evaluation, all 3 intervention groups improved substantially in “vocal confidence” in the post-intervention assessment. In the pre-intervention videos, the students’ low confidence was manifested in their restricted posture and body movement, and low volume in interpreting delivery. For example, the lay assessor noticed that when the LTG students were not sure about something (e.g., vocabulary), “their voice became very quiet, and they tended to pull their shoulders up or bend their tarsal backwards, but when they were confident, the volume could be very loud which caused an impression of inconsistent and unsteady speed of delivery”. The interpreting instructor assessor noted that the ARG students demonstrated a lack of vocal confidence as they were focusing overly on transferring messages linguistically and ignoring their voice presentation.

In the post-intervention videos, over 60 per cent of the STG students (i.e., 3 out of 5), 80 per cent of the LTG students (i.e., 4 out of 5) and 60 per cent of ARG students showed obvious improvement in vocal confidence, which was referred by the interpreting instructor assessor as “a good stage manner”, “passionate” and “persuasive” by the lay assessor and “much poised and impressive vocal presence” by the voice trainer assessor. This suggests that vocal confidence, one of the spoken-voice dynamic features, was interconnected with the features of spoken-voice production, such as voice projection, body movement, and posture and alignment, indicating that

when the students improved these attributes, they sounded much more confident. In turn, when their confidence level improved, these features were improved as well. Overall, the outcome indicates that the intervention had a beneficial effect on vocal confidence.

#### **4.2.1.2 Audience connection and expressiveness**

Apart from vocal confidence, “audience connection” and “expressiveness” were another two attributes in the category of “spoken-voice dynamism” that improved substantially in both the STG and the ARG. In the pre-intervention assessment, the assessors noted that many students were stuck in their notes, overly focusing on interpreting the SL speech and lacked audience awareness and expression in their delivery, which “reduced the good impression of performance” (the lay assessor). Other dynamic qualities, such as emphatic and intonation patterns, were missing. As such, their delivery was “more like a monotone”, sounded “tedious” (the interpreting instructor assessor), “lacked emphasis” (the lay assessor) and were “quite stiff” and “non-expressive” (the voice trainer assessor).

In the post-training videos, the students manifested “better eye contact” and “good rapport with their audience” (the voice trainer assessor). They sounded “convincing” with “a really great emphasis” and showed “more meaning and expression in their voice” (the lay assessor), “not as tedious as the first time” and “more expressive and dynamic” (the interpreting instructor assessor).

Taking Student M (from the ARG) for example, the lay assessor commented that “she held her notebook too high, and it seemed like she was clutching to the notebook whereas the notebook should be only something to fall back on” and “sounds very unsure of what she was saying” in the pre-intervention video. In the post-intervention assessment, the lay assessor spoke highly of her improvement in these attributes:

She sounded and looked more confident in video 2. **More connection with the audience.** She seemed to know what she was talking about. She **looked up and looked at the audience rather than looking the sideways.** You could really look at her and listen to her speaking and think that comes from her experience rather than something she was just reading. It seemed to

mean more to her. And **has more meaning for me** as well [bold added by the author].

Notably, the lay assessor noticed that the student “knew what she was talking about” and consciously monitored her eye contact with the audience, which demonstrated the student showed not only high confidence but also solid metacognitive awareness, allowing her to monitor her performance and connection with the audience appropriately and this left a good impression with the lay assessor.

The results suggest that the ARG students demonstrated significant self-monitoring ability following the voice-training and awareness-raising intervention sessions, which may have helped these students increase their expressiveness and connection with the audience. Additionally, the results indicate that “expressiveness” and “audience connection” had significant impacts on how an audience perceived the students’ vocal performance during interpretation assignments. According to the assessors’ comments, the LTG students did not display substantial growth in “audience connection” or “expressiveness,” implying that the training method may have had little effect on these students’ spoken-voice dynamism.

#### **4.2.1.3 Speaking clarity**

When evaluating the pre-intervention video recordings, the lay assessor and the interpreting instructor assessor commented that the students’ pronunciation across all groups was inconsistent and difficult to understand. The voice trainer assessor noticed that a heavy Chinese accent impeded their understanding, such as a prolonged vowel sound, for example “will” as “weal”, a lack of vowel and consonant clarity, for example muffled “th” sound in the word “the”, and a heavy accent affected by the mother tongue, for example a retroflex consonant in “r” sound at the end of words like “year” and “Australia”.

Both the STG and the LTG students improved substantially in “speaking clarity” in the post-training assessment based on the assessors’ comments. The interpreting instructor assessor noticed that the STG students attained “marked improvement” in pronunciation and enunciated English

“more precisely than before”. The lay assessor praised the STG students’ articulation as “rather clearer than the first video” except when they mispronounced some words. The voice trainer assessor said the STG students had “better consonant and vowel clarity”. The voice trainer assessor commented that the LTG students’ pronunciations of vowels and consonants were much clearer although they “could work a little bit more on accent enhancement”.

According to the lay assessor, Student I (from the LTG) ‘s imprecise pronunciation hampered her comprehension throughout the pre-intervention assessment. However, the lay assessor lauded Student I’s notable improvement in the post-intervention assessment, remarking that “it was as if a different person was speaking and interpreting. Her articulation was far more understandable.” However, the lay assessor observed that while this student’s articulation had improved, they “still do not sound that fantastic” due to their heavy Chinese accent.

The findings indicate that the spoken-voice training had discernible influences on students’ speaking clarity; however, further training might be necessary for students who must interpret English as their B language. The ARG did not show significant improvement in speaking clarity, indicating that raising awareness alone may not be sufficient to improve this attribute.

#### **4.2.1.4 Posture and alignment**

In the pre-training assessment, the assessors pointed out that the students’ poor posture and alignment, caused by their over-concentration on reading from their interpreting notes, negatively affected the students’ vocal quality. Specifically, the voice trainer assessor noted that some STG students were “bending forward ... hips jutted forward ... sort of deflated” while interpreting. The lay assessor noted that some STG students were “standing and sort of looking down, trying to find their place in their notes”. The interpreting instructor assessor commented that some ARG students “lowered or tilted their head”, their body “leaned to one side” and their “shoulders were not balanced, with one side higher than the other”. Taking Student O as an example, the interpreting instructor assessor noted that she “had a muffled volume because her head was down while she focused on reading through the notes and was not using her whole body to deliver”.

In the post-training assessment, the assessors reported the STG and ARG students had improved substantially in “posture and alignment”. The voice trainer assessor commented that the STG students “stood upright with their shoulders back” and “their spines were aligned”. The lay assessor said that “they held their notes properly”. The interpreting instructor assessor noted that the ARG students were “standing with their head back and reading notes without interfering with their interpreting performance” and looked “dignified and much more professional”. Although the voice trainer assessor did not report significant improvement for the ARG group, she mentioned that the ARG students demonstrated “a better release in their alignment and physical tension balance”. The interpreting instructor assessor praised Student O, saying that “her head was up, and her volume was at an excellent level”.

As above, posture and alignment, which is a critical component of speech behaviour, may have influenced the students’ voice projection. Students might benefit from improved posture and alignment by reducing tension and increasing their vocal confidence. This outcome shows that with adequate intervention, through either consistent voice training or awareness-raising intervention, students’ bad posture and alignment habits could be changed, hence improving their overall voice quality.

#### **4.2.1.5 Voice projection**

According to the voice trainer assessor and the interpreting instructor assessor, the STG students made clear improvement in “voice projection”. In the pre-training assessment, the voice trainer assessor identified Student D’s low vocal volume and “unnatural loudness” being due to “pushing up volume hard”. The interpreting instructor assessor identified that poor posture in Student E was due to her “lowering the head and reading from notes”. In the post-training assessment, the assessors viewed the students’ improvement from different angles, but both assessors pointed out that the students’ obvious improvement in speaking volume was related to their better posture and alignment. For example, the voice trainer assessor commented that student D “improved a lot with her alignment, posture and appropriate loudness”. The interpreting instructor praised Student E’s

better voice projection, which “made her voice more audible” when she lifted her head to speak.

The lay assessor made few observations on the students’ progress in vocal projection.

The findings indicate that assisting students to correct their posture and alignment during interpretation delivery, as well as modifying undesirable vocal projection habits, may have helped them enhance their voice projection. However, no substantial improvement in voice projection was reported in the LTG or the ARG.

#### **4.2.1.6 Rhythm and fluency**

Both the lay assessor and the interpreting instructor assessor expressed satisfaction with the significant improvement in the STG’s “rhythmic and fluent rendition” and commented that these students’ fluency and rhythm could be related to their improved coping skills and confidence in dealing with difficulties encountered in interpreting delivery. They observed that “staccato and fillers were reduced greatly” (the lay assessor) in some students’ delivery and the interpreting text sounded “more rhythmic and smoother with fewer unnecessary pauses and even pace” (the interpreting instructor).

Student D is an example who made conspicuous improvement in this attribute. In the pre-training assessment, she concentrated on reading words rather than rendering sentences. As a result, there were many fillers and staccatos in her interpretation and “the interpreting text was broken up in terms of pauses and pace” (the lay assessor). In the post-training assessment, the lay assessor noted that Student D sounded “a lot more rhythmic and smoother than before”. The interpreting instructor assessor noted that her improvement was partly associated with her “better competence to cope with some emerged problems in interpreting delivery”. This shows that from the lay assessor’s and the interpreting instructor’s points of view, students could improve their English rhythm and fluency through appropriate spoken-voice training, consistent self-practice and practice in controlled settings.

Despite the above, from the voice trainer assessor's perspective students in all groups did not improve much in this attribute. Instead, the voice trainer assessor commented that the students encountered many barriers to achieving English rhythm and fluency when interpreting into their B language. The voice trainer assessor gave an example as below showing that the students tended to mix up the Chinese rhythm with the English rhythm at both word and sentence level, which adversely affected their delivery:

Some students **gave English syllables almost the same strength, length and pitch as if they were speaking Mandarin**. They **randomly paused** at wherever the problematic points, especially when they were **reading mechanically** from their note or **tried to formulate a long sentence**. As such, they sounded **unnatural and not rhythmic, very choppy and unclear**.

This result reflects the different perspectives regarding rhythm and fluency between the voice trainer assessor and the other two assessors. The dissatisfaction with their performance of the voice trainer assessor, an L1 user of English, points to the importance of students changing their rhythm habits. This suggests that it is essential for English B-language interpreting students to consistently acquire and practise English rhythm and fluency patterns. The LTG and ARG students did not demonstrate conspicuous improvement in this attribute, indicating that the intervention for these two groups may have been less effective in improving the students' rhythm and fluency.

#### **4.2.1.7 Intonation and tone**

As shown in Table 4.7, the lay assessor was the only one who commented that the STG students made substantial progress in "intonation and tone" patterns. In the pre-test assessment, this assessor complained that these students' flat intonation hindered her understanding. However, when watching the post-intervention video she stated that the students "put intonation into the phrases correctly" and that she could understand the message, which was better rendered than when she watched the first video. However, the voice trainer assessor and the interpreting instructor assessor



maintained that the students' intonation and tone pattern remained unchanged. In particular, the voice trainer assessor expressed her concerns about students' monotonous tone in general:

I've noticed with these guys ... **have more of a monotone** ... I can see that they have got **less pitch rise and fall**.

Although the students in the training groups had been guided to explore their pitch range and encouraged to do intonation exercises such as shadowing and imitating native English speakers' intonation, little improvement was observed by the students themselves. In her journal, Student A expressed that she found it difficult to improve her intonation although she had made a lot of effort to change it:

It was **challenging for me as English is my B language** when I **unconsciously used intonation patterns from my mother tongue**.

This result shows that although the lay assessor claimed that she could understand the interpreting context because of the students' improved intonation and tone patterns, there was still room for further training in intonation and tone patterns, particularly for interpreters working in their B language. As in Subsection 4.2.1.6, neither the LTG nor the ARG students showed substantial progress in this attribute, showing that the intervention for these two groups may have been less effective at improving students' intonation and tone pattern.

#### **4.2.1.8 Breathing support and modulation of pitch range**

As illustrated in Table 4.7, the voice trainer assessor reported that "breathing support" and "pitch range modulation" both improved significantly in the LTG. The voice trainer assessor noted that poor breathing management was a common problem for the students in all groups including the LTG, which manifested short breath (such as gasping) while interpreting due to nerves, stress and tension or feeling awkward (not released) in their standing position. As a result, their body became tense, their voice sounded "quite breathy" and their delivery pace was out of control. The voice trainer assessor also mentioned this was: "Because they didn't have enough breath for the thought and that made them less easy to listen to".

The voice trainer assessor's responses in relation to the LTG group in the post-training assessment were that they showed a high level of improvement, demonstrating "better breathing support, release and reflexivity", "better breathing management" and "better breathing release" than in the pre-training video. This suggests that breathing exercises could be beneficial for students to release tension in their physical body, which is essential for the entire vocal performance in an interpreting presentation.

Modulation of pitch range is another attribute that was reported to be "much improved" in the LTG. The voice trainer assessor commented that the LTG students did not present much pitch variety in the pre-intervention video and their interpreting delivery sounded "kind of hard to listen to" and "just not very interesting". But the voice trainer assessor stated that these students had improved "well enough" in their pitch range in the post-training video and that "they had far more range and interest in their voice than it was before." Taking Student F as an example, the voice trainer assessor noted that this student showed a habit of speaking with a "high rising terminal" in the pre-training video – her voice went higher at the end of sentences, which made her interpreting lose authority. In the post-training video, the assessor found that she had a consistent pitch range improvement and commented, "There is far less going up at the end of the line in her pitch, which adds to her credibility".

The above shows that the training of breathing skills and pitch modulation was effective for the LTG students, which contributed to their high improvement in the post-intervention assessment.

The voice trainer assessor commented that the STG students also improved in breathing support and pitch range modulation, although not as significantly as the LTG. However, the students in the ARG remained unchanged or even regressed in these attributes. The following is an excerpt from the assessor's comment on Student M's (from the ARG) post-intervention evaluation:

She suffers from **a lot of nerves** in that context. Really just **the nervousness made her breath too short**, which **broke up the speaking into two repetitive rhythm** and made the expressivity too **monotonous** ... And she

had very **monotonous tone without any pitch change**. So that was the thing she needs to work on.

The assessor's comments show that the student appeared unable to control her excessive tension, which had a noticeable effect on her breathing and pitch modulation. As advised by the voice trainer assessor, the student should improve these two attributes, indicating that the awareness-raising intervention might be insufficient.

#### **4.2.1.9 Pace and pauses**

According to the interpreting instructor assessor, the ARG students achieved a high level of improvement in “pace and pauses”, an attribute in the category of “spoken voice quality”, in the post-intervention assessment. The interpreting instructor assessor critiqued some ARG students with comments that they spoke either too slowly or too fast or with inconsistent speed in the pre-intervention assessment. Others also delivered their interpretation with strange pauses in the middle of the delivery. However, most ARG students (i.e., 3 out of 5) improved in this attribute, demonstrating significantly more consistent delivery and a steadier pace with fewer inappropriate pauses between and within sentences, according to the assessor's post-intervention assessment. For instance, the interpreting instructor assessor commended Student L's progress:

**I particularly like her speed of speaking, very steady. No frequent pauses** in between the sentences like in the first video.

This demonstrates that without vocal training activities, ARG students might have improved their “pace and pauses” with increased awareness of this attribute. The STG and the LTG also improved in this attribute according to the voice trainer assessor's comments. However, their improvement was not as significant.

In summary, Section 4.2.1 has presented the spoken-voice attributes that recorded high levels of improvement in each intervention group. A total of 11 spoken-voice attributes improved significantly, appearing either in one group or in combination with other groups across the 3 intervention groups. These attributes were: “posture and alignment”, “breathing support”, “voice

projection”, “pace and pauses”, “speaking clarity”, “modulation of pitch range”, “voice confidence”, “intonation and tone pattern”, “expressiveness”, “rhythmic and fluent rendition” and “audience connection”. The results indicate that the intervention method effectively improved these attributes for interpreting students. Among the improved attributes, significant improvement was recorded in “vocal confidence” in all 3 intervention groups, indicating that both the voice-training and awareness-raising interventions had a beneficial effect on “vocal confidence”.

The STG had the highest number of highly improved attributes across all groups. Apart from “vocal confidence,” STG students improved in a variety of other vocal features, including all 4 attributes in “spoken voice dynamism” as well as “speaking clarity,” “posture alignment” and “voice projection,” as identified by all 3 external assessors. This corroborates the quantitative results by identifying which attributes contributed to the STG’s outstanding performance.

According to the assessors, the LTG students demonstrated significant improvements in “speaking clarity”, “breathing support” and “pitch range modulation”, mainly from the perspective of the voice trainer assessor. The results indicate that the spoken-voice training significantly benefited the LTG students in enhancing their skills in “spoken voice production” and “spoken voice quality”. However, with the exception of “vocal confidence,” the LTG did not improve in any other component of “spoken voice dynamism,” which may explain why the lay assessor and the interpreting instructor assessor did not consider the LTG’s improvement substantial.

The ARG students were reported to have improved 3 aspects of their “vocal dynamism”, namely, “vocal confidence”, “audience connection” and “expressiveness”, as well as “pace and pauses” and “posture and alignment”. This indicates that, even in the absence of formal voice training, a self-regulated learning strategy aimed at enhancing spoken-voice awareness could effectively improve these attributes, most notably in “spoken voice dynamism”. This result corroborates the quantitative findings of the ARG’s significant improvement in rating scores from the lay assessor and the interpreting instructor assessor. However, the ARG students improved less in the attributes of the skills in “spoken-voice production” and “spoken-voice quality”, indicating that raising awareness

alone may not be sufficient to improve the fundamental spoken-voice techniques, for example, “speaking clarity” and “modulation of pitch range”, which may explain why the voice training assessor did not consider the ARG students’ improvement significant.

Apart from the above results, the external assessors’ comments indicate that interpreting students working in their B language still need to enhance their prosodic features such as “speaking clarity”, “rhythm and fluency” and “intonation and tone patterns”, despite the significant progress in the STG.

A few spoken-voice attributes remained unchanged in all groups, mainly in “spoken-voice production”, including “physical tension balance”, “movement”, “resonance and timbre” and “speech muscles tension release”, suggesting that the intervention may not have been as effective for these attributes as for the other attributes. Excessive tension in particular remained a significant impediment to all students’ improvement in vocal performance. According to the assessors’ comments, the CoG students regressed in a number of spoken-voice attributes, mostly in “spoken-voice dynamism” as a result of excessive tension and mechanically reading their notes (see Appendix N-4), demonstrating the critical nature of consistent intervention for excessive tension in spoken-voice intervention and interpretation delivery.

#### **4.2.2 Factors that contributed to high performance**

Section 4.2.1 has documented the spoken-voice attributes that recorded high levels of improvement from the 3 external assessors’ perspectives. This section answers the 5th research question, which explores the data samples elicited from the students’ reflective journals and interview responses in an attempt to find the factors that contributed to the significant improvements in the 3 intervention groups. Eight factors were identified and are presented in the following subsections.

##### **4.2.2.1 Kinaesthetic awareness**

The STG and the LTG students’ accounts of bodywork practice at the voice-training workshops show that they developed kinaesthetic awareness through a series of voice-training exercises, for

example, vocal warm-up exercises, breathing exercises and other vocal work, which contributed to the great improvement in their skills of voice production, particularly “posture and alignment” and “breathing techniques” as evaluated by the external assessors (refer to Subsections 4.2.1.4 and 4.2.1.8). Selected journal excerpts from the student participants in the two voice-training groups are presented below.

Student G and Student I from the LTG stated that “diaphragmatic breathing exercises” (exercise 7 in Appendix D-3) assisted them in overcoming shallow breathing induced by mental and physical tensions:

During this [training] process, when I was lying down and doing the inhalation and exhalation exercise, I **felt relaxed** and it seemed that I had **eliminated all the tensions.** (Student G)

I think **taking a deep breath** could also act as a useful pause button when the wave of emotion is coming ... **This method should be practised frequently; otherwise, it is so easy to get trapped in the mood again.** (Student I)

This shows that the students were aware of the relationship between the physical exercises and their body sensations, tension and mood. Student I also showed an understanding that the workouts must be performed consistently in order to maintain the effect.

Student C (from the STG) stated that “body mapping” exercise (exercise 2) had changed her awareness of her perception of body structure:

We were asked to draw our own spine and our partner’s on a piece of paper. Both my partner and I found something **interesting** after we shared our drawings. We **thought** our own spines are as stiff as a stick, but **the truth is** that our spines are curve shaped.

Student B (from the STG) indicated that the “semi-supine” exercise (exercise 5) helped improve her posture, breathing and tension balance:

While doing these exercises, I felt **as if I was in a yoga class, relaxing and breathing well**. These vocal warm-ups are **helpful** in both improving posture, breathing and balancing stress.

This demonstrates a significant improvement in these students' kinaesthetic awareness. In contrast, no students in the ARG displayed significant awareness of their voice production or body movement, which might be attributable to a lack of voice training which can raise students' kinaesthetic awareness.

#### **4.2.2.2 Vocal care awareness**

The students in the 3 intervention groups demonstrated improved awareness of vocal care. The journal excerpts from the students exemplify their improved knowledge of voice management and their commitment to improving their vocal use. For example, Student B (from the STG) reflected:

**I learnt** that managing your voice or not to abuse your voice is utterly essential for interpreters. This is a **new issue** that has just unveiled and become **appealing to me**. It **raised my awareness** of the need to protect my vocal cords on a daily basis and also to prepare my vocal cords in order to deliver correctly and perform effectively. Throughout the workshops, **I have learnt the possible results** of not using voice properly, basic knowledge of the vocal system, the factors that hamper good vocal delivery and also methods of preserving and warming up our voices. **I have detected examples of me misusing my voice** from personal experiences of daily practising and in-class presentations. Relating to what I have learnt, **I've also deliberately tried to make improvements**.

Student B's journal reflection detailed her path towards developing a new perspective on vocal care. Following the training, she recognised the importance of voice care and was determined to improve through workshop activities and self-practice.

Similarly, Student J (from the LTG) described how she came to realise the need for voice management in her profession after the training class and how important it is to care for the voice in interpreting:

The most important thing that **I learnt from this class** is that we need to pay more attention to our voice as an interpreter. **Before this, I had no idea** why I need to care about my voice, you know, I am not the one who does the speech, I am just an interpreter. But **now I realise** that as an interpreter, I should not only just care about how to translate, but also make sure that my information is clear and understood by my clients through my voice. A clear and healthy voice is particularly important in this sense.

Student J's journal reflected her increased awareness of voice management and her appreciation for how critical a healthy voice is to the delivery of an interpreter. Nevertheless, Student J did not indicate how she had developed her vocal management, such as warm-up exercises or others, or her self-practice in maintaining a healthy voice.

The ARG students also demonstrated an improved awareness of vocal care over the two interview sessions. In the initial interviews, few ARG students seemed aware that they could suffer "vocal fatigue" after doing a long interpreting practice. For example, Student N stated, "The only time I really noticed my voice was when I had a cold or when my throat was sore." Other students indicated that they did not know how to control their vocal volume when they became nervous and as such they often pushed their voice so as to be heard by their audience.

In the subsequent interviews, some ARG students had realised that they needed to make "conscious effort" (Student K) to manage voice hygiene. For example, Student L had realised that pushing volume could harm her voice and begun looking for a better technique to project her voice:

**I learnt** that forcing voice projection is no good for voice health and neither for voice quality. So **I try to** use the diaphragm to speak. **It works. My classmates said I sound much louder and more energetic than before.**

Student L had realised the importance of not forcefully pushing her voice up and attempted to improve through self-practice. In addition, she was seeking feedback from her peers due to being unable to get feedback from the workshop training. This shows that this student turned her awareness into self-motivation and self-directed learning, which was beneficial for her improvement.



The above shows that the students increased their awareness of vocal care and management following the voice-training intervention. Some students in the STG and the ARG were motivated to turn their awareness into self-practice, which was beneficial for improving their voice quality and voice projection.

#### 4.2.2.3 Prosodic feature awareness

Students' journal reflections revealed that the instructor-led training activities in the two voice-training groups focusing on the prosodic features considerably improved the students' awareness and skills in intonation, rhythm, pitch, articulation and volume, which contributed to the overall quality of interpreting delivery in the post-training assessment. For instance, in the following journal extract Student A (from the STG) demonstrated self-awareness of a strong accent when speaking English following workshop activities and was encouraged to engage in self-practice aimed at enhancing English-speaking clarity:

After watching my interpreting video, **I realise** that I have such a strong Chinese accent, so **I have kept practising** the pronunciation exercises taught to us by the voice coach every day, such as English vowels and consonants, tongue wagging and lip stretching.

Moreover, the STG students reported that the “spoken voice production seminar” and the “voice projection exercise” (exercise 10) assisted them to overcome the voice projection blockage caused by speaking a second language and to change their poor vocal projection habit. The students showed knowledge of how the articulation organs work and how to “project” their voices instead of “pressing” their volume up in a forced manner. Student E, for example, reflected on the “resonance exercise” (exercise 9) in which she and other students were guided to project sound by exercising the resonator muscles (neck, mouth and throat), and demonstrated mastery of the technique of projecting her voice from the deep chest rather than the mouth or throat through regular repetition:

**I learnt to employ resonance cavities to project my voice** instead of mechanically pushing up my vocal cord and **felt chest vibration through repeated practice** of gentle humming to activate the resonator cavities including head, nasal cavity, mouth, larynx and chest.

The LTG students mentioned that the guided pitch exercise work in breathing, pitch melody and resonance was helpful in improving their pitch modulation and intonation skills. However, Student H noticed that she still misused intonation patterns subconsciously:

**I tried to find appropriate intonation and emphatic patterns** that highlight the key elements and convey the tone of the piece helpfully and convincingly. **But it was challenging for me as English is my B language** when I **unconsciously used intonation patterns from my mother tongue.**

This echoed the external examiners' comments that the prosodic features were particularly difficult for interpreters who spoke English as a second language and that the students, despite their development, needed to improve their prosodic feature skills through consistent practice. As shown, although the LTG students showed enthusiasm in workshop activities, they mentioned very little about self-practice. This could explain why the LTG students did not demonstrate substantial improvement in prosodic features as measured by the assessors' rating scores and comments. This demonstrates that enhancing prosodic features requires constant practice both inside and outside workshops.

The ARG students were exposed to a list of spoken-voice attributes and evaluated their own performance during the interview sessions and self-evaluation activities. A comparison of their initial and follow-up responses revealed a significant increase in their awareness of prosodic features. When asked which spoken-voice features they thought were important for interpreters (see Question 3 in Appendix C-1), 80 per cent of the students (i.e., 4 out of 5) responded in the initial interviews that a clear and loud sound was the most essential vocal quality for an interpreter to be understood by their audience. However, during the subsequent interviews the students found that, in addition to loudness, many other spoken-voice attributes are important for interpreters. Following that, their replies expanded to include tone and intonation, pitch modulation, good tempo and pauses, and articulation, among other prosodic features. The students also reflected on the spoken-voice attributes that they thought they needed to improve. For example, Student K stated:

It's eye-opening to know **there are so many voice skills to learn**. I think **I have a lot to improve, especially to control my speaking speed**. And I think I **also need to control my breath**. I **will voice record or video record my practice and review later** for **practising these skills with my friends**.

This indicates that MRI raised this group of students' awareness of prosodic features, which might have triggered self-directed learning and further improved their spoken-voice skills relating to these prosodic attributes.

As shown above, increased awareness of prosodic features, either directly through voice training or indirectly from awareness-raising intervention, contributed to a significant improvement in the quality of the spoken voice. However, it was still challenging for the students to gain mastery over these prosodic features through short-term training or unguided sessions. In particular, intonation and tone patterns, rhythm and fluency, and pitch modulation are specialised areas of spoken-voice training and deserve considerable attention and constant effort from the students' side, especially for Mandarin–English interpreters.

#### **4.2.2.4 Audience awareness**

The STG and the ARG students demonstrated improved audience awareness, which may have contributed to their significant improvement in “spoken voice dynamism”, including better “vocal confidence”, “audience connection” and “expressiveness”. For example, the journal of Student E (from the STG) showed that she developed an inclusive awareness (refer to Subsection 2.2.3.3), that is, connecting emotion, body movement and spoken-voice production when interacting with the audience to better fulfil speakers' communicational needs and purposes. As she became more dynamic in interpreting delivery, she turned out to have a better connection with the audience as well and presented herself as a confident interpreter. Her reflection is as follows:

**I used to speak in a low key with a flat tone** which sometimes hinders communication between myself and other people ... In the workshop, **I received feedback from the tutor** that I have to deliver the meaning of the words through my voice as best as I can because this will help me to make my opinion more persuasive ... So **I realised that the register of the tone is**

**vital** because it represents the attitude and can reveal the personality of the speaker in a snapshot. **I think it is necessary for me to be more aware of my voice**, including uttering my sound in a more positive way and selecting the words more wisely to express my thoughts. **By doing so, the world around me could be changed in a better way since the quality of communication has improved.**

Student E demonstrated in her journal that she was aware of her weakness in audience connection following the feedback from the tutor. She realised that the register of her tone and sound quality were essential for her audience communication. It is worth pointing out that Student E was one of the two students (the other was Student A) who were identified by all 3 assessors as having made a remarkable improvement in expressiveness in the post-training video. For example, the lay assessor made the following comment about Student E, expressing satisfaction with her progress:

**In the first video, she was not expressive** because she was trying to understand herself. I guess as well because it was **softer and less confident**, her delivery and phrasing perhaps left more to be desired. **In the second video, she was definitely more dynamic than she was before. She sounded very expressive and more enjoyable.**

When Student O (from the ARG) was asked in the initial interview whether she thought she communicated well with her interpreting audience (her classmate or instructor) using spoken-voice skills (see Question 8), her interview response below shows that she may have paid too much attention to linguistic correctness and was unaware of the vocal skills related to audience connection:

**Not that I am aware of.** I think I **paid more attention to accuracy** than voice skills. I tended to **look at my notes a lot**. I think **as long as the audience could understand me, that should be ok.**

In the follow-up interview, Student O replied that she had started to remind herself to connect with the audience by using eye contact, reading less from her notes and being mindful of the delivery pace:

**I do remember to tell myself** to move my eyes from the notes a bit and have eye contact with the audience and **try to lower the pitch and use a bit of tone change to increase vocal variety for the audience.**

Despite their improved self-monitoring skills and self-regulation, the ARG students also reported that they sometimes forgot to make audience connections during their interpreting. Student L stated:

**I try to look at my audience and smile at them. I did make efforts ... I pay attention to these skills while I practise on my own. But when it comes to the time that I need to stand in front of people, I tend to forget about all these because of nerves.**

This demonstrates that a high-level awareness of spoken-voice attributes can significantly improve spoken-voice production skills. However, in the interpreting setting this awareness may need to be further substantiated in actual practice to reach automaticity.

According to the assessors' comments reported in Subsection 4.2.1.2, the LTG students did not demonstrate good improvement in "audience connection." Correspondingly, in their journals the LTG students demonstrated a low level of awareness of audience connection. This indicates that, for some reasons, the LTG students' audience awareness and other dynamic aspects were not properly enhanced by the long interval voice-training regimen.

#### **4.2.2.5 Changed perceptions of spoken-voice training for interpreters**

The students in all 3 intervention groups demonstrated shifts in their perceptions of spoken-voice training for interpreters. Prior to the instruction, for example, Student E (from the STG) believed that "voice is all about the neck and chest" and was unaware that the spoken voice could be trained to perform more effectively. Following the instruction her outlook transformed, as she stated in her journal reflection:

**I was wrong. Voice can be trained. We are not only using vocal folds to produce "voice", but also it relates to other parts of bodies, such as resonance in the head, chest, nose, back of neck etc. Understanding how to manage voice is important, because if we don't take care or control, it could be damaged and lead to severe consequences, especially as an interpreter.**

Student G (from the LTG) had thought that the spoken-voice training would be the same as singing training and was irrelevant to interpreting training. Her notion evolved over time and she recorded this in her journal:

There is some **similarity** between these two **but not necessarily the same** ... As an interpreter, **I understand how important voice is to me**, yet **it never came to my mind that voice training and management would be an option for me to improve and better manage my interpreting performance.**

The ARG students also demonstrated shifts in their perspective on the importance of speaking-voice management in their future interpreter careers. When asked, “How important to you is your voice as a means of expressing yourself and in the field of interpreting?” (see Question 1), the students’ initial responses including “never thought about it before” (Student K), “quite important” (Student L) and “not at all important” (Student M) reveal their lack of awareness of spoken-voice management. However, further interviews revealed that the students’ perceptions had changed. For instance, as stated by Student K:

**Very very important.** Speed, pronunciation, accent, confidence ... **all important** for me to take and do the job in the future. To present an idea to an audience both in English and Chinese. **Crucial** for me to have a nice and pleasant voice while speaking.

Questions 13 and 14 were designed to elicit the students’ opinions on whether spoken-voice training should be provided to interpreting students and whether they would be willing to participate if voice training were included in the interpreting course. The students appeared unsure in the initial interviews whether voice training should be included in the interpreting course. Student O thought she would be too preoccupied to participate in the interpretation study. Student M believed that by consciously monitoring her voice, she might improve her vocal performance. During the following-up interviews, the ARG students expressed their new understanding of the importance of spoken-voice training for interpreting students and their willingness to participate in the practice. For example, Student O altered her perspective:

After we had the last interview and **I watched my video recording, I find it's interesting and does help the interpreter to perform well so I think it should be included in the main course.** Like voice projection, care of voice and elocution. **I think it is necessary to include voice training into interpreting courses. Because interpreting is not all about language, other skills are also important.**

Student M proposed that spoken-voice training should be included as a required component in a basic-level interpreting course:

Yes, **definitely.** Voice training **should be a part of an interpreter's training.** Even the most adept interpreters are useless without a strong, clear voice. For me, I would like to be training on voice projection, care of voice, elocution and the tone of my voice. **I believe that voice training should be provided at the right start of interpreting training and be compulsive, so lay a good foundation** to our interpreting career.

All of the above indicates that the students' perceptions of spoken-voice training for interpreters changed over time, which may have affected their motivation and self-regulation to improve their spoken-voice performance.

#### **4.2.2.6 Awareness of dealing with tension**

Both the LTG and STG students reported increased awareness when confronted with stressful conditions. Student B (from the STG) described how she managed her tension by paying attention to physical sensations:

**I knew that I was in the fright state again ... I paused a bit to sort myself out, took a few deep breaths and focused on my core and my body sensation** as the voice instructor taught us. **I felt my left hand and my right hand.** They were there. Then **I felt my left leg and right leg ... I felt I calmed down ... I think taking a deep breath could act as a useful pause button** when the wave of emotion is coming. But this method should be practised frequently ... **I learnt that tension itself is not a bad thing and it is very hard to get rid of tension completely.** Rather than attempting to get rid of tension and becoming more nervous, **I should focus more on managing my breath and the body sensation and the message that the speaker wanted to express.**

Student G (from the LTG) also gave an in-depth reflection on her way of dealing with physical tension while speaking or interpreting in public. In her journal, she reflected on several reasons for being anxious while presenting speeches and interpreting. She put these reasons down to a “lack of voicing muscle warm-up, anxiety issue of public speaking, poor preparation for the speech, and also interpreting skills.” She also claimed that her tension was largely due to her “lacking proper voice warm-up.”

The above shows that the students became more conscious of their emotions, accepted stress as a natural part of their lives and careers, and came up with a solution for emotional regulation. Some ARG students reflected during the interviews that they often disguised their nerves by dropping their heads, not looking at their audience or rushing to finish the work when addressing how they dealt with stress during challenging interpreting assignments. Others mentioned taking deep breaths or making gestures to relieve stress, demonstrating a certain amount of awareness in coping with a difficult circumstance. However, all ARG students mentioned that interpreting in front of an audience can be very intimidating.

Although the students from the 3 intervention groups demonstrated improved awareness in dealing with tension, the comments from the 3 external assessors’ perspectives point to the fact that the students did not make significant improvement in “physical tension balance”. In particular, the voice trainer assessor noted that the LTG students had a significant amount of “speech muscle – pharyngeal and laryngeal tension”. As the tension increased, they had to “squash their voice down”. As a result, their voice projection was constricted and this resulted in vocal or glottal fry. The interpreting instructor observed that some LTG students gulped a couple of times when they were nervous or formed a habit of smacking their lips under stress. The voice trainer commented that the ARG students still held “a lot of tension in their breath and their back and hold the facial articulators and tongue back”.

This shows that excessive physical and speech muscle tension was still one of the greatest inhibiting factors that prevented the students in all groups from making substantial improvement in their vocal



performance. As Student O (from the ARG) claimed during the interview, “I think we need to do routine practice. But it is hard to keep doing it every day because we have a heavy workload. It is a long-term improvement.” This suggests that raising students’ awareness alone may not be sufficient to help them improve their physical tension balance. Instead, long-term training and practice may be required to turn tension monitoring from conscious to automatic.

#### **4.2.2.7 Capacity for self-monitoring and self-reflection**

The LTG and the STG students exhibited solid capacity to self-monitor their vocal performance and evaluate their weaknesses and strengths in spoken-voice skills, which is a fundamental element of self-regulation (refer to Section 2.3.2). For instance, Student G (from the LTG) gave an in-depth reflection in her journal on how she monitored her physical sensations and emotions during and after a speech presentation in a voice-training workshop:

During my presentation, as soon as I started talking **I noticed that** I had a very shaky voice and repeated myself a couple of times while I was trying to reorder my thoughts ... **I couldn’t control** my pace and tempo. If I were to pinpoint where exactly these nerves came from, I would have to say that the video recording placed a lot of pressure on myself to do well ... As we later discussed in the class, I **realised** that I hold a lot of tension in my hands and neck, which can become particularly distracting during a presentation ...

**After the speech, I felt relieved and much more relaxed**, and I could talk more in my usual voice while answering questions from the audience.

**Evidently, I was not only dealing with physical tension but also emotional tension that fuels up my anxiety level.**

Student G also discussed in her journal various causes for her anxiety while giving a speech. She described these reasons as “lack of voicing muscle warm-up, anxiety issue of public speaking and poor preparation for the speech.” She also claimed that her tension was largely due to her “lacking proper voice warm-up.” She concluded that physical tension could be managed through ongoing training:

**Muscles can be trained, like public speaking and interpreting skills.**

**Anxiety is caused by** a lack of confidence, which is caused by **insufficient training** in those areas. Confidence is gained, we are not born with it, so in

order to improve my vocal performance and eventually speaking and interpreting performance, **ongoing training for my body and mind is necessary.**

This demonstrates Student G's ability to self-regulate her physical tension and emotion while presenting a speech and her ability to self-evaluate her performance. With these skills, she recognised her weakness and identified that she needed a long-term coping strategy to progress towards the learning goal. Interestingly, Student G was the only student in the LTG who significantly improved her physical tension balance, as determined by the voice trainer assessor's rating scores. "She has improved a lot", the voice trainer assessor observed during the post-intervention evaluation and "There is much more freedom in the body and expression/voice".

The STG students reflected in their journals in great depth on how they monitored their vocal performance during self-practice. For example, Student E reported in her journal how she felt about her self-practice of shadowing at home:

Today I **learnt to practise shadowing at home for the first time** (for the purpose of correcting pronunciation) ... And **I felt breathless** while doing shadowing. After doing shadowing, I almost lost my voice. My throat was sore ... **I talked to the tutor about it.** She said that I need to use what we learnt from the voice workshop to the interpreting workshop and try to manage our breathing better and use chest resonance rather than the throat. **I tried this in my own practice. It worked. Now I understand** the practical side of voice training. **More importantly we need to use it in our practice.**

Furthermore, Student E noticed a substantial improvement in her English pronunciation following regular shadowing self-practice at home, as she wrote in her journal:

**I chose a TED talk** and repeated after the speaker. The result was stunning. **I did better** in imitating the pronunciation and the tone of the speaker than before, **which made me feel good** and after several days of practising, **I got to like my own voice** and the way I speak English.

This shows that Student E consciously monitored her self-practice and evaluated her progress objectively, which demonstrates solid self-regulated learning capacity. When she found problems,

she proactively sought advice from her instructors. Through practising, Student E also concluded that shadowing could serve as an effective warm-up exercise before an actual interpreting exercise and that it was better to do some vocal warm-up exercise before starting a shadowing exercise. This also shows that Student E was consciously applying the spoken-voice training sessions skills to her interpreting skills and seeking the most effective solution to improve her voice quality. In her comment, the interpreting instructor assessor praised Student E, noting that the student demonstrated “clear improvement” in the post-training assessment:

**Her voice projection is much better ... I think her improvement is more apparent among so many students. Her better voice projection has made her voice more audible with better fluency in pronunciation.**

The two voice-training groups’ strong self-monitoring abilities could be a result of the voice-training activities. As Student D (from the STG) mentioned, the speech and interpreting self-evaluation activity (Exercise 22) enabled her to conduct self- and peer reviews of her vocal performance using a list of spoken-voice features, which aided her in becoming familiar with important spoken-voice characteristics, monitoring her practice and setting specific and measurable individual learning goals.

Student H (from the LTG) reflected that she enjoyed the group discussion session when everyone finished the exercises, sat in a circle and was encouraged to reflect on anything they had encountered during the practice. She felt relaxed and talked about her problems freely during the group discussion. This shows that the metacognitive instruction and self-evaluation activities were beneficial for the students in improving their skill of self-regulation.

The ARG students also showed enhanced capacity to evaluate their own performance and determine the most effective learning strategies. Student K, for example, described her vocal quality as “monotonous” and “lacking in energy” during the initial interview, based on comments purely from others – her interpreting instructors and peers:

**According to my tutor and classmates, I'm too loud when I get nervous ...** I speak up and speak very fast. So, people do not understand me in that way and ask me to keep the volume down and speak slowly ... I thought it was all about personality: What and how I want to be and how I speak to people. I thought I need to practise on my personality and be gentler.  
**I don't know if there is another way to help me.**

In the subsequent interview, Student K was able to provide more objective reflection based on her own observation from her video recording. She sounded more confident than in the initial interview:

**I still looked very nervous in this video but seemed composed and more professional.** I did not have much eye contact. I need to work on that with my classmates. **I spoke much clearer than before and more fluently. I need to practise more on pitch and vocal variety** to make my interpreting sound dynamic and interesting. **I will also practise reading aloud and how to break down sentences and syntax.**

This demonstrates that Student K acquired the ability to critically analyse her vocal performance for an interpreting task and to develop a personal learning objective and plan.

The ARG students initially relied on their peers' and instructor's feedback and comments, which was often discouraging when they could not find suitable tools to improve their spoken-voice skills. Through the metacognitive reflection activity, such as listening to their recordings and monitoring their performance using rubrics, the ARG students developed self-assessment capacity and intrinsic ability to regulate their emotions and motivations. This instilled in them a strong motivation to engage in self-directed learning. After seeing her interpreting video recordings, Student L commented on her voice quality as follows:

**I would rate my spoken voice quality as average plus, and I find my voice quality is actually very good. I just need to put more dynamics in it. It takes a lot of practice.**

Apart from improved self-reflection skills, the ARG students showed strong awareness of self-monitoring their spoken-voice performance during interpreting practice. They made conscious effort to monitor their spoken-voice performance and remind themselves to connect with their

audience through eye contact, minimise reading from notes and be mindful of their delivery pace.

For example, Student O stated in the follow-up interview:

I haven't done anything, but I **often remind myself to speak a bit slowly and to speak more clearly**. And to make sense at a slower speed. **I have to remind myself all the time**.

However, some ARG student, for example, Student L, mentioned that they frequently forgot to monitor their performance due to excessive tension during a challenging interpreting assignment.

This shows that, while students enhanced their capacity for self-monitoring, this capacity may not have developed to the point of automaticity without actual exercise, as Happe (n.d.) also noted:

Skill acquisition starts off as laboured, conscious learning and after consistent, frequent practice become more automatic and unconscious. Once the action is well learned, the behaviour becomes automatic in the sense that it does not require constant conscious monitoring. This automaticity allows us no longer to think about the details, and instead to think about the act at a higher level.

This demonstrates that the students learnt to carefully and properly monitor their emotions, reactivity and performance skills through the metacognitive method used in both the voice-training sessions and the self-awareness raising activity, which resulted in different outcomes depending on the intervention method employed. According to the ULD (2018), individuals vary significantly in their capacity and preference for metacognition, and some learners will require much explicit training and modelling to successfully learn how to do so. For others, simply acknowledging that they are progressing towards greater independence is extremely motivating. On the other hand, being able to recognise their own development through self-monitoring and self-reflection is highly motivating for learners. Therefore, it is critical for learners to be provided with different self-evaluation and self-monitoring strategies so that they can discover and select the most effective ones.

#### **4.2.2.8 Motivation and self-regulated practice**

Both the STG and the LTG students displayed self-awareness about what they found motivating, be it intrinsic or extrinsic (refer to Section 2.3.3), in developing their speaking-voice abilities, which is

another critical aspect of self-regulation (ULD, 2018). For example, the students of both voice-training groups reported extensively about how the instructor's clear explanation and demonstration of each spoken-voice exercise and moves motivated them to develop the necessary skills to complete these exercises. In the following journal excerpt, Student C (from the STG) described how the instructor's modelling of vocal warm-up exercises motivated her for autonomous practice:

**Before doing the warm-up exercises, I didn't know the exact aim of these exercises.** Those body activities reminded me of yoga, which helps relax and stretch, but I couldn't relate these exercises to my voice ... After completing all the exercises listed in the handouts following the voice trainer's step-by-step instruction, **I began to understand** that warming up the body before speaking is very important and has positive effects on increasing blood flow to muscles and tissues, and the elimination of excess tension. **I then looked up more information** about voice warm-up online and **practise these exercises daily**, warming up my vocal organs and body before interpreting practice.

As seen above, following the instructor's demonstration Student C displayed an internal motivation to set realistic learning targets and maintain a positive attitude towards accomplishing those objectives. She took the initiative to set up a vocal warm-up routine before starting daily interpreting practice, and selected some of her favourite exercises to work on, such as stretching, massaging her face and jaws, relaxing her lips and moving her tongue. She stated in her journal:

**I chose some of the postures** that can help me stretch and relax as the body warm-up and **follow the other steps in the handouts.** After several times of practice, **I feel the organs of articulation become flexible** than before. And most importantly, **it makes me aware of the tension in my body and my mind. Doing these exercises wakes up my whole body to a certain level, strengthens my interpreting work and improves my confidence consequently.**

This shows that the students could be motivated to do exercises and develop their agency in learning to become independent learners through a great deal of guided practice steps as the

external motivation. The independent practice allowed the students to apply what they had understood from the explicit instruction in their interpreting practice.

Aside from modelling, some students found that receiving personalised feedback from the instructor on a specific spoken-voice issue was motivating. Student J (from the LTG) reflected on how the voice instructor's "reminder" and "special task" benefited her in the following journal excerpt:

**One of the things she [the voice trainer] noted was my knees were standing tightly together and the way I held my notebook** made my shoulders tense, which **led to feeling nervous** at the top half of the body; the teacher said if I had stood more relaxed it would have helped me speak. **The teacher suggested some breathing exercises** that I can do, for example, lying semi-supine and practising breathing and imagining all those muscles releasing and then extending exercises in a standing position. **I take her advice and practice every day in front of a mirror. I've noticed that my standing posture is much better than before.**

Student J received specific personal feedback from the voice workshop that she "inhaled and exhaled a lot during one sentence and it wasn't a smooth flow" and was instructed to read a speech following a natural rhythm and flow while focusing on breathing and pausing at the right places, particularly when encountering a long sentence. She reflected that this exercise helped her improve delivery fluency by using proper breathing techniques. She felt her body was relaxed, her vocal cords were powerful and her voice was freer. She wrote:

I didn't know that I was doing that when I was giving the speech, **but as she [the voice trainer] pointed out**, I was mainly focusing on how to use my voice appropriately (such as tone, volume) for the topic of my speech, as well as trying to recall the next sentence of my speech, and the incorrect breathing rhythm caused my speech to be less fluent and rhythmic ... **After the exercises, I felt relaxed. When I was using my voice again, I felt the voice was clearer and louder** as if there was more power coming out from my chest. **It was easier to speak clearly without much effort; I didn't notice any blockage of sound in my voice** after the exercises.

This demonstrates that the individual instruction and feedback served as external motivation not only making the students aware of their problems, but also making it easier for them to monitor their performance and regulate their learning strategies. Both the LTG and the STG students

showed clear understanding of the intention of the voice training, the expected learning outcome, the instructor's expectations and directions, and the personal vocal issues to be addressed.

Apart from actively participating in the workshop activities, the STG students reflected extensively on regular self-practice outside of the spoken-voice workshops, demonstrating strong intrinsic motivation and self-regulated learning strategies (Zimmerman, 1990). For example, Student E realised that “knowing how the voice works is far from enough”. Student B stated that the vocal exercises should be a “daily job” to maintain their competence. Student D practised the tongue-twisters every day to improve her tongue's flexibility and “observed improvements from practising time after time”.

Student B practised, in her own time, emphatic intonation patterns with paragraphs in different manifestations such as news, scientific articles and business reviews using distinguishing tones. She recorded herself reading out paragraphs using different tones and then did self-appraisal by listening to her audio recordings. She realised that “controlling the pace of speech is crucial and articulation needs to be precise”.

These examples demonstrate the STG students' clear understanding of why they chose specific exercises and reflect the outcomes of such practice. As a result, they were able to turn their motivation into self-regulated learning skills. This evidence supports Eom (2019)'s claim that intrinsic and extrinsic motivation activates the self-regulation process, which has a beneficial effect on learning outcomes.

The LTG students expressed enthusiasm about attending the voice-training workshops in their reflective journals and reflected extensively on how spoken-voice training allowed them to become aware of their unique voice quality and acquire vocal skills to improve their interpreting performance. However, in contrast to the students of STG, who reported daily about how they were motivated by their self-practice, few students in the LTG mentioned how often they conducted self-directed exercises and what exercises they did outside of the workshops. The students in the LTG



also expressed concern about losing motivation for self-practice during the long intervals between the voice-training workshops. Following are journal excerpts from Student G and Student F (both from the LTG). Student G mentioned that she did not practise vocal skills regularly because she was worried that she would make mistakes in self-led practice due to not getting timely feedback from the voice instructor during the long break. This student's journal entry indicates that the training interval had a significant effect on students' motivation to conduct self-practice. Student G stated:

I think **four sessions of voice training workshops are not enough** and I believe **it's better to have the workshop at least once a week** rather than once every couple of weeks. In that way, I can practise the vocal skills at home and get feedback from the instructor the next week. **I stopped the vocal practice because I wouldn't receive feedback regularly and I'm afraid I would practise in the wrong way, which is worse than not practising at all.**

Student F disclosed that although she was aware that self-practice was important, she did not practise vocal skills regularly due to a heavy study load. Moreover, in her journal she expressed her disappointment at not noticing much change in her vocal performance after 4 sessions of voice-training workshops and how she had lost motivation:

Despite the benefits of the above exercises, I unfortunately **noticed very little change** in my voice after the voice exercises, **nor did I notice any change in my interpreting practice. My classmate said she did not notice many changes in her voice or mine, either.** The voice training and the interpreting practice seemed **not very connected.** These exercises are obviously meant to be done regularly to see any clear or permanent effects. At the beginning of the voice training I was enthusiastic, but now I seem to have **lost interest.** Maybe because of the **heavy study load.** I just **don't have time and enthusiasm to practise** now.

This reveals that Student F did not proceed with the spoken-voice self-exercises because she lacked intrinsic motivation and self-efficacy in conducting self-practice in spoken-voice skills outside the spoken-voice workshops and was unable to apply the spoken-voice skills in interpreting practice, which may have resulted in lessened achievement throughout the process. As the results show,

Student F declined in several aspects, including “physical tension balance,” “rhythm and fluent rendition” and “expressiveness,” according to the external assessor’s ratings (see Appendix N-2). The lay assessor compared the pre- and post-assessment videos and concluded that she enjoyed Student F’s performance in “the first video” more than in “the second video”. The following is an excerpt from the lay assessor’s comment:

**The voice about the first one is better. Everything about the first one is better.** She looked fairly professional in the second video but a little **more hesitant** ... In the first video, she interpreted with confidence and expression and lightness. No tension. Excellent and I enjoyed very much. She did not read her notes a lot. **But in video 2, she concentrated on the reading more so than the expression.** Her tone was little bit flat. She was reading her notes a lot. So stilted and tone down. Tension in face. Fluency is a little stifled. Expression is a little “flat” reading her notes. And intonation is a little flat as well. Just a little slower and a little more hesitant and attentive than the first time. I think the biggest thing was the fluency with which she spoke from her notes. I like the first one better, but I can see the difference. Compared with other students doing the same thing, she was very good. Her voice quality is good. **It was mainly the fluency that hindered my ability to follow what she was saying.**

According to the interpreting instructor assessor, Student F appeared “slightly more confident and her presentation more professional. Her voice seems to have dropped slightly, showing a certain degree of maturity.” However, “She has too many ‘ums and ahs’... and flat tone, and fluency is a big problem.” From the assessors’ comments, it can be seen that Student F was concentrating on reading her notes rather than voice skills and she was agitated, although she may have shown a bit more maturity and professionalism through interpreting skill training over one semester. This also shows that interpreting training may not necessarily help students improve their spoken-voice skills. Student F may have improved interpreting accuracy and linguistic transfer skills, and looked more “mature” than in the first video, but her overconcern for accuracy brought her much tension and sacrificed her vocal performance in relation to tone, fluency and audience connection.

These results confirm previous research indicating a significant and positive relationship between intrinsic motivation and academic achievement (Lepper et al., 2005; Pérez-López & Contero, 2013), implying that intrinsic motivation is critical in influencing students' academic achievement. The fact that Student F was not satisfied with the result of training shows that there is a strong link between learning motivation and learning satisfaction (Huang, 1992). According to Martin (1988), learning satisfaction is the level of enjoyment that students experience after undertaking learning activities and is the degree of consistency between people's expectations and their actual experiences. Huang (1992) noted that after students engage in learning activities, learning satisfaction comes from fulfilling the learning needs generated by their learning motivation. Low learning satisfaction reduced the students' motivation for self-practice.

The above two journal excerpts from LTG students show that the absence of systematic and routine self-exercises appears to have hindered the LTG students in attaining the same vocal performance as the STG students. This indicates that lack of motivation for practising vocal skills outside of the workshops is a factor that is likely to have inhibited the LTG students' improvement, despite their enthusiasm (extrinsic motivation) for practising vocal warm-up exercises at the start of the workshops and self-awareness of motivation for improving their vocal skills. More crucially, the results indicate that the separation of vocal and interpretation practice, as well as the length of time between training sessions, may have demotivated students to engage in self-directed practice. In contrast, the STG students showed greater learner autonomy to continuously practise their vocal skills and integrate them with their interpreting practice, which was a key reason for their remarkable growth in spoken-voice skills. This indicates that the duration and training interval had significant effects on the training outcomes and motivation of the students.

According to the ARG students, viewing their interpreting video recordings and self-evaluating their vocal performance were motivating factors for them to set up personal plans to improve their vocal performance. Although the ARG students did not receive spoken-voice training, many of them showed enthusiasm in engaging in self-practice to enhance their spoken-voice performance,

for example, sourcing articles about spoken-voice training on the internet, practising reading speeches aloud and imitating skilled speakers, recording and listening to their interpreting output and evaluating the learning outcome. As Student N stated in the second interview:

**I didn't get to do any voice training. I think voice training is expensive and time consuming. But I did some voice exercises at home. I found some exercises online. I make efforts not to speak too fast. I found reading aloud is useful,** newspaper, poems, and record and listen afterwards. And to see if my speed is too fast. **I also record my practice** from interpreting workshop **and listen at home.**

In particular, Student L went out of her way to find a spoken-voice coach to guide her practice, in addition to conducting an online search for books and articles on voice training. As she put it:

**I take a tutoring class** with voice training. **I recorded myself speaking** out loud. **I listen to my recording** in both languages. I also **practise breathing techniques. I think I've improved a lot. I sound more dynamic** than before and **show more confidence.** I will do a lot of shadowing and imitate speakers who speak at international meetings.

The ARG students also made conscious efforts to monitor their spoken-voice performance during interpreting practice. As the ARG students showed considerable improvement in the post-intervention assessment, this demonstrates that they were motivated in self-directed learning.

However, self-practice without an instructor's guidance may still have acted as a hindrance to their long-term progress in their vocal skills. One of the impediments, as Student L described below, was that her growth was not constant or sustainable due to a lack of actual spoken-voice training and practice:

I need practice, but **it is very hard to do on your own ... I need audience's feedback about how my voice sounds.** I mean, I can practise at home, but I don't know how my voice sounds in other people's ears. For example, **voice volume, in my recording it's still loud enough, but would be different from the feedback from my classmates in the interpreting setting.**

In summary, Section 4.2.2 has presented the factors that contributed to the significant improvements of the 3 intervention groups. The results show that the intervention methods employed in this project served as external factors that facilitated intrinsic changes in the students, resulting in different levels of progress in the 3 groups. The results indicate that metacognitive instruction, whether via explicit spoken-voice instruction or an implicit vocal awareness-raising approach, facilitated students' metacognitive awareness and self-regulation, as evidenced by the students' changed perceptions of spoken-voice training, vocal care awareness, improved self-awareness regarding their motivation for improving spoken-voice skills, improved self-assessment and self-monitoring skills, and enhanced awareness of dealing with tension and emotion in challenging interpreting assignments.

The results have also revealed that the STG students outperformed the LTG students due to regular self-practice, better skills in vocal dynamism and better coping and self-regulation skills in a challenging situation. Although the LTG students showed great enthusiasm in practising spoken-voice production skills and substantially improved kinaesthetics, vocal care and prosodic feature awareness, the LTG students reported that excessive physical and speech muscle tension and lack of intrinsic motivation for self-practice hindered their further improvement. Although the ARG students did not receive spoken-voice training, they reported substantially improved spoken-voice awareness, self-evaluation and reflection skills following the metacognitive reflective intervention. However, their improvement was limited, most probably due to the absence of instructor-guided spoken-voice guidance and regular practice.

## Chapter 5 Discussion of Results

In this chapter, the results presented in Chapter 4 concerning the research questions of this study are discussed in light of relevant studies. The discussion aims to find the connections between the results for the 5 research questions. The following is a summary of the major findings in relation to the 5 research questions.

1. The 3 spoken-voice intervention groups statistically recorded significantly higher levels of vocal performance at the post-intervention stage compared to the pre-intervention stage. In contrast, the non-intervention group recorded no improvement in the post-intervention interpreting assessment.
2. The 3 intervention groups attained significantly higher levels of vocal performance than the non-intervention group. The STG significantly outperformed the other groups and attained the best intervention results.
3. A significant positive correlation was discovered between students' spoken-voice performance and their overall interpreting performance in both pre- and post-intervention assessments, indicating that the higher the students' spoken-voice performance scores, the higher their overall interpreting performance scores.
4. Among the 11 highly improved spoken-voice attributes, "vocal confidence" was the most prominent attribute contributing to the significant improvement of all 3 intervention groups. The STG had the highest number of highly improved attributes across all groups. Apart from "vocal confidence," STG students improved in a variety of other vocal features including all attributes of "spoken voice dynamism", as well as "speaking clarity," "posture alignment" and "voice projection," as identified by the 3 external assessors. The LTG students demonstrated significant improvement in some attributes of "spoken voice production" and "spoken voice quality," including "speaking clarity," "breathing support" and "pitch range modulation", mainly from the perspective of the voice trainer assessor. However, except for "vocal confidence," the LTG did not improve in any other aspect of "spoken voice dynamism". The

ARG students improved in the 3 attributes of “vocal dynamism”, namely “vocal confidence”, “audience connection” and “expressiveness”, as well as “pace and pauses” and “posture and alignment.” However, the ARG students improved less in “spoken voice production” and “spoken voice quality”. Notwithstanding these highly improved attributes, a few spoken-voice attributes remained unchanged in all groups, including “physical tension balance”, “movement”, “resonance and timbre” and “speech muscle tension release”, suggesting that the intervention may not have been as effective for these attributes as for the other attributes.

5. The results have revealed that the intervention strategies used in this study acted as external influences that aided students’ internal changes, resulting in varying amounts of progress among the 3 intervention groups. The explicit spoken-voice training enhanced the STG and LTG students’ spoken-voice awareness and skills in vocal production and quality. The ARG students significantly improved their vocal skills due to improved vocal awareness and self-evaluation skills following the MRI intervention. At the same time, the result also reveal that the STG students outperformed the LTG students in spoken-voice dynamism owing to improved audience awareness and higher motivation for self-practice. Excessive physical and speech muscle strain, a lack of intrinsic motivation for self-practice and a lack of audience awareness were shown to be stumbling blocks for the LTG students. The ARG students’ improvement was hampered by a lack of fundamental spoken-voice technique training.

In this chapter, Section 5.1 discusses the possible causes for the 3 groups achieving positive intervention results. Section 5.2 analyses the effects of spoken-voice performance on the delivery component of interpreting performance. Section 5.3 discusses the considerable improvements in spoken-voice categories and qualities following the intervention. Section 5.4 outlines the viewpoints that the 3 external assessors contributed to this study.

### **5.1 Intervention methods and intervention effects**

As shown by the findings presented in Chapter 4, the intervention methods yielded significant levels of elevated performance in the 3 intervention groups. The fact that all 3 intervention groups

achieved significantly higher levels of vocal performance at the post-intervention stage compared to the pre-intervention stage indicates that the metacognitive approach embedded in the intervention activities, whether explicit spoken-voice training workshops or spoken voice awareness-raising sessions, assisted students in changing their perceptions of spoken-voice training for interpreters, enhancing their vocal awareness and spoken-voice skills, and improving their self-regulation. However, the outperformance of the STG over the other two intervention groups demonstrates that the level of improvement varied according to the different training regimens and approaches employed in the intervention.

The intervention methods utilised in each group are summarised in Table 5.1, including voice-training regimens, intervention activities and metacognitive instruction approaches. These external effects helped to raise the student participants' internal factors including vocal awareness, metacognitive awareness and self-regulation, resulting in varying progress in spoken-voice presentation skills in interpreting delivery. The following subsections contain in-depth discussion.



Table 5.1 Intervention methods and intervention effects.

Intervention methods External factors		Experimental groups				Internal factors	Intervention effects
		STG	LTG	ARG	CoG		
Intervention activities	Instructor-guided spoken- voice activities (explicit instruction)	√	√	×	×	Improved kinaesthetic awareness (LTG and STG) Improved vocal care awareness (LTG and STG) Improved prosodic feature awareness (LTG and STG) Improved audience awareness, inclusive awareness (STG)	Improved spoken- voice production <ul style="list-style-type: none"> <li>• Posture &amp; alignment (STG)</li> <li>• Breathing support (LTG)</li> </ul> Improved spoken- voice quality <ul style="list-style-type: none"> <li>• Speaking clarity (STG&amp; LTG)</li> <li>• Voice projection (STG)</li> <li>• Modulation of pitch range (LTG)</li> </ul> Improved spoken- voice dynamism <ul style="list-style-type: none"> <li>• Vocal confidence (STG &amp; LTG)</li> <li>• Intonation &amp; tone (STG)</li> <li>• Expressiveness (STG)</li> <li>• Rhythmic &amp; fluent rendition (STG)</li> <li>• Audience connection (STG)</li> </ul>
	Awareness-raising interviews and self-evaluation activities (implicit instruction)	×	×	√	×	Improved vocal care awareness Improved prosodic feature awareness Improved audience awareness	Improved spoken- voice production <ul style="list-style-type: none"> <li>• Posture &amp; alignment</li> </ul> Improved spoken- voice quality <ul style="list-style-type: none"> <li>• Pace &amp; pauses</li> </ul> Improved spoken- voice dynamism <ul style="list-style-type: none"> <li>• Vocal confidence</li> <li>• Expressiveness</li> <li>• Audience connection</li> </ul>
Voice-training regimens	Short voice training sessions with short breaks between sessions	√	×	×	×	Enthusiasm at workshops Motivated for regular self-practice	<ul style="list-style-type: none"> <li>• High training effects at and after workshops</li> </ul>
	Long interval intensive training sessions	×	√	×	×	Enthusiasm at workshops Demotivated for regular self-practice	<ul style="list-style-type: none"> <li>• High training effects at workshops but reduced effects after workshops</li> </ul>
Metacognitive instruction	Student- centred instruction PBL After- practice discussion Reflective journal writing	√	√	×	×	Changed perceptions of spoken-voice training for interpreters (STG and LTG) Improved awareness in dealing with tension (STG & LTG) Improved awareness to self-monitor and evaluate own practice (STG and LTG)	<ul style="list-style-type: none"> <li>• Improved metacognitive awareness</li> <li>• Improved self-regulation</li> </ul>
	MRI	×	×	√	×	Changed perceptions of spoken-voice training for interpreters Improved awareness in dealing with tension Improved awareness to self-monitor and evaluate own practice	<ul style="list-style-type: none"> <li>• Improved metacognitive awareness</li> <li>• Improved self-regulation</li> </ul>

### **5.1.1 Intervention activities and improved spoken-voice awareness**

The findings (see Section 4.2.2) indicated that spoken-voice intervention activities, such as instructor-led workshops (for the STG and LTG) and awareness-raising activities (for the ARG), significantly increased the students' spoken-voice awareness and spoken-voice skills. The following subsections discuss these in detail.

#### **5.1.1.1 Kinaesthetic awareness and spoken-voice production**

According to the external assessor's comments reported in Section 4.2, the leading causes of vocal problems for the interpreting students in the pre-training assessment included looking down at their interpreting notes with low awareness of their posture, body alignment and movement, mechanically pushing up their voice, focusing exclusively on reading from their notes without expression or awareness of their audience, speaking with shallow breathing and/or too much or too little nasal and/or oral sounds and high and unchanging pitch, and speaking at an inconsistent pace with less articulation and a heavy accent, which corroborates Barton (1997) description of disadvantages for learners who are predominant visual learners (see Subsection 2.2.4.3).

This study has discovered that the STG and the LTG students' improved spoken-voice production skills, including better posture and alignment and breathing skills as demonstrated in the post-training assessment, were associated with their increased kinaesthetic awareness following instructor-guided kinaesthetic activities such as semi-supine Alexander techniques, posture and alignment exercises, vocal warm-up activities and breathing exercises. The results show that when the students were exposed to visual auditory kinaesthetic (Fleming, 2006) modes of learning in the voice workshops and became aware of their potential for vocal skill improvement, they were likely to upgrade to auditory and kinaesthetic senses from being visual learners (Barton, 1997). With enhanced kinaesthetic awareness, the students were more likely to develop an inclusive awareness (Malde et al., 2013, p. 5) which enabled them to connect bodily movement, emotion and spoken-voice production when interacting with their audience. This would thus better meet the speakers' communicational requirements and purposes. This indicates that explicit instruction of kinaesthetic

activities was highly effective in guiding students to develop “psycho-physical awareness” (Linklater, 2006, p. 31), which enabled them to notice and monitor their mental and physical sensations and tension in vocal production caused by challenging interpreting assignments, as well as developing the high physical and mental stamina required for interpreting tasks.

While the ARG students demonstrated high improvement in posture and alignment (the only improved attribute in spoken-voice production), none indicated a link between their kinaesthetic understanding and their improvement in this attribute in their interview responses. Two reasons may explain this phenomenon. First, the kinaesthetic sense is more likely to have been trained effectively through explicit instruction of the relevant spoken-voice exercises than implicit instruction in the awareness-raising sessions. Explicit teaching (see Subsection 2.2.4.9) entails focusing students’ attention on specific learning in a highly structured environment, a type of instruction that is directed towards achieving specific learning outcomes. As reported in Subsection 4.2.2.8, using the explicit instruction method the voice coach provided instructions on precisely executing a movement. This way, students in the two voice-training groups were introduced to the movements deliberately and in small steps, and thus gained the essential skills to perform these exercises. This result reaffirms the research evidence (Ellis, 2009; Yenkimaleki, 2018) which has established that students who are exposed to explicit teaching strategies and receive appropriate feedback make more significant improvement in learning than students who are not exposed to these practices.

Second, the ARG students may have already developed some kinaesthetic awareness but were not conscious of the learning that occurred without the utilisation of central attentional resources; as suggested by many studies, implicit learning is largely independent of attention as a resource and is driven in response to a stimulus input independently of control processes or selection (Wolf & Müller, 2012). As a result, the ARG students may not have been able to verbally explain what they had learnt even while they exhibited clear behavioural responses. This suggests that although non-conscious metacognitive knowledge may be developed through implicit instruction, or self-

instruction through many years of trial and error, if the knowledge is not able to be eventually brought to consciousness, the strategy used may remain simply a cognitive response rather than a metacognitive process, which involves an awareness of themselves as “agents of their own thinking” (Kluwe, 1982, p. 222). This indicates that it is critical for students to acquire metacognitive skills that enable them to reflect on their own mental processes. This understanding of their own learning process enables them to exercise greater control over it. Additionally, it strengthens their ability to regulate and manage their own motivation for learning. Furthermore, the data implies that conscious learning and practice are more likely to culminate in unconscious automaticity. As Happe (2001, p. 991) stated:

Skill acquisition starts off as labored, conscious learning and after consistent, frequent practice becomes more automatic and unconscious. Once the action is well learned, the behavior becomes automatic in the sense that it does not require constant conscious monitoring. This automaticity allows us no longer to think about the details, and instead to think about the act at a higher level.

#### **5.1.1.2 Prosodic feature awareness and spoken-voice quality**

The results that all 3 intervention groups improved in some prosodic features such as “voice projection”, “speaking clarity”, “modulation of pitch range” and “pace and pauses” (See Table 4.7) are consistent with previous studies on the positive effects of training of interpreters’ theatrical performance skills and prosodic features on their interpreting performance (Ahrens, 2015; Cho & Roger, 2010; Kayi, 2012; Lim, 2008; Naimushin, 2014; Shlesinger, 1994; Williams, 1995). Additionally, the findings validate a study (Yenkimaleki, 2018) in which prosodic feature training was found to increase the quality of speech production in CI, notably in terms of intonation, tempo and voice quality.

However, as the results show, prosodic features, as fundamental aspects of speech communication, could pose significant challenges to interpreters’ spoken-voice quality, especially for those who speak English as their B language. Given the difficulty of students acquiring mastery of these prosodic features through short-term intensive training sessions, this study corroborates Gut et al.’s

(2008) assertion that instructors should adapt their training methodologies to different situations, such as native and TLs and training objectives, and proposal that training on awareness of these prosodic features should be included as a core component of spoken-voice training for interpreters and encourage consistent self-practice outside workshops.

### **5.1.1.3 Audience awareness and spoken-voice dynamism**

The study has discovered that the significant improvements in spoken-voice dynamism demonstrated by the STG and ARG students were related to increased audience awareness. These findings verify the assertions of several interpreting scholars that training in performative aspects and speech behaviour is crucial for the quality of CI performance (Lim, 2008). Instructor-led vocal performance training has been shown to improve the quality of CI by helping students convey a message effectively and provide high performing quality for the audience (Cho & Roger, 2010).

What has been shown is that through explicitly instructed activities such as role play, voice persona, impromptu speech and public-speaking activities, students became mindful of vocal delivery and audience connection while performing interpretation. The STG students reflected extensively on the spoken-voice training activities that made them aware of their strengths and weaknesses, and helped them present a professional image to the audience, for example, through the training of eye contact and hand gesture, the practice of reading out selected texts with appropriate vocal energy and the role-play activities.

The ARG students also developed awareness of connecting with the audience through the interviews and self-evaluation activity by constantly reminding themselves to maintain eye contact, to reduce note reading and to regulate the delivery pace and pauses, in addition to peers reminding each other. However, the ARG students needed to monitor their audience connection behaviour consciously because sometimes they forgot to do so, which shows that their awareness remained at the “conscious monitoring” stage and had not yet reached the state of automaticity (Tzelgov, 1999). This means that raised awareness may need further substantiation in actual practice in the interpreting environment so as to reach automaticity in audience connection and vocal performance.

This may explain why the interpreting instructor assessor noticed that some ARG students had “fake eye contact. Gave people a feeling like staring at you. Doesn’t look polite. She was not conscious of it” and the lay assessor commented on Student L that “she needs to look at different spots in the audience – you connect better if you are looking around the audience, not at the same spot”. This demonstrates that, while the ARG students had greater knowledge of their audience and planned to work with it, they were unable to do this adequately due to a lack of instruction. In contrast to this, the STG students demonstrated automaticity and being “in the flow” (voice trainer assessor). This indicates that although implicit instruction could raise the students’ awareness in audience connection and performance skills, the explicit instruction on spoken-voice activities was more effective in developing “smooth coping” in highly refined skills such as CI skills (Bergamin, 2017, p. 403). This also means that audience awareness needs to be integrated into real interpreting practice to reach automaticity.

The LTG students showed strong interest in practising fundamental spoken-voice production skills such as doing breathing exercises and vocal warm-exercises. However, they did not demonstrate explicit audience awareness or communication behaviour, which may be one of the factors in the LTG students being reported to have improved minimally in spoken-voice dynamism by the assessors. This suggests that interpreters’ audience awareness is important for presenting spoken-voice dynamism. As Horváth (2012) pointed out, interpreters’ speech behaviour impacts on the audience by providing information or entertainment or by asserting knowledge, among other things. A lack of audience awareness can negatively affect speech behaviour, including voice production, breathing, pronunciation, prosodic factors, eye contact, gestures, movement and physical appearance (Horváth, 2012). This may account for the audience’s lack of enthusiasm for the LTG students despite their enhanced spoken-voice production skills.

### **5.1.2 Spoken-voice training regimens and motivation for regular self-practice**

The fact that the STG outperformed the LTG points to the conclusion that brief voice-training sessions with short breaks between sessions (e.g., 20 mins/session/workday) have higher

intervention efficacy than long-duration voice-training sessions with extensive gaps between sessions (e.g., 120 mins/session/fortnightly). This result is consistent with past findings that, apart from the overall amount of practice and the content of training delivered, the outcome of perceptual learning can be affected by the distribution of training, such as the length of each training session (Wright & Sabin, 2007) and the space between sessions (Donovan & Radosevich, 1999; Goedert & Miller, 2008).

One explanation for the shorter training sessions providing a benefit over more extended sessions, based on past experiments, is that the short and regular training regimen “allowed for more latent, between-session and post-training learning to emerge” (Molloy et al., 2012, p. 1). As suggested by Banai et al. (2010), although within-session and between-session learning can be two independent processes of training, these two processes can interact and have a synergetic effect on the outcome of perceptual training. In other words, improvements may be apparent while training or during within-session learning, but improvement can sometimes occur during a latent period after training has finished or as between-session learning (Wright & Sabin, 2007).

This may account for the STG’s significant improvement as a result of workshop activities and self-practice in between scheduled sessions. The STG students’ motivations and actions for persistent and frequent self-practice and self-reflection happened outside of the workshops, as evidenced by their self-reflecting journals, and could be due to the students having been exposed to spoken-voice training on a regular basis and encouraged to perform vocal exercises regularly to maintain their competency. As Anderson (1992) asserted, the persistent practice of a particular skill can transfer to conscious automaticity, without the need for conscious monitoring of one’s voice production, and thus reduce the cognitive load for interpreting delivery.

While the LTG students showed a high level of participation in the workshop activities, the lengthy time between sessions may have had a negative influence on the training effect, as the students described in their journals a sense of saturation and exhaustion after a two-hour training session and a loss of motivation for self-practice during the extended periods between sessions. As suggested by

the findings from past experiments (Hauptmann & Karni, 2002; Hauptmann et al., 2005), the within-training effect can reach its peak towards the end of a training session and restart once a new session has begun, while between-session improvements are greatest at the start of training and gradually reduce as training progresses. This implies that the LTG students may have met the challenges of over-training and fatigue due to extended periods of sustained engagement with some training activities. Thus, “a failure of overnight consolidation of improvements attained within a session” (Molloy et al., 2012, p. 2) could be an explanation for the LTG students’ lower performance improvement gained during training. In contrast, according to Halliday’s assertion brief and continuous training sessions can improve learners’ compliance with instructional standards and their attention spans (Halliday et al., 2008).

The long-interval training schedule may have demotivated the LTG students from engaging in regular self-practice outside of the workshops, which is a significant factor that is likely to have inhibited the LTG students’ substantial improvement, particularly in relation to “spoken-voice dynamism”. The lengthy gaps between spoken-voice training sessions may have left the LTG students out of practice and unable to transfer the spoken-voice skills obtained in a highly organised learning setting to self-practice and the actual work environment. These results conform with what Anderson (1992) observed, that in skilled performance, when the quantity of forgetting over long intervals appears to be significant compared to the amounts of improvement with practice, well-practised skills degrade exponentially with disuse. As a result, the LTG students may have been unable to maintain and strengthen their spoken-voice skills obtained during the workshops.

These results confirm that training is most effective when the opportunity for latent improvement between practice sessions is maximised, especially during the early training stage (Molloy et al., 2012). This shows that short breaks between sessions could provide an opportunity for students to retain the effects of training and take advantage of latent learning following training sessions. The results support the claim (Molloy et al., 2012) that optimal training regimens at an early stage of



spoken-voice training should be employed in short sessions which are spaced within a smaller number of days rather than packed sessions over a more extended period of days or weeks.

### **5.1.3 Metacognitive instruction and self-regulation**

The findings have demonstrated that when metacognitive instruction is paired with spoken-voice intervention approaches, whether explicit spoken-voice training or implicit vocal awareness-raising, students can gain metacognitive awareness and self-regulation. The metacognitive instruction on explicit spoken-voice training improved the LTG and STG students' spoken-voice skills and increased their awareness of vocal exercise, suggesting that a metacognitive mechanism prompted their self-regulation for practice. The explicit instruction and PBL approach focusing on individual-based vocal-management problems (Dunlap, 2005; Paris & Paris, 2001; Turan et al., 2009) enabled the students to become aware of their own voice challenges and strengths and, more importantly, to set goals to overcome their problems and reach their potential. The students were guided to self-monitor their own vocal performance and to find the rationale behind each specific exercise.

The group discussion, which was based on a learner-centred approach (McCombs & Whisler, 1997), allowed students to ask questions and receive clear feedback on their performance against learning outcomes, which improved the students' self-monitoring and assessment skills and allowed instructors to use the students' feedback to determine if they had understood what they were taught or needed to repeat the instruction.

The reflective journals proved to be another effective tool to raise students' metacognitive awareness as a means of improving their vocal performance. The students' reflective journals, including self-reflections on their own and peers' work and provision of feedback on the instructor's teaching, show that they were highly motivated and became enthusiastic participants in their own learning, restructuring knowledge rather than just being passive recipients (Levett-Jones, 2005). In writing a learning journal at the end of each training session, students were stimulated to contemplate their own cognitive processes and regulation of cognition (two metacognitive dimensions) (Henter & Indreica, 2014). In the journals, the students reflected on their cognitive

processes in performing vocal and speaking tasks, their strengths and weaknesses in using voice and speech, and other aspects related to voice management such as their motivation in applying their learning to practice and looking for voice-related resources. They also appraised their peers' vocal performance that they observed by applying the knowledge they learnt from the workshops. The students' reflective journals also helped the voice trainer to understand the students' feelings, beliefs, knowledge and skills regarding the training and to improve the instruction. As a result, the training can be considered to have become more student-centred. This is consistent with previous research (Henter & Indreica, 2014; Olson & Johnson, 2012) about the effectiveness of reflective journals in terms of improving students' motivation and self-regulation.

This is congruent with previous studies showing that metacognitive instruction or strategy is associated with successful learning (Borkowski et al., 1987) and can be taught (Halpern, 1996). The results suggest that instructors should encourage learners to become more strategic thinkers by helping them focus on the ways they process information. These can include “self-questioning, reflective journal writing, and discussing their thought processes with other learners as being among the ways that teachers can encourage learners to examine and develop their metacognitive processes” (U.S. Department of Education, 2011).

The interview questions presented to the ARG students allowed them to contemplate and arrive at their own conclusions. The implicit approach embedded in the MRI procedures made students aware of “an experience from which they learnt about themselves as learners and how they might alter their strategies and their knowledge construction” (Anderson et al., 2009, p. 191). The self-evaluation activity allowed the students to reflect on each spoken-voice attribute for their own vocal performance using a metacognitive strategy. Subsequently, the students showed raised awareness of prosodic features and elevated levels of motivation for self-practice. These findings verify Schraw's (1994) conclusion that metacognitive awareness plays a compensatory role in cognitive performance by improving strategy use, demonstrating that metacognitive intervention strategies can have a correlative effect on students' academic performance and contribute to students' self-

efficacy, goal-directedness and self-regulation of learning abilities and academic performance (Clark, 1988; Emerick, 1992; Young & Fry, 2008).

The fact that neither the ARG nor the CoG received spoken-voice training but only the ARG achieved noticeable improvement and received high audience preference (especially from the lay assessor) validates the argument put forward by Huang et al. (2004) that prosodic feature awareness is not an automatic result of language learning and can only be acquired through targeted awareness intervention. This study indicates that the spoken-voice awareness intervention was pertinent and that the students' improvement in their vocal performance was not automatic due to some halo-effect caused by the interpreting curriculum or linguistic practice. Instead, this study argues that the gain in vocal performance was obtained because of the customised training on various spoken-voice attributes specific to interpreters' requirements and better awareness of the use of the spoken voice in interpreting performance. The results show that in circumstances where the provision of spoken-voice training is limited, raising spoken-voice awareness and metacognitive awareness could also improve students' vocal performance to satisfy the quality criteria of the interpreting audience and interpreter trainers.

It is also worth mentioning that, while the ARG students achieved great improvement in the eyes of clients and interpreting professionals, the voice trainer assessor did not see their progress as noteworthy. This result suggests that the ARG students' improvement was negligible from the voice trainer assessor's perspective. It is understood that the voice trainer assessor, as a vocal professional trained to examine spoken-voice characteristics, voice-using techniques and skills based on specific vocal analysis criteria, would view the students' vocal performance from the perspective of a specialised audience. This means that the ARG students may have made insignificant progress in their spoken-voice skills and techniques due to a lack of explicit spoken-voice training, implying that while students may improve their metacognitive awareness and self-regulation, their improvement in spoken-voice skills may be limited unless they seek additional assistance from a

voice professional. This corroborates Yenkimaleki and van Heuven's (2018) finding that explicit instruction in the use of prosodic features improved interpreter trainees' speech production.

This study supports Yenkimaleki and van Heuven's (2018) recommendation that interpreting students with English as their second or B language should receive extensive explicit instruction and practice in order to distinguish between the two languages' word stress systems and to acquire the correct knowledge of the stress pattern in English words. Furthermore, other recommendations from them in relation to phrase prosody, target English pronunciation, sentence structure and accentuation, as well as the use of suitable speech intonation, are also shared.

## **5.2 Spoken-voice performance and interpreting presentation quality**

The result that the interpreting students' spoken-voice skills had a significant impact on their overall interpreting quality is consistent with the interpreting research and experiments on the significant effects of performative aspects such as public-speaking skills, voice quality, stress-coping strategies and prosodic skills on overall interpreting quality (Ahrens, 2015; Cho & Roger, 2010; Flerov & Jacobs, 2016; Herbert, 1952; Horváth, 2017; Makarová, 1994, Naimushin, 2014). Additionally, correlations between students' spoken-voice performance and interpreting performance appeared more substantial in the post-intervention assessment than in the pre-intervention assessment, indicating that the intervention was effective and significantly improved overall interpreting quality.

The results indicate that training spoken-voice skills as a component of overall interpreting skills (see Subsection 2.1.3.2) may be able to facilitate automatic processing of vocal skills, freeing up processing power for other interpreting components that require additional processing capacity during interpreting tasks (Gile, 2009; Muller-Townsend, 2017), reducing the risks of cognitive overload and emotional stress induced by a focus on practising the entire interpreting task (Schneider, 1985). However, the provision of interpreting training may not automatically facilitate the acquisition of students' spoken-voice skills. These findings are consistent with those of other studies such as that of De Groot (2000). Her study made the finding that training in a particular

component of performance had positive effects on individual performance on the entire task and was thus more effective than whole-task training, particularly for novice interpreters. With heightened spoken-voice skills and self-regulation, students' spoken-voice production capacity may transfer from "non-automatic operation" (Gile, 2008, p. 165) to a conscious automatic state. As a skill that is honed, it becomes less disruptive to concurrent tasks happening in an interpreting assignment (Anderson, 1992). Accordingly, the self-regulatory capacity between spoken-voice production and other interpreting activities that require high cognitive processing capacity is no longer competing for limited resources (Gile, 2008; Matthei & Roeper, 1985; Schmeichel, 2007). As a result, students who have increased cognitive processing ability to cope with interpreting effort will generate higher quality interpreting presentations. As such, the study results lend support to spoken-voice training, which should be integrated into the interpreter training curriculum with the aim of achieving automaticity in spoken-voice skills and improving overall interpreting quality.

### **5.3 Highly improved spoken-voice categories and attributes**

The result that "spoken-voice dynamics" improved most notably in the STG and the ARG indicates that the intervention approaches utilised in this study, either spoken-voice instruction or awareness-raising, fulfilled the purpose of enhancing the students' vocal communication skills. Positive audience ratings were recorded when student interpreters were able to demonstrate more optimal performative skills through appropriate vocal modulation, use of intonation, rhythm and expressiveness, and speech behaviours such as appropriate eye-contact, vocal confidence, tone and physical appearance. This result supports studies (Kurz, 2009; Pöchhacker, 1994; Seleskovitch, 1986) in which authors have argued that interpreters' spoken voice serves as the chain of communication in the interpreting process that is important for the interpreters' audience. The study results also show "spoken-voice dynamics", in particular "vocal confidence", "expressiveness" and "audience connection", are particularly amenable to being improved due to the intervention approaches used in the study.

Apart from the “spoken-voice dynamics”, the results also show that the students improved considerably in some attributes of “spoken-voice quality”, including “speaking clarity”, “voice projection”, “modulation of pitch range” and “pace and pauses”. This result is consistent with previous studies regarding articulation, pitch, loudness and tempo as important prosodic features and suprasegmental properties of speech (Podesva, 2007; Roach, 2001) that strongly affect spoken-voice quality.

The output of correlational analysis (see Section 4.1.3) showing that these vocal attributes strongly affected the assessors’ impressions of the students’ interpreting quality demonstrates that the intervention on spoken-voice quality achieved the aim of training the students to manage voice quality in speech processing (Keller, 2004). This validates the importance for interpreters who work into their B language of receiving spoken-voice training because the quality of students’ production in English is always going to require more significant cognitive effort in interpreting into their B language, especially in relation to accent reduction, reducing tension, enhancing confidence and improving pitch range, pace and volume. At the same time, comments from the external assessors, who were native English speakers, reveal that despite their significant progress, interpreting students working in their B language still needed to improve their prosodic skills in order to accommodate an interpreting audience speaking English as its native language. Additionally, the result that the ARG improved significantly in “posture and alignment” and “pace and pauses” shows that these attributes can be improved if the student becomes aware of the relevant problems and consciously strives to overcome them.

Although “spoken-voice production” appears to be the least improved spoken-voice attribute, the results of this study show that this was fundamental in spoken-voice skills and that the significant improvement in the presentation of “vocal confidence” and “voice projection” and many other attributes was attributable to the training of students’ spoken-voice production, particularly “posture and alignment” and “breathing techniques”. This attests to the correlation between various spoken-voice production attributes and overall interpreting performances that recorded high levels in both

pre- and post-intervention assessments, suggesting the spoken-voice production impacted on the overall interpreting performance, although the effect may not be as conspicuous as with the other two categories.

The fact that several spoken-voice attributes, including “physical tension balance”, “movement” and “speech muscle tension release,” improved only marginally indicates that habitual tension remains a significant barrier to quality voice production, particularly for novice voice users, despite the students’ increased awareness of how to deal with tension in a difficult situation. The fact that all experimental groups improved minimally in these attributes reveals that restoring the physical tension balance is a slow process that requires sustained and long-term effort. Additionally, the fact that students in all 3 groups improved minimally in “resonance and timbre” demonstrates that improving vocal resonance may be difficult for interpreting students because it is a sensation-based skill that may not be easy to acquire until they master the correct technique and are able to feel it (John Henny Vocal Studio, 2021). As a result, more effective practices are necessary to heighten students’ awareness of their resonators, sympathetic vibrations and tone quality.

#### **5.4 Perspectives of multiple external assessors**

A similar conclusion from the multiple external assessors was that the STG students achieved considerable progress and outperformed the other groups while the non-training group made little improvement, which shows that the brief and regular voice-training sessions were most beneficial for the interpreting students. The varied perspectives between the voice trainer assessor and the other two assessors were mostly in their perceptions of the improvement of the LTG and the ARG. Although the ARG students did not develop much in their speaking-voice skills from the voice specialist’s perspective, their improvement was clear and satisfactory from the interpreting audience’s and interpreting instructor assessor’s viewpoints. Likewise, although in the vocal professional’s opinion the LTG students had significantly developed their spoken-voice skills, their growth was not clear from the interpreting audience’s and interpreting instructor assessor’s perspectives.

Similar perspectives of the interpreting instructor assessor and the lay assessor indicate that the interpreting instructor assessor may have employed an interpreting quality standard from a real interpreting market and from the end-users of interpreting services – the audience’s perception of how well they could comprehend messages conveyed by the interpreter through their vocal delivery (Pöchhacker, 1994; Shlesinger, 1994). The results show that the audience (the lay assessor) was delighted with the ARG’s improvement, particularly in “spoken-voice dynamism” in the post-assessment performance, including “voice confidence”, “expressiveness” and “audience connection”. This demonstrates that, while the ARG students’ spoken-voice production skills may not have improved from the voice trainer’s standpoint, their increased vocal awareness enabled them to make a more favourable impression on their audience.

Similarly, the results show that the lay assessor and the interpreting instructor assessor had unfavourable views of the LTG students’ improvement in spoken-voice dynamism despite the voice trainer assessor stating that the LTG students’ improvement was significant in spoken-voice production and quality. This suggests that the vocal assessment criteria employed by specialised voice trainer assessors for other vocal professionals need to be tailored according to the interpreting end-user’s perceptions in a real interpreting market.

On the other hand, the interpreting instructor assessor was more conservative towards the improvement in the students’ vocal performance in comparison with the lay assessor. An explanation can be inferred from one of Bühler’s (1986) experimental studies which contended that the demands on the quality of interpretation expressed by professional interpreters were consistently higher than those expressed by delegates. This means that interpreters may be inclined to impose higher quality standards on their own and their peers’ performance. Another interpretation in accounting for why the interpreting instructor assessor appeared to have responses different from the other two assessors is that the interpreting instructor assessor may have had a tendency to be overly focused on interlingual transfer and information content, rather than looking at information packaging (see Subsection 2.1.3.1) such as the spoken-voice quality, presentation and interactional



features during the interpreting assessment in this study. Some evidence for this comes from the fact that the interpreting instructor assessor provided few comments on some attributes, for example, “physical tension balance”, “movement”, “breathing techniques”, “release of speech muscle’ and “pitch modulation”, which may not be apparent in the end product (the content) of interpretation. This was contrary to the voice trainer assessor, who appeared to give more attention to the process of the students’ spoken-voice production. As mentioned, the voice trainer assessor was trained to check voice features, vocal performance and voice-using skills based on a specific vocal analysis method because the voice trainer assessor regarded the students’ vocal performance and skills more from the perspective of a specialised vocal professional and a trainer rather than an ordinary interpreting audience. Given the divergent perspectives of the multiple assessors and other research participants, this study suggests that interpreting pedagogy and training programs draw upon all views instead of taking sides.

In summary, Chapter 5 has discussed in detail the findings in relation to the 5 research questions and the implications for this research project. The spoken-voice intervention activities and metacognitive instruction yielded outstanding intervention effects, significantly raising the spoken-voice awareness and spoken-voice skills of the 3 intervention groups. The results have demonstrated that brief training sessions with short intervals were more effective than mass training sessions with long intervals and that explicit vocal instruction was more effective than awareness-raising intervention. The STG students outperformed the other two intervention groups due to a greater ability for self-regulated learning, which enabled them to develop automaticity in performance skills. Due to a lack of motivation for self-practice, audience awareness and vocal communication skills, the LTG students were unable to perform on a par with the STG students. While the ARG students displayed considerable improvements in spoken-voice awareness, their further growth was hampered by a lack of instructor-guided direction and kinaesthetic awareness. The results have shown that the training of spoken-voice awareness and skills as a component of interpreting production skills could improve students’ overall interpreting performance by

maximising automatic vocal skill processing and reducing cognitive overload. The students who participated in the intervention program demonstrated increased vocal dynamism and vocal quality, specifically in “vocal confidence”, “audience connection”, “expressiveness”, “speaking clarity” and “posture alignment”, indicating that the training of these attributes accomplished its goal of enhancing the students’ communication skills and voice quality. Although “spoken-voice production” was the least improved of the 3 spoken-voice categories, spoken-voice production skills were critical for students making considerable progress in the presentation of “vocal confidence” and other vocal qualities. This chapter has also discussed the importance of drawing upon all perspectives instead of taking sides when designing interpreting pedagogy and training programs.

## Chapter 6 Conclusion

### 6.1 Summary of results

This study has assessed the effectiveness of spoken-voice intervention approaches for interpreting students. A mixed-methods approach has been employed to explore the answers to 5 research questions. After comparing the spoken-voice performance of 4 groups of interpreting students, the statistical analysis has concluded that: 1) the 3 spoken-voice intervention groups recorded significantly higher levels of vocal performance at the post-intervention stage than the pre-intervention stage while, in contrast, the non-intervention group recorded little improvement; 2) the daily 20-minute spoken-voice training approach was more effective than the biweekly two-hour voice-training approach and the spoken-voice awareness-raising only intervention; and 3) spoken-voice performance significantly impacted on interpreting quality in the delivery stage.

According to the qualitative analysis, altogether 11 spoken-voice attributes improved considerably after the intervention activity: “vocal confidence”, “audience connection”, “expressiveness”, “intonation and tone”, “rhythmic and fluent rendition”, “posture and alignment”, “speaking clarity”, “breathing support”, “posture and alignment”, “voice projection” and “pace and pauses”. Among these, “vocal confidence” was the most prominent attribute contributing to the significant improvement of all 3 intervention groups.

The improvement of these attributes has demonstrated that certain aspects of spoken-voice dynamism, such as “vocal confidence,” “audience connection” and “expressiveness,” as well as “posture and alignment” and “pace and pauses,” are particularly amenable to improvement via either spoken-voice training activity or a spoken-voice awareness-raising approach. These skills can also be converted relatively easily into genuine improvements in the quality of students’ interpretation if they are aware of the underlying issues and consciously work to overcome them.

The results have also shown that the less favourable speaking habits included poor posture and alignment, shallow breathing, physical and speech muscle tension, mumbling, high and unchanging

pitch, inconsistent speaking pace and non-functional pauses, which negatively affected the students' interpreting performance. The spoken-voice attributes, such as "voice projection", "breathing skills", "speaking clarity" and "modulation of pitch range", were more effectively improved with explicit instruction. Moreover, the unchanged attributes, including "physical tension balancing" and "speech muscle tension", as well as "body movement" and "resonance and timbre", may benefit from long-term intervention, through either consistent spoken-voice training or conscious self-monitoring during an interpreting assignment.

The findings have also indicated that metacognitive instruction embedded in the spoken-voice intervention approach yielded a highly positive outcome, resulting in significant levels of improvement in the 3 intervention groups. Teaching specific spoken-voice skills in a structured and sequenced manner, with clear explanations of the skills and information structures and effective use of reflective activities, allowed the students to gradually internalise these techniques and use them to take control of their own learning and performance. The students who took part in spoken-voice training activities enhanced their kinaesthetic awareness and prosodic feature awareness, resulting in better voice production and quality. In particular, the students who attended daily voice-training sessions with a short and intense training regimen demonstrated high learner autonomy, improved inclusive awareness and speech behaviour, and better coping skills in a challenging situation, which allowed them to outperform other groups. The students who attended fortnightly intensive voice training also improved considerably in their vocal performance. However, these students may have been unduly focused on practising spoken-voice production skills in the workshops, ignoring the importance of communicating with the audience. In addition, due to long breaks between sessions and intense training content, they exhibited less motivation for self-practice outside of workshops, which may have hampered their improvement. These findings have indicated that successful learning requires a combination of metacognitive strategies, intervention activities and good training regimens from the instructor's side, along with self-regulated learning behaviour and

intrinsic motivation on the part of students. As Saxena (2020) mentioned, self-motivated students have a greater capacity for reflecting on and regulating their learning and emotions.

The students who attended spoken-voice awareness-raising activities showed significant improvement in vocal care awareness, prosodic feature awareness, audience awareness and vocal confidence. The MRI intervention improved their self-regulation, including better self-monitoring and self-evaluation skills and awareness in dealing with tension. However, this group of students showed a lack of kinaesthetic awareness and fundamental spoken-voice production skills due to a lack of instructor-guided spoken-voice training, which may have hindered their progress. Although the students demonstrated strong motivation in self-directed learning, they may not have been able to integrate the spoken-voice practice into actual interpreting practice and attain long-term and consistent improvement due to a lack of instructor guidance.

Hence, the 3 preliminary hypotheses presented in Section 1.2 are addressed as follows.

Firstly, interpreting students who receive spoken-voice intervention can make substantial progress. A student group that is not exposed to spoken-voice intervention, on the other hand, may not make improvements in spoken-voice skills in relation to interpreting delivery and improved spoken-voice dynamics.

Secondly, explicit spoken-voice training instruction, a metacognitive teaching approach and short and daily sessions may be the most effective method for improving interpreting students' skills in spoken-voice production, quality and dynamics.

Thirdly, raising vocal awareness of the spoken voice even without providing formal spoken-voice training may enable interpreting students to consciously monitor their interpreting delivery. However, because the students are lacking in fundamental spoken-voice skills, supervision and constructive feedback from the voice instructor, their improvement may be limited.

## 6.2 Pedagogical and methodological implications

Based on the findings, this study draws the following conclusions in relation to both interpreting training and spoken-voice training.

1. The finding that interpreting students who received spoken-voice intervention made significant progress while those who did not receive such intervention achieved little progress suggests that, while improving spoken-voice skills can significantly improve overall interpreting performance in terms of delivery, interpreting training alone is not necessarily an intervention that leads to improvement in students' spoken-voice skills. This may be bad news for interpreters who maintain a "get-by" attitude, believing that the whole task of interpreting training will naturally improve their interpreting delivery or that speaking-voice training is a waste of time and work, as they may struggle to achieve a higher level of interpreting performance without making effort to improve their spoken-voice skills.

2. Sporadic mode (see Subsection 2.2.4.1) is an effective training regimen for beginner students because it promotes concentration, voice awareness and a higher level of motivation for self-practice both during and between sessions. In comparison, spoken-voice training in block mode is less effective due to the heavy training content and lengthy intervals between sessions, making it difficult for interpreting students to focus, undertake consistent self-practice outside of workshops and receive prompt feedback from the vocal instructor. Additionally, the sporadic mode facilitates students' integration of acquired spoken-voice skills into daily interpreting skills, but the block mode may hamper the students in transferring voice skills to actual interpreting practice as they may regard the voice training and the interpreting training as two separate training activities.

3. The students may have strengthened some spoken-voice attributes, particularly attributes of vocal dynamism, through increased audience awareness and mindfully monitoring their spoken-voice performance without receiving explicit spoken-voice instruction. However,

formal spoken-voice instruction and consistent self-practice may be required to develop spoken-voice production skills and acquire command of prosodic features over time. This may serve as a reminder call for those who assume that simply having vocal awareness will improve their performance and so do not wish to devote significant effort to practising spoken-voice skills.

4. The 3 spoken-voice categories are interconnected, although they function separately. Spoken-voice production serves as a foundation for the other two categories. Spoken-voice dynamism may be improved efficiently by improving students' vocal awareness and their ability to consciously monitor their own performance if explicit spoken-voice instruction is absent. Spoken-voice production and prosodic features, on the other hand, may improve more significantly with explicit spoken-voice instruction and sustained self-practice outside workshops, particularly when interpreting students work into their B language.

5. Metacognitive instruction included in spoken-voice intervention significantly increases students' metacognitive awareness and self-regulation, as seen through increased self-monitoring and self-reflection on their performance and emotion, as well as increased motivation for self-practice. Thus, employing metacognitive instruction benefits not only spoken-voice skills but also interpretation skills.

Based on the conclusions described above, this study recommends incorporating metacognitive instruction into interpreting training. In this sense, metacognitive instruction is not only beneficial for increasing students' speaking-voice skills but may also be good for enhancing students' interpretation skills.

In addition, the program of interventions that formed the teaching/learning guide used in this study could be used as a guide and could be modified to train interpreting students' vocal performance. An instructor could select some of the interpreter-tailored spoken-voice training activities at the beginning of each interpreting workshop as a warm-up exercise before starting interpreting

activities. In particular, an instructor could choose some voice-production exercises that the students received well in this study, such as breathing exercises and vocal warm-up exercises, resonance exercises and articulation exercises, which may lay a good foundation for students' spoken-voice production skills. In this way, students could apply spoken-voice training exercises to actual interpreting practice and be constantly prompted to take care of their vocal organs in demanding workloads and produce the optimal sound effects in interpreting delivery.

Alternatively, in a case where spoken-voice training is not available, an instructor could select some activities to enhance students' prosodic feature awareness and audience awareness to improve the students' voice quality and vocal dynamism. Given that the results have indicated that "physical tension balance", "resonance and timbre" and "body movement" remained unchanged across all groups following the intensive intervention, an instructor may need to provide consistent support and awareness-raising in relation to these. It is strongly encouraged that an instructor applies metacognitive strategies in an intervention to enhance students' metacognitive awareness and self-regulation capacity, which will benefit the students' spoken-voice skills and interpreting skills in the long run. An instructor could also use a list of spoken voice features to aid students in becoming familiar with important spoken-voice features, monitoring their own practice and providing feedback to their peers, and setting specific and measurable individual learning goals.

Encouragingly, spoken-voice skills have been included in the rhetorical skills in the latest NAATI Certified Conference Interpreter Test Assessment Rubrics and are specified as one of the 3 important transfer competencies for candidates to pass the assessment. The rubrics are classified into two parts – "Transfer competency" and "Language competency". Rhetorical skills along with meaning transfer skills and application of interpreting mode have been classified into "Transfer competency". The pass requirement for rhetorical skills for candidates is specified as "consistently/mostly demonstrates use of rhetorical technique appropriate to the specifications provided in the interpreting brief. Excellent voice projection. Demonstrates clear pronunciation, fluent delivery, good tone, voice quality and volume" (NAATI, 2020). This study provides evidence



to support NAATI's rubrics that include spoken voice as a vital part of overall interpreting quality. With such an assessment rubric, this means that the assessment of spoken-voice skills and of overall quality of interpreting can be mutually informed. In other words, the spoken-voice performance can inform the overall interpreting quality and the interpreting quality can encompass attributes that include spoken-voice skills. As such, the study results lend support to spoken-voice training, which should be integrated into the interpreter training curriculum with the aim of improving overall professional interpreting quality.

This research presents the empirical findings of a spoken-voice intervention for interpreting students, bridging the gap between spoken-voice training and interpreting teaching. Furthermore, the study has applied a methodologically innovative approach by examining how both a spoken-voice intervention approach and a metacognitive instruction method can be tested in interpreting research. These spoken-voice interventions have been found to be useful tools for improving interpreting students' vocal presentation and interpreting delivery. Furthermore, the study makes a valuable contribution by narrowing the gap between existing spoken-voice pedagogy and interpreting pedagogy by providing specific spoken-voice activities integrated into interpreting activities. These spoken-voice activities contribute to the body of research in interpreting pedagogy on those interventions that may assist interpreting instructors to better understand the vocal challenges of their interpreting students and guide them to improve their interpreting delivery.

### **6.3 Limitations and future research directions**

While this study has made some preliminary contributions to understanding the effects of spoken-voice intervention on interpreting students, it is equally as important to note the study's limitations. Firstly, the data for the study are based on only a small number of Mandarin–English first year interpreting students and hence the study findings can be generalised only tentatively across interpreting students working with other language pairs or in other working contexts. Secondly, because all student participants in this study had Mandarin as their native language, the findings may not be generalisable to Mandarin–English interpreters who have English as their native

language and are learning Mandarin as a second or foreign language. Thirdly, because the co-voice coach Ms P was unavailable to travel to China at the time, the researcher had to deliver the spoken-voice training workshops at the Suzhou campus without Ms P's presence. This may potentially have affected the training quality for the STG. Fourthly, when collecting data from the student participants from 2013 to 2015, the MITS was a one-and-a-half year (3-semester) translation and interpretation postgraduate program that had an intake once a year and was upgraded to a two-year program from 2017. The number of students who took up places in the Mandarin–English stream in each calendar year was modest, between 5 and 12 students, which resulted in a small intake of student participants in this study. Therefore, there were only 5 participants in each group in the study and hence the performance of each participant could substantially influence the generalisability of the data in the study. Finally, as the study was only conducted in the CI mode, further research is recommended to ascertain the effect of spoken-voice intervention on interpreters working in other language pairs in other interpreting modes and contexts including SI and RI.

As the coronavirus has been quickly spreading all over the world, many interpreting assignments have been moved online. The threat and fear of dehumanisation in the interpreting sector through the rapidly increasing use of voice-recognition technology that, combined with automatic interpretation tools, can promise “instant automatic interpretation” even if the quality is questionable may be a response to this. Superior spoken-voice quality can be a compelling reason for a client to choose a person over a machine, as a machine is unlikely to produce vocal production with the same quality or dynamism as a human voice. Undertaking spoken-voice training is the key. As this study describes intervention results in CI mode from Mandarin to English, it would be fascinating to see the training completed with different language pairings, directionalities and interpreting modes. In addition, it may be helpful to compare interpreters' spoken-voice strategy use in SI with their strategy use in CI to understand how different interpreting modes and language directions affect interpreters' spoken-voice strategy use. Finally, it would be beneficial to extend the investigation from student interpreters to professional interpreters using the expert-novice research

paradigm to see how different the training effect is for student interpreters compared with professional interpreters and how the results could be integrated into interpreting teaching pedagogy.

## References

- Abrahamson, M. (1983). *Social research methods*. Englewood Cliffs, NJ: Prentice Hall.
- Adams, C. (1979). *English speech rhythm and the foreign learner*. The Hague: De Gruyter Mouton.
- Ahrens, B. (2004). *Prosodie beim Simultandolmetschen*. Frankfurt: Peter Lang.
- Ahrens, B. (2005). Prosodic phenomena in simultaneous interpreting: A conceptual approach and its practical application. *Interpreting*, 7(1): 51-76.
- Ahrens, B., Kalderon, E., Krick, C. M., Reith, W., Gile, D., Hansen, G., & Pokorn, N. K. (2010). fMRI for exploring simultaneous interpreting. In D. Gile, G. Hansen & N. Pokorn (Eds.) *Why translation studies matter*. (pp. 237-248) Amsterdam/Philadelphia: John Benjamins.
- Ahrens, B. (2015). Intonation. In F. Pöchhacker (Ed.) *Routledge Encyclopedia of Interpreting Studies*. (pp. 212-214) Abingdon, Oxon: Routledge.
- AIIC. (2016). *Practical guide for professional conference interpreters*. Retrieved from [https://aiic.org/document/547/AIICWebzine\\_Apr2004\\_2\\_Practical\\_guide\\_for\\_professional\\_conference\\_interpreters\\_EN.pdf](https://aiic.org/document/547/AIICWebzine_Apr2004_2_Practical_guide_for_professional_conference_interpreters_EN.pdf)
- AIIC [International Association of Conference Interpreters] (2019) *Glossary*. Retrieved from <https://aiic.org/site/world/conference/glossary#:~:text=%E2%94%80%20A%20Language,A%20Language,both%20consecutive%20and%20simultaneous%20interpretation>
- Alcantara, P. D. (1997). *Indirect procedure – A musician’s guide to the Alexander technique*. NY: Oxford.
- Alexander, F. M. (1932). The use of the self. *British Medical Journal*, 1(3728), 1149.
- Allwright, R. (1984). Why don’t learners learn what teachers teach? In D. Singleton & D. Little (Eds.), *Language learning in formal and informal contexts* (p. 318). Dublin: IRAAL.
- Amini, M. (2015). Malaysian and non-Malaysian users’ quality expectations in conference interpreting. *Journal of Translator Education and Translation Studies*, 2(4), 22–44.
- Amini, M. (2016). *Speaking speed vs. consecutive interpreting speed: A case of English–Persian*. *Translation Journal*. Retrieved from <https://translationjournal.net/January-2016/speaking-speed-vs-consecutive-interpreting-speed-a-case-of-english-persian.htm>
- Anderson, D., Nashon, S. M., & Thomas, G. P. (2009). Evolution of research methods for probing and understanding metacognition. *Research in Science Education*, 39(2), 181–195.

- Andrade, D. F., Heuer, R., Hockstein, N. E., Castro, E., Spiegel, J. R., & Sataloff, R. T. (2000). The frequency of hard glottal attacks in patients with muscle tension dysphonia, unilateral benign masses and bilateral benign masses. *Journal of voice : official journal of the Voice Foundation*, 14(2), 240–246.
- Anderson, J. (1995). *Cognitive psychology and its implications*. NY: W. H. Freeman.
- Anderson, J. R. (1992). Automaticity and the ACT theory. *American Journal of Psychology*, 105(2), 165–180.
- Anderson, N. J. (2002). *The role of metacognition in second language teaching and learning*. ERIC Digest.
- Anderson, R., Klofstad, C., Mayew, W., & Venkatachalam, M. (2014). Vocal fry may undermine the success of young women in the labor market. *PLoS One*, 9(5), E97506.
- Anderson, V. A. (1977). *Training the speaking voice* (3rd ed.). NY: Oxford University Press.
- APA Work Group of The Board of Educational Affairs (1997, November). *Learner-centered psychological principles: A framework for school reform and redesign*. Washington, DC: American Psychological Association.
- APA Task Force on Psychology in Education (1993, January). *Learner-centered psychological principles: Guidelines for school redesign and reform*. Washington, DC: American Psychological Association and Mid-Continent Regional Educational Laboratory.
- Appel, E. (2019). *5 issues when dealing with non-English speaking customers*. Retrieved from <https://www.allgraduates.com.au/5-issues-when-dealing-with-non-english-speaking-customers/>
- Appelman, D. R. (1974). *The science of vocal pedagogy*. Bloomington: Indiana University Press. (Original work published 1967)
- Arboleda, B. M. W., & Frederick, A. L. (2008). Considerations for maintenance of postural alignment for voice production. *Journal of Voice*, 22(1), 90–99.
- Archie, W. C. (1949). The interpreter in peace and war. *The French Review*, 22(3), 249–255.
- Arumí, R. M. (2012). Problems and strategies in consecutive interpreting: A pilot study at two different stages of interpreter training. *Meta*, 57(3), 812–835.
- ASTM International. (2008). *Standard guide for language interpretation services 1*, 1–14.

- Bacher, K., Baltrus, A., Barrie, B., Bliss, K., Cardea, D., Chandler, L., and Lacombe, B. (2007, March 1). *Foundations of Interpretation*. Retrieved from <https://www.nps.gov/idp/interp/101/foundationscurriculum.pdf>
- Baird, J. R. (1986). Improving learning through enhanced metacognition: A classroom study. *European Journal of Science Education*, 8(3), 263–282.
- Baleghizadeh, S., & Derakhshesh, A. (2017). Measuring the effectiveness of explicit and implicit instruction through explicit and implicit measures. *Iranian Journal of Applied Linguistics*, 20(1), 81–111.
- Baker, S., Weinrich, B., Bevington, M., Schroth, K., & Schroeder, E. (2008). The effect of task type on fundamental frequency in children. *International Journal of Pediatric Otorhinolaryngology*, 72, 885–889.
- Banai, K., Ortiz, J. A., Oppenheimer, J. D., & Wright, B. A. (2010). Learning two things at once: Differential constraints on the acquisition and consolidation of perceptual learning. *Neuroscience*, 165(2), 436–444.
- Barik, H. C. (1975). Simultaneous interpretation: Qualitative and linguistic data. *Language and Speech*, 18(3), 272–297.
- Bartłomiejczyk, M. (2014). Interpreting quality as perceived by trainee interpreters. *The Interpreter and Translator Trainer*, 1(2), 247–267.
- Barton, R. (2006). *Acting: Onstage and off* (4th ed.). Belmont, CA: Thomson, Wadsworth.
- Barton, R. (1997). Voice in a visual world. *The Vocal Vision*. New York: Applause Books.
- Barrows, H., & Tamblyn, Robyn. (1980). Problem-based learning: *An approach to medical education*. New York: Springer Pub.
- Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego, CA: Academic Press.
- Beardmore, H. B. (1986). *Bilingualism: Basic principles* (2nd ed.). Clevedon: Multilingual Matters.
- Bellon-Harn, M. L. (2011). Targeting prosody: A case study of an adolescent. *Communication Disorders Quarterly*, 32(2), 109–117.
- Benge, T. J. (1996). *Movement utilized by conductors in the simulation of expression and musicianship* (doctoral dissertation). University of Southern California.

- Bereiter, C., & Scardamalia, M. (1993). *Surpassing ourselves: An inquiry into the nature and implications of expertise*. Chicago, IL: Open Court.
- Bergamin, J. A. (2017). Being-in-the-flow: Expert coping as beyond both thought and automaticity. *Phenomenology and the Cognitive Sciences*, 16(3), 403–424.
- Berry, C. (1973). *Voice and the actor*. London: Harrap.
- Berry, C. (1987). *The actor and his text*. London: Virgin.
- Berry, C. (2001). *The text in action: A definitive guide to exploring text in rehearsal for actors and directors*. London. Virgin Books.
- Bolinger, D. (1986). *Intonation and its parts: Melody in spoken English*. Stanford: Stanford University Press.
- Boone, D. R., & McFarlane, S. C. (1993). A critical view of the yawn: Sigh as a voice therapy technique. *Journal of Voice*, 7, 75–80.
- Borkowski, J. G., Carr, M., & Pressley, M. (1987). “Spontaneous” strategy use: Perspectives from metacognitive theory. *Intelligence*, 11(1), 61–75.
- Bouma, G. (2000). *The research process* (4th ed.). Melbourne: Oxford University Press.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. Sage Publications, Inc.
- Bracken, D. W., Dalton, M. A., Jako, R. A., McCauley, C. D., Pollman, V. A., & McCauley, C. D. (1997). *Should 360-degree feedback be only used for developmental purposes?* Greensboro, NC: Center for Creative Leadership.
- Braun, S. (2015). Remote interpreting. In F. Pöchhacker (Ed.) *Routledge Encyclopedia of Interpreting Studies*. (pp. 346-348). Abingdon, Oxon: Routledge.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Brief, A. P., & Aldag, R. J. (1977). The intrinsic–extrinsic dichotomy: Toward conceptual clarity. *Academy of Management Review*, 2(3), 496–500.
- Brown, O. (1996). *Discover your voice: How to develop healthy voice habits*. San Diego, CA: Singular.

- Bühler, H. (1986). Linguistic (semantic) and extra-linguistic (pragmatic) criteria for the evaluation of conference interpretation and interpreters. *Multilingua*, 5(4), 231–235.
- Cañada, M. D., & Arumí, M. (2012). Self-regulating activity: Use of metacognitive guides in the interpreting classroom. *Educational Research and Evaluation*, 18(3), 245–264.
- Case, J., Seyfarth, S., & Levi, Su, V. (2018). Does implicit voice learning improve spoken language processing? Implications for clinical practice. *Journal of Speech, Language, and Hearing Research*, 61(5), 1251–1260.
- Cecot, M. (2001). Pauses in simultaneous interpretation: A contrastive analysis of professional interpreters' performances. *Interpreters' Newsletter*, 11, 63–85.
- Chapman, J. (2006). *Singing and teaching: A holistic approach to classical voice*. San Diego, CA: Plural Publishing, Inc.
- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535–544.
- Chang, I. Y., & Chang, W. Y. (2012). The effect of student learning motivation on learning satisfaction. *International Journal of Organizational Innovation*, 4(3), 281.
- Chen, C.F., Fan, C.Y. & Lin, H.P. (1996). *A new perspective on teaching English pronunciation: Rhythm*. ERIC Document Reproduction Service (ED397638), 1-16. Retrieved from <https://eric.ed.gov/?id=ED397638>
- Cho, J., & Roger, P. (2010). Improving interpreting performance through theatrical training. *The Interpreter and Translator Trainer*, 4(2), 151–171.
- Cho, J. Y., & Lee, E. (2014). Reducing confusion about grounded theory and qualitative content analysis: Similarities and differences. *The Qualitative Report*, 19(32), 1–20. Retrieved from <http://nsuworks.nova.edu/tqr/vol19/iss32/2>
- Chong, Y. [莊旖瀾]. (2012). *Vocal characteristics of English and Mandarin produced by Mandarin-English and English-Mandarin bilingual speakers: a long-term average spectral analysis*. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.
- Christodoulides, G. (2013). Prosodic features of simultaneous interpreting. In *Proceedings of the Prosody–Discourse Interface Conference* (pp. 33–37).
- Clare-Newman, C. (n.d.). *Most of Us Live on Autopilot*. Retrieved from <http://www.alexandertechnique.com/articles2/autopilot/>
- Clark, B. (1988). *Growing up gifted* (3rd ed.). Columbus, OH: Merrill.



- Cleveland, J. N., Murphy, K. R., & Williams, R. E. (1989). Multiple uses of performance appraisal: Prevalence and correlates. *Journal of Applied Psychology*, 74(1), 130–135.
- Cohen, J. W. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Collados Aís, A. (1998) *La evaluación de la calidad en interpretación simultánea. Laimportancia de la comunicación no verbal*, Granada: Comares.
- (2008) Evaluación de la calidad en interpretación simultánea: contrastes de exposición e inferencias emocionales. Evaluación de la evaluación, in Hansen, G., Chesterman, A. and Gerzymisch-Arbogast, H. (eds) *Efforts and Models in Interpreting and Translation Research*. Amsterdam/Philadelphia: John Benjamins, pp. 193-214.
- Collados Aís, Á.; Pradas Macías, E.; Stévaux, E. & García Becerra, O. (2007) (ed.) *La evaluación de la calidad en interpretación simultánea: parámetros de incidencia*, Granada, Comares.
- Colton, R., & Casper, J. K. (1996). *Understanding voice problems: A physiological perspective for diagnosis and treatment* (2nd ed.). Baltimore, MD: Williams & Wilkins.
- Cosmides, L. (1983). Invariances in the acoustic expression of emotion during speech. *Journal of Experimental Psychology: Human Perception and Performance*, 9(6):864–881.
- Couper-Kuhlen, E., & Selting, M. (Eds.) (1996). *Prosody in conversation*. Cambridge, UK: Cambridge University Press.
- Creswell, J., & Poth, Cheryl N. (2018). *Qualitative inquiry & research design: Choosing among five approaches* (Fourth ed.).
- Crowley, B. P., & Delfico, J. F. (1996). *Content analysis: A methodology for structuring and analyzing written material*. United States General Accounting Office (GAO), Program Evaluation and Methodology Division. Retrieved from <http://www.gao.gov/assets/80/76281.pdf>
- Cruttenden, A. (1997). *Intonation*. Cambridge, UK: Cambridge University Press.
- Crystal, D. (2005). *How language works*. London: Penguin Books.
- Curtis, D. (2000). Competition policy and the future of higher education institutions in Australia. *Proceedings AARE Conference*. Retrieved from <http://www.aare.edu.au/00pap/cur00137.htm>

- Cutler, A. (1987). *Components of prosodic effects in speech recognition*. Proceedings of the Eleventh International Congress of Phonetic Sciences (pp. 84-87).
- Cutler, A. (2012). *Native listening: Language experience and the recognition of spoken words*. Cambridge, MA: MIT Press.
- Dam, N. (2013). *Inside the learning brain*. TD Magazine (8 April 2013) Retrieved from: <https://www.td.org/magazines/td-magazine/inside-the-learning-brain>
- Davidson, J. W. (1993). Visual perception of performance manner in the movements of solo musicians. *Psychology of Music*, 21(2), 103–113.
- Davidson, J. W. (2001). The role of the body in the production and perception of solo vocal performance: A case study of Annie Lennox. *Musicae Scientiae*, 5(2), 235–256.
- Davies, W. M. (2006). Intensive teaching formats: A review. *Issues in Educational Research*, 16(1), 1–20.
- Davis, S. (n.d.). *How does your posture affect your vocal potential?* Retrieved from [https://www.voxfitconsulting.com/blog/how-does-your-posture-affect-your-vocal-potential#:~:text=Poor%20posture%20results%20in%20reduced,muscle%20tension%20dysphonia%20\(MTD\)](https://www.voxfitconsulting.com/blog/how-does-your-posture-affect-your-vocal-potential#:~:text=Poor%20posture%20results%20in%20reduced,muscle%20tension%20dysphonia%20(MTD))
- De Jong, N. & Perfetti, C. A. (2011). *Fluency Training in the ESL Classroom: An Experimental Study of Fluency Development and Proceduralization*. *Language Learning*, 61(2), 533-568.
- De Groot, A. M.B. (2000). A complex-skill approach to translation and interpreting. In S Tirkkonen-Condit and R. Jääskeläinen *Tapping and Mapping the Processes of Translation and Interpreting: Outlooks on empirical research*. (pp. 53-68.) Amsterdam: John Benjamins.
- Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior* Boston, MA: Springer.
- Deci, E. L., & Ryan, R. M. (2016). Optimizing students' motivation in the era of testing and pressure: A self-determination theory perspective. In W.C. Liu, J.C.K. Wang, & R.M. Ryan (Eds.) *Building autonomous learners: Perspectives from research and practice using self-determination theory* (pp. 9–29). Singapore: Springer.
- Déjean le Féal, K. (1990). Some thoughts on the evaluation of simultaneous interpretation. In D. Bowen & M. Bowen (Eds.), *Interpreting – Yesterday, today and tomorrow* (pp. 154–160). Binghamton: State University of New York.

- deMarrais, K. B., & Lapan, S. D. (2003). Qualitative interview studies: Learning through experience. In Kathleen B. deMarrais & Stephen D. Lapan (Eds.), *Foundations for Research: (pp. 67-84)*. Routledge.
- Department of Health & Human Services. (2015). *Breathing to reduce stress*. Retrieved from <https://www.betterhealth.vic.gov.au/health/HealthyLiving/breathing-to-reduce-stress>
- DeVellis, R. F. (2012). *Scale development: Theory and applications* (3rd ed.). Thousand Oaks, CA: Sage.
- Dillinger, M. (1994) Comprehension during interpreting: What do interpreters know that bilinguals don't? In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation*. (pp. 155–189). Amsterdam: John Benjamins,
- Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the Conceptual Lens on Metacognition, Self-regulation, and Self-regulated Learning. *Educational Psychology Review, 20*(4), 391-409.
- Dlugan, A. (2012). *Speech pauses: 12 techniques to speak volumes with your silence*. Retrieved from <http://sixminutes.dlugan.com/pause-speech/>
- Donkersgoed, M. G. V. (2016). *Using the Metacognition and Reflective Inquiry Method as a Metacognitive Stimulant for Gifted Students*. Retrieved from [http://dspace.library.uu.nl › bitstream](http://dspace.library.uu.nl/bitstream)
- Donovan, C. (2011). Ethics in the teaching of conference interpreting. *The Interpreter and Translator Trainer, 5*(1), 109-128.
- Donovan, J. J., & Radosevich, D. J. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Psychology, 84*(5), 795.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford, UK: Oxford University Press.
- Dunlap, J. C. (2005). Problem-based learning and self-efficacy: How a capstone course prepares students for a profession. *Educational Technology Research and Development, 53*(1), 65–83.
- Edwards, R., & Temple, B., & Alexander, C. (2005). Users' experiences of interpreters. *Interpreting, 7*(1), 77–95.
- Eggen, P., & Kauchak, D. (1994). *Educational psychology: Classroom connections*. New York: Macmillan College.

- Ellis, N. (1993). Rules and instances in foreign language learning: Interactions of explicit and implicit knowledge. *European Journal of Cognitive Psychology*, 5(3), 289-318.
- Ellis, R. (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Multilingual Matters.
- El-Menoufy, A. (1988). Intonation and meaning in spontaneous discourse. In M. Cummings Benson & W. S. Greaves (Eds.), *Linguistics in a systemic perspective* (pp. 1–26). Amsterdam: John Benjamins.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115.
- Emerick, L. J. (1992). Academic underachievement among the gifted: Students' perceptions of factors that reverse the pattern. *Gifted Child Quarterly*, 36(3), 140–146.
- Encyclopaedia Britannica. (1998). *Pitch*. Retrieved from <https://www.britannica.com/topic/pitch-speech>
- Eom, S. (2019). The effects of student motivation and self-regulated learning strategies on student's perceived E-learning outcomes and satisfaction. *Journal of Higher Education Theory & Practice*, 19(7), 29-42.
- Esen, S. (2018). *Pitch in English pronunciation: Definition and examples*. Retrieved from <https://owlcation.com/humanities/Pitch-Definition-And-Examples-Of-Pitch-In-English-Pronunciation>
- Evans, G. (2016). *The power of pace and pauses in a speech*. Retrieved from <http://www.greatspeech.co.uk/power-pace-pauses-speech.html>
- Evans, J., Bell, J. L., Sweeney, A. E., Morgan, J. I., & Kelly, H. M. (2010). Confidence in critical care nursing. *Nursing Science Quarterly*, 23(4), 334–340.
- Feltz, D. L. (2007). Self-confidence and sports performance. *Studies*, 33(41), 278–291.
- Ferris, E. (2018). *Posture and breathing: The physiological effects of shallow breaths*. The Breath Effect. <https://www.thebreatheffect.com/posture-breathing-physiological-effects/>
- Fidan, M. B. (2017). *The third person in the room: The impact of the interpreter on the counselling process with non-English speaking clients*. University of Leicester. <https://hdl.handle.net/2381/40907>

- Flavell, J. H. (1976). Metacognitive aspects of problem-solving. In L. B. Resnick (Ed.), *Perspectives on the development of memory and cognition* (pp. 231–235). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *The American Psychologist*, *34*(10), 906-911.
- Flavell, J. H. (1987). Speculation about the nature and development of metacognition. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 21–29). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fleenor, J. W., & Prince, J. M. (1997). *Using 360-degree feedback in organizations*. Center for Creative Leadership, Greensboro, NC.
- Fleming, N. (2006). *Teaching and learning styles: VARK strategies* (2nd ed.). Christchurch, N.Z.: N.D. Fleming.
- Flerov, C., & Jacobs, M. (2016). *Improving the interpreter's voice*. London, England: Lulu Publishing Services.
- Freeman, J. L. (2010). Identifying movement pattern and severity of associated pain in sign language interpreters. *College Student Journal*, *44*(2), 1–7.
- Fujii, S., & Wan, C. Y. (2014). The role of rhythm in speech and language rehabilitation: The SEP hypothesis. *Frontiers in Human Neuroscience*, *8*, 777.
- Gardiner, M. D. (1957). *The principles of exercise therapy*. London: G. Bell and Sons.
- Garner, R. & Alexander, P. A. (1989). Metacognition: Answered and unanswered questions. *Educational Psychologist*, *24*(2), 143-158.
- Garzone, G. (2002). Quality and norms in interpretation. In G. Garzone & M. Viezzi (Eds.), *Interpreting in the 21st century: Challenges and opportunities* (pp. 107–119). Amsterdam; Philadelphia: John Benjamins.
- Garzone G. (2003) Reliability of quality criteria evaluation in survey research (Eds.), *La evaluación de la calidad en interpretación: investigación* (pp. 23-30 ). Granada: Comares.
- Gastaldo, D., Magalhães, L., Carrasco, C., & Davy, C. (2012). *Body-map storytelling as research: Methodological considerations for telling the stories of undocumented workers through body mapping*. Retrieved from <http://www.migrationhealth.ca/undocumented-workers-ontario/body-mapping>

- Gates, L. (2011). *Voice for performance* (2nd ed.). Milwaukee, WI: Limelight Editions.
- Gebhard, S. (2000). Learning to tell the difference. *AIIC Bulletin*, XXVII (2), 7.
- Gentile, A. (1997). Community interpreting or not? Practices, standards and accreditation. In *The critical link: Interpreters in the community* (p. 109). Amsterdam/ Philadelphia: John Benjamins.
- Gibson, E. J. (1963). Perceptual learning. *Annual Review of Psychology*, 14(1), 29–56.
- Gile, D. (1989). *La communication linguistique en réunion multilingue: Les difficultés de la transmission informationnelle en interprétation simultanée* (thèse). Université Paris III.
- Gile, D. (1990). L'évaluation de la qualité de l'interprétation par les délégués: Une étude de cas. *Interpreters' Newsletter*, 3, 66–71.
- Gile, D. (1991) Methodological Aspects of Interpretation (and Translation) Research, *Target* 3(2), 153-174.
- Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (rev. ed., pp. 25–51). Amsterdam: John Benjamins Publishing Company.
- Gillies, A. (2004). *Conference interpreting. A new students' companion*. Kraków: Tertium.
- Glenn, E. C., Glenn, P. J., & Forman, S. H. (1984). *Your Voice and Articulation*. United States of America: Macmillan Publishing Company.
- Goedert, K. M., & Miller, J. (2008). Spacing practice sessions across days earlier rather than later in training improves performance of a visuomotor skill. *Experimental Brain Research*, 189(2), 189–197.
- Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585–612.
- Gomel, R. (2017). The relationship between motivation and self-efficacy. *International Journal of Advanced Multidisciplinary Research and Review*, 5(8), 1–11.
- Gopalan, V., Bakar, J. A. A., Zulkifli, A. N., Alwi, A., & Mat, R. C. (2017). A review of the motivation theories in learning. In *AIP Conference Proceedings*, 1891(1), 020043.
- Gorton, A. (2012). 'B' Language Interpreting: The interpreter's perspective. *FORUM*, 10(2), 61-88.
- Graeme. (2016, December 12). *Posture*. Retrieved March 11, 2022, from <https://www.sportscarephysio.com.au/posture/#:~:text=Postural%20alignment%20is%20the%20composite,prerequisite%20for%20optimal%20movement%20function>

- Grant, S. (2011). An approach to the affective dimension of speaking. *Parrhesia*, 13, 112–125.
- Gray, J. (1990). *Your guide to the Alexander technique*. NY: St. Martin's Press.
- Grbić, N. (2008). Constructing interpreting quality. *Interpreting*, 10(2), 232–257.
- Greer, D. A., Cathcart, A., & Neale, L. (2016). Helping doctoral students teach: Transitioning to early career academia through cognitive apprenticeship. *Higher Education Research & Development*, 35(4), 712–726.
- Gudgon, A. (2009). *Voice coaching for interpreters*. AIIC Webzine, 2009. Retrieved from <http://aiic.net/page/3127/voice-coaching-for-interpreters/lang/1>.
- Guerrero, L. K., & Floyd, K. (2006). *Nonverbal communication in close relationships*. Mahwah: Taylor and Francis.
- Gut, U., Trouvain, J., & Barry, W. J. (2007). Bridging research on phonetic descriptions with knowledge from teaching practice: The case of prosody in non-native speech. In *Non-native prosody* (pp. 3–24). Berlin, New York: De Gruyter Mouton.
- Hacker, D. J. (1998). Metacognition: Definitions and empirical foundations. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 1-23). Mahwah, NJ: Erlbaum
- Hahn, E., Lomas, C., Hargis, D., & Vandraegen, D. (1957). *Basic voice training for speech*. New York: McGraw.
- Halliday, L. F., Taylor, J. L., Edmondson-Jones, A. M., & Moore, D. R. (2008). Frequency discrimination learning in children. *Journal of the Acoustical Society of America*, 123(6), 4393–4402.
- Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Arnold.
- Halpern, D. F. (1996). *Thought and knowledge: An introduction to critical thinking* (3rd ed.). Mahwah, N.J.: L. Erlbaum Associates.
- Hammarberg, B., Fritzell, B., Gaufin, J., Sundberg, J., & Wedin, L. (1980). Perceptual and acoustic correlates of abnormal voice qualities. *Acta Oto-laryngologica*, 90(1–6), 441–451.
- Happe, F. (2001) Automaticity of action, psychology of. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences*. Retrieved from <https://scholar.harvard.edu/files/dwegner/files/wheatleywegner.pdf>

- Hargie, O. (2011). *Skilled interpersonal communication: Research, theory and practice* (5th ed.). Routledge/Taylor & Francis Group.
- Hargie, O. (2016). The importance of communication for organisational effectiveness. *Psicologia do Trabalho e das Organizações*, 15–32.
- Hauptmann, B., & Karni, A. (2002). From primed to learn: The saturation of repetition priming and the induction of long-term memory. *Cognitive Brain Research*, 13(3), 313–322.
- Hauptmann, B., Reinhart, E., Brandt, S. A., & Karni, A. (2005). The predictive value of the leveling off of within session performance for procedural memory consolidation. *Cognitive Brain Research*, 24(2), 181–189.
- Heimlich, E. P., & Mark, A. J. (1990). *Paraverbal communication with children: Not through words alone*. NY; London: Plenum Press.
- Heine, S. J., Takata, T., & Lehman, D. R. (2000). Beyond self-presentation: Evidence for self-criticism among Japanese. *Personality and Social Psychology Bulletin*, 26, 71–78.
- Heirich, R. (2011). *Voice and the Alexander technique: Active explorations for speaking and singing*. Berkeley: Mornum Time Press.
- Henter, R., & Indreica, E., (2014) Reflective Journal Writing as a Metacognitive tool. *International Conference of Scientific Paper*.
- Herbert, J. (1952). *The interpreter's handbook: How to become a conference interpreter*. Geneva: Georg.
- Herder, J. G. (1986). Essay on the origin of language. In *On the origin of language: Two essays, Jean-Jacques Rousseau and Johann Gottfried Herder* (trans. J. H. Moran & A. Gode). Chicago, IL: University of Chicago Press.
- Heskett, J. L., Sasser, W. E., & Hart, C. W. L. (1995). *Service breakthroughs: Changing the rules of the game*. New York: Free Press.
- Heyne, F. (2013). *Understanding paraverbal communication*. Retrieved from [https://www.ipersonic.com/blog\\_files/Understanding-Paraverbal-Communication.html](https://www.ipersonic.com/blog_files/Understanding-Paraverbal-Communication.html)
- Hild, A. (2014). The role of self-regulatory processes in the development of interpreting expertise. *Translation and Interpreting Studies*, 9(1), 128–149.
- Hirvonen, P. (1973). *Measurement of achievement in English at the end of upper secondary school: 4. Speaking Test*. Turku: AFinLA Publications 8.



- Hlavac, J. (2022). Training Interpreters. In T. M. Derwing (Ed.), *The Routledge Handbook of Second Language Acquisition and Speaking* (1st ed.), (pp. 427–441). New York & London: Taylor and Francis Group.
- Hoban, J. D., Lawson, S. R., Mazmanian, P. E., Best, A. M., & Seibel, H. R. (2005). The self-directed learning readiness scale: A factor analysis study. *Medical Education*, 39(4), 370–379.
- Hockett, C. D. (1960). *The origin of speech*. San Francisco: W.H. Freeman.
- Hofmeyr, Z. (2016). Exploring some of Cicely Berry’s voice and language exercises for a directing concept of Athol Fugard’s Sorrows and rejoicings. *South African Theatre Journal*, 32(2), 123-141.
- Holub, E. (2010). Does intonation matter? The impact of monotony on listener comprehension. *Interpreters Newsletter*, 15, 117–126.
- Horváth, I. (2012). *Interpreter behaviour: A psychological approach*. Budapest: Hang Nyelviskola.
- Horváth, I. (2017). The speech behaviour of interpreters. *Across Languages and Cultures*, 18(2), 219–236.
- Houseman, B. (2015). *Finding your voice: A step-by-step guide for actors*. London: Nick Hern Books.
- Hsieh, H-F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, 15(9), 1277–1288.
- Hu, G. (2006). Adaptation in consecutive interpreting. *Perspectives: Studies in Thanatology*, 14(1), 3–21.
- Huang, F. S. (1992). *Learning motivation of adults*. Kaohsiung: Fu Wen.
- Huang, C., Lin, C., & Su, B. (2004). The effects of phonological awareness training on technological university students’ phonics and vocabulary knowledge. *Journal of National Taipei Teachers College*, 17, 59–90.
- Iglesias Fernández, E. (2006). La incidencia del parámetro `agradabilidad de la voz` y los estudios de calidad de la interpretación simultánea. En VarelaSalinas, M.J. (coord.). *La evaluación en los estudios de traducción e interpretación*. (pp.225-239). Bienza: Sevilla.
- Iglesias Fernández, E. (2007). La incidencia del parámetro “agradabilidad de la voz”. In A. Collados Aís, E. Pradas Macías, E. Stévaux & O. García Becerra (Eds.) *La evaluación de la*

- calidad en interpretación simultánea: Parámetros de incidencia*, (pp. 37-51). Granada: Comares.
- Iglesias Fernández, E. (2013). Unpacking delivery criteria in interpreting quality assessment. In D. Tsagari & R. Van Deemter (Eds). *Assessment issues in language translation and interpreting*, (pp. 51-66). Frankfurt: Peter Lang.
- Ilardo, J. A. (1981). *Speaking persuasively*. NY; London: Macmillan.
- International Association of Conference Interpreters (AIIC). (2004). *Practical guide*. AIIC. Retrieved from [https://aiic.org/document/547/AIICWebzine\\_Apr2004\\_2\\_Practical\\_guide\\_for\\_professional\\_conference\\_interpreters\\_EN.pdf](https://aiic.org/document/547/AIICWebzine_Apr2004_2_Practical_guide_for_professional_conference_interpreters_EN.pdf)
- International Organization for Standardization. (2014). *International Interpreting – Guidelines for community interpreting*.
- International Organization for Standardization. (2018). Draft International Standard ISO/DIS 18841 *Interpreting services – General requirements and recommendations*. Jain, S., Janssen, K., & DeCelle, S. (2004). *Alexander and Feldenkrais clinics*. Physical Medicine Rehabilitation Clinic North America. Retrieved from [http://www.synidetics.com/resources/CAIM/alexander-and-feldenkrais\\_PMR-clinics.pdf](http://www.synidetics.com/resources/CAIM/alexander-and-feldenkrais_PMR-clinics.pdf)
- Järvinen, K., Laukkanen, A., & Aaltonen, O. (2013). Speaking a foreign language and its effect on F0. *Logopedics Phoniatrics Vocology*, 38, 47–51.
- Järvinen, K., Laukkanen, A. M., & Geneid, A. (2017). Voice quality in native and foreign languages investigated by inverse filtering and perceptual analyses. *Journal of Voice*, 31(2), 261.e25–261.e31.
- Jeffries, K. J., Fritz, J. B., & Braun, A. R. (2003). Words in melody: An H215O PET study of brain activation during singing and speaking. *Neuro Report* 14(15), 749–754.
- John Henny Vocal Studio. (2021). *Finding singing resonance – Episode 215*. Retrieved from <https://johnhenny.com/simple-exercises-for-finding-singing-resonance/>
- Jones, R. (1998). *Conference interpreting explained*. Manchester, UK: St. Jerome Publishing.
- Joseph, A. S. (2013). *Vocal power: Harness your inner voice to conquer everyday communication challenges*. New York: McGraw-Hill Education.

- Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology*, 7(4), 396.
- Karaksha, A., Anoopkumar-Dukie, S., Grant, G., Davey, A. K., Nirthanan, S. N., Arora, D., ... & Haywood, A. (2013). Benefits of intensive mode teaching to improve student performance. *Proceedings of ICERI 2013 Conference*. Retrieved from [http://www98.griffith.edu.au/dspace/bitstream/handle/10072/59795/90003\\_1.Pdf](http://www98.griffith.edu.au/dspace/bitstream/handle/10072/59795/90003_1.Pdf)
- Kayi, H. (2012). *Teaching speaking: Activities to promote speaking in a second language*. Retrieved from <http://iteslj.org/Techniques/Kayi-TeachingSpeaking.html>
- Keating, P., & Kuo, P. (2012). Comparison of speaking fundamental frequency in English and Mandarin. *Journal of Acoustic Society of America*, 132, 1050–1060.
- Keller, E. (2004). The analysis of voice quality in speech processing. In *International School on Neural Networks, initiated by IIASS and EMFCSC* (pp. 54–73). Berlin; Heidelberg: Springer.
- Kellett Bidoli, C.J. (1995): Video-aided testing of student delivery and presentation in consecutive interpretation, in *The Interpreters' Newsletter*, 6, 43-66.
- Kelly, N. (2008). *Telephone interpreting: A comprehensive guide to the profession*. Bloomington: Trafford.
- Kelly, J. (2015). *What's in a voice? The debate over vocal fry and what it means for women*. Retrieved from <https://news.virginia.edu/content/whats-voice-debate-over-vocal-fry-and-what-it-means-women>
- Kemper, M. (2008). *Implicit and explicit instruction of spelling rules* (PhD dissertation). Radboud Universiteit Nijmegen.
- Kintsch, W. (1974). *The representation of meaning in memory*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kirchhoff, H. (1976/2002). SI: interdependence of variables in the interpreting process, interpreting models and interpreting strategies. In F. Pöchhacker & M. Shlesinger (Eds.), *The interpreting studies reader* (pp. 112-130). London/New York: Routledge.
- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698–2704.

- Kluwe, R. H. (1982). Cognitive knowledge and executive control: Metacognition. In D. R. Griffin (Ed.), *Animal mind – human mind* (pp. 201–224). NY: Springer.
- Kong, Y. (2009). A brief discussion on motivation and ways to motivate students in English language learning. *International Education Studies*, 2(2), 145–149.
- Kooijman, P. G., De Jong, F. I. C. R. S., Oudes, M. J., Huinck, W., Van Acht, H., & Graamans, K. (2005). Muscular tension and body posture in relation to voice handicap and voice quality in teachers with persistent voice complaints. *Folia Phoniatica et Logopaedica*, 57(3), 134–147.
- Kopczyński, A. (1980). *Conference interpreting: Some linguistic and communicative problems*. Poznań: Wydawnictwo Naukowe Uniwersytetu im. Adama Mickiewicza.
- Koponen, M., & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5–24). Ann Arbor: University of Michigan Press.
- Kreiman, J., & Sidtis, D. (2011). *Foundations of Voice Studies: An Interdisciplinary Approach to Voice Production and Perception*. Wiley-lackwell.
- Kucsera, J. V., & Zimmaro, D. M. (2010). Comparing the effectiveness of intensive and traditional courses. *College Teaching*, 58(2), 62–68.
- Kurz, I. (2001). Conference Interpreting: Quality in the Ears of the User. *Meta*, 46(2), 394–409.
- Kurz, I. (2008). The impact of non-native English on students' interpreting performance. In *Efforts and models in interpreting and translation research: A tribute to Daniel Gile* (pp. 179–192). Amsterdam & Philadelphia: John Benjamins.
- KV, A. (2008). *Lessac–Madsen resonant voice therapy: Clinician manual*. San. Diego: Plural Publishing, Inc.
- Ladefoged, P. (2001). *A course in phonetics*. San Diego, CA: Harcourt Brace.
- Lajoie, S. P., & Azevedo, R. (2006). Teaching and learning in technology-rich environments. In P. Alexander & P. Winne (Eds.), *Handbook of educational psychology*. Mahwah, NJ: Erlbaum.
- Lajoie, S. P. (2008). Metacognition, self-regulation, and self-regulated learning: A rose by any other name? *Educational Psychology Review*, 20(4), 469–475.
- Laver, J. (1991). *The gift of speech: Papers in the analysis of speech and voice*. Edinburgh: Edinburgh University Press.

- Le, E., Nguyen, V.N. and Ménard, N. (2009), *Interpreters' identities: an exploratory study of Vietnamese interpreters in Vietnam*, *Transcultural*, Vol. 1 No. 2, p. 93.
- Lec, A., & Khaleel, A. (2009). *The influence of speech delivery rates on simultaneous interpreters' performance*. Retrieved from <https://www.iasj.net/iasj/download/c449c7579249a113>
- Lee, B., & Sidtis, D. V. L. (2017). The bilingual voice: Vocal characteristics when speaking two languages across speech tasks. *Speech, Language and Hearing*, 20(3), 174–185.
- Lee, J., & Buzo, A. (2009). *Community language interpreting: A workbook*. NSW: The Federation Press.
- Lee, N., & Horsfall, B. (2010). Accelerated learning: A study of faculty and student experiences. *Innovative Higher Education*, 35(3), 191–202.
- Lee, Yun-Hyang. (2003). *Comparison of Error Frequency in Simultaneous Interpretation A to Band B to A (Korean-English)*. Unpublished DEA (pre-doctoral) thesis, University of Geneva, Ecole de Traduction et Interprétation.
- Legault, L., Green-Demers, I., & Pelletier, L. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of Educational Psychology*, 98(3), 567.
- Lepper, M. R., Corpus, J. H., & Iyengar, S. S. (2005). Intrinsic and extrinsic motivational orientations in the classroom: Age differences and academic correlates. *Journal of Educational Psychology*, 97(2), 184.
- Levett-Jones, T. L. (2005). Self-directed learning: implications and limitations for undergraduate nursing education. *Nurse Education Today*, 25(5), 363–368.
- Li, S.Z., & Jain, A. (Ed.). (2009). Fundamental Frequency, Pitch, F0. *Encyclopedia of Biometrics*. Springer, Boston.
- Li, T., & Lynch, R. (2016). The relationship between motivation for learning and academic achievement among basic and advanced level students studying Chinese as a foreign language in years 3 to 6 at Ascot International School in Bangkok, Thailand. *Scholar: Human Sciences*, 8(1), 1–1.
- Liao, M. Y. (2002). *The effects of gesture and movement training on the intonation and tone quality of children's choral singing* (doctoral dissertation). University of Sheffield.

- Liddicoat, A. (1993). *Bilingualism: An introduction*. NLIA Occasional Papers No. 2; Retrieved from <http://eric.ed.gov/?id=ED404847>
- Lim, H. O. (2008). The art of public speaking and the art of interpretation. *Forum: Revue Internationale D'interprétation et de Traduction/International Journal of Interpretation and Translation*, 6(2), 125–142.
- Linklater, K. (2006). *Freeing the natural voice*. NY: Applaud.
- Linklater, K. (2016). The art and craft of voice and speech training. *Journal of Interdisciplinary Voice Studies*, 1(11(1)), 57–70.
- Livingston, J. A. (2003). *Metacognition: An overview*. (ERIC Document Reproduction Service No. ED474273)
- Löfqvist, A., & Mandersson, B. (1987). Long-time average spectrum of speech and voice analysis. *Folia Phoniatica et Logopaedica*, 39(5), 221–229.
- Luis, J., & Morell, R. (2011). *Toward the development of a metacognitive intercultural communicative competence in the education of students of interpreting: General theoretical/pragmatic foundations*. *Translation and Interpreting: The International Journal of Translation and Interpreting Research*, 3(1), 106–118.
- Lumen Learning. (n.d.). *Boundless communications: Effective vocal delivery*. Retrieved from <https://courses.lumenlearning.com/boundless-communications/chapter/effective-vocal-delivery/>
- Lynch, M. (2021). *The difference between tone and intonation*. Retrieved from <https://www.theedadvocate.org/difference-between-tone-and-intonation/#:~:text=Tone%20refers%20to%20the%20emotion,one's%20voice%20as%20they%20speak.&text=Intonation%20is%20how%20you%20use,flow%20to%20convey%20a%20message>
- Mackay, I. (1985). *Introducing practical phonetics*. Taipei: Crane.
- Mackey, A., & Gass, S. M. (2016). *Second language research* (2nd ed.). Georgetown University and Michigan State University.
- Makarová, V. (1994). Whose line is it anyway? Or teaching improvisation in interpreting. In *Teaching translation and interpreting 2* (p. 207). Amsterdam & Philadelphia: John Benjamins.

- Malde, M., Allen, M., Zeller, K. A., Conable, B., & Nichols, T. R. (2020). *What every singer needs to know about the body* (Fourth edition.). San Diego, California: Plural Publishing, Inc.
- Malde, M., Allen, M. J., & Zeller, K.-A. (2013). *What every singer needs to know about the body*. USA: McNaughton & Gunn.
- Male, S. A., Baillie, C., Hancock, P., Leggoe, J., MacNish, C., & Crispin, S. (2016). *Intensive mode teaching good practice report*. Retrieved from <https://unistars.org/papers/STARS2017/03C.pdf>
- Marquezin, D. M. S. S., Viola, I., Ghirardi, A. C. de A. M., Madureira, S., & Ferreira, L. P. (2015). Executives' speech expressiveness: Analysis of perceptive and acoustic aspects of vocal dynamics. *CoDAS*, 27(2), 160–169.
- Marrone, S. (1993). Quality: A shared objective. *Interpreters' Newsletter*, 5, 35–41.
- Martin, C. L. (1988). Enhancing children's satisfaction and participation using a predictive regression model of bowling performance norms. *Physical Educator*, 45(4), 196.
- Matthei, E., & Roeper, T. (1985). *Understanding and producing speech*. NY: University Books.
- Maykut, P. S., & Morehouse, R. E. (1994). *Beginning qualitative research: A philosophic and practical guide*. London; Washington, DC: Falmer Press.
- McCombs, B. L. (2004). The learner-centered psychological principles: A framework for balancing academic achievement and social-emotional learning outcomes. *Building Academic Success on Social and Emotional Learning: What does the Research Say*, 1, 23–39.
- McCombs, B. L., & Whisler, J. S. (1997). *The learner-centered classroom and school: Strategies for increasing student motivation and achievement*. San Francisco: Jossey-Bass.
- McCrinkle, A. R., & Christensen, C. A. (1995). The impact of learning journals on metacognitive and cognitive processes and learning performance. *Learning and Instruction*, 5(2), 167–185.
- McDonough, D. (2012). Applying learner-centered principles and strategies: From face to face instruction to a hybrid course learning format. *Journal of Learning in Higher Education*, 8(2), 31–39.
- Meak, L. (1990) Interprétation simultanée et congrès médical: Attentes et commentaires, *The Interpreter's Newsletter*, 3, 8–13.
- Mehrabian, A. (1971). *Silent messages*. Belmont, California: Wadsworth Publishing Company.

- Melton, J. (1993). *Pilates training and the actor/singer*. <http://www.joanmelton.com/pilates-training-and-actorsinger>
- Melton, J. (2012). *One voice: Integrating singing and theatre voice techniques* (2nd ed.). Long Grove, IL: Waveland Press.
- Mikkelson, H. (2009). *Interpreting is interpreting – or is it?* Retrieved from <http://aiic.net/page/3356/interpreting-is-interpreting-or-is-it/lang/1>
- Mikkelson, H. (2003) Telephone interpreting: boon or bane? in L. Pérez González (ed), *Speaking in tongues: language across contexts and users*. València: Universitat de València, 251–69.
- Miller, J. (2017). Engaging your audience. In *Stand up, speak out: The practice and ethics of public speaking* (p. 1). Communication Faculty Books. [https://dc.uwm.edu/comm\\_facbooks/1](https://dc.uwm.edu/comm_facbooks/1)
- Mills, J. (2004). *The broadcast voice*. Amsterdam: Focal Press.
- Mithen, S., & Parsons, L. (2008). The brain as a cultural artefact. *Cambridge Archaeological Journal*, 18(03), 415.
- Molloy, K., Moore, D. R., Sohoglu, E., & Amitay, S. (2012). Less is more: Latent learning is maximized by shorter training sessions in auditory perceptual learning. *PloS One*, 7(5), e36929.
- Moser-Mercer, B. (2011). Remote interpreting. *Handbook of Translation Studies*, 2, 131–134.
- Mossop, B. (1987). Who is addressing us when we read a translation? *TextContText* 2,2,1-22
- Muller-Townsend, K. L. (2017). *Transfer of automatic skills: the role of automaticity in skill acquisition and transfer*. <https://ro.ecu.edu.au/theses/1954>
- NAATI. (2020). *NAATI – A connected community without language barriers*. Retrieved from <https://www.naati.com.au/become-certified/certification/certified-conference-interpreter/>
- Nafisi, J. (2013). Gesture and body-movement as teaching and learning tools in the classical voice lesson: A survey into current practice. *British Journal of Music Education*, 30(3), 347–367.
- Naimushin, B (2014). The Interpreter as Performer. Grozeva, Maria & Naimushin, Boris. (Eds.). Globalisierung — Interkulturelle Kommunikation — Sprache. Akten des 44. *LinguistischenKolloquiums*. Frankfurt: Peter Lang



- Namy, C. (1978). Reflections on the Training of Simultaneous Interpreters A metalinguistic approach. In: Gerver, D., Sinaiko, H.W. (Eds.). *Language Interpretation and Communication. NATO Conference Series, 6*. Springer, Boston, MA.
- Napier, J., Rohan, M., & Slatyer, H. (2005). Perceptions of bilingual competence and preferred language direction in Auslan/English interpreters. *Journal of Applied Linguistics, 2*(2), 185-218.
- National Institutes of Health. (2002). *What is the NIH?* Retrieved from [https://www.nidcd.nih.gov/health/voice/pages/whatis\\_vsl.aspx](https://www.nidcd.nih.gov/health/voice/pages/whatis_vsl.aspx)
- National Osteoporosis Foundation. (n.d.). *Proper body alignment*. Retrieved from <https://www.nof.org/patients/treatment/exercisesafe-movement/proper-body-alignment/#:~:text=Alignment%20refers%20to%20how%20the,the%20following%20positions%20or%20movements%3A&text=Twisting%20of%20the%20spine%20to%20a%20point%20of%20strain>
- Neiva, T. M. A., Gama, A. C. C., & Teixeira, L. C. (2016). Expressividade vocal e corporal para falar bem no telejornalismo: resultados de treinamento. *Revista CEFAC, 18*(2), 498–507.
- Nevo, L., Nevo, C., & Oliveira, G. (2015). A comparison of vocal parameters in adult bilingual Hebrew–English speakers. *Codas, 27*(5), 483–491.
- Ng, M. L., & Chen, Y. (2010). Comparison of vocal characteristics between Cantonese and English produced by Cantonese–English bilingual adult speakers. *Journal of the Acoustical Society of America, 128*, 2475.
- Nolan, J. (2008). *Interpretation: Techniques and exercises*. Clevedon: Multilingual Matters.
- Notas, N. (2012). *How to develop a confident attractive voice*. Retrieved from <https://www.nicknotas.com/blog/how-to-develop-a-confident-attractive-voice/>
- Novak, A., Dlouha, O., Capkova, B., & Vohradnik, M. (1991). Voice fatigue after theatre performance in actors. *Folia Phoniatica et Logopaedica, 43*, 74–78.
- Nutritious Movement. (n.d.). *Alignment is not posture*. Retrieved from <https://www.nutritiousmovement.com/alignment-is-not-posture/>
- Oates J. (2009). Auditory-perceptual evaluation of disordered voice quality: pros, cons and future directions. *Folia phoniatica et logopaedica: official organ of the International Association of Logopedics and Phoniatrics (IALP), 61*(1), 49–56.

- Ochs, E. (1979). Transcription as theory. In E. Ochs & B. B Schiefflin (Eds.), *Developmental pragmatics* (pp. 43–72). NY: Academic Press.
- Olson, J., & Johnson, C. S. (2012). Implementing journal writing in a grade 8 mathematics class. *Academic Research International*, 3(3), 191.
- Ota, M. (2016). Prosodic phenomena: Stress, tone and intonation. *Oxford handbook of developmental linguistics*. Retrieved from [https://www.researchgate.net/publication/304023857\\_Proasodic\\_phenomena\\_Stress\\_tone\\_and\\_intonation](https://www.researchgate.net/publication/304023857_Proasodic_phenomena_Stress_tone_and_intonation)
- Ozcelik, E., Cagiltay, N. E., & Ozcelik, N. S. (2013). The effect of uncertainty on learning in game-like environments. *Computers & Education*, 67, 12–20.
- Paananen-Porkka, M. M. (2007). *Speech rhythm in an interlanguage perspective: Finnish adolescents speaking English* (PhD dissertation). University of Helsinki.
- Paige, D. D., Rasinski, T., Magpuri-Lavell, T., & Smith, G. S. (2014). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, 46(2), 123–156.
- Pallant, J. (2016). *SPSS Survival Manual: A Step By Step Guide to Data Analysis Using SPSS Program* (6th ed.). London, UK: McGraw-Hill Education.
- Pammu, A. B. I. D. I. N., Amir, Z. A. I. N. I., & Maasum, T. (2011). *Metacognition in reading: Reviewing the literature*. Retrieved from [http://irep.iium.edu.my/27403/4/SOLLS\\_Proceeding\\_2011.pdf](http://irep.iium.edu.my/27403/4/SOLLS_Proceeding_2011.pdf)
- Papaleontiou Louca, E. (2008). *Metacognition and Theory of Mind* (pp. 12-16). London: British Library.
- Paris, S. G., & Paris, A. H. (2001). Classroom applications of research on self-regulated learning. *Educational Psychologist*, 36(2), 89–101.
- Patterson, D., & Ladd, D. R. (1999). Pitch range modelling: New dimensions of variation. *Proceedings of 13th International Congress of Phonetic Sciences*, San Francisco (pp. 1169–1172).
- Perez, I. A., Hartley, A., Mason, I., & Peng, G. (2003). *Peer- and self-assessment in conference interpreting training*. Retrieved from [www.lang.ltsn.ac.uk/prf.aspx#lang1](http://www.lang.ltsn.ac.uk/prf.aspx#lang1)

- Pérez-López, D., & Contero, M. (2013). Delivering educational multimedia contents through an augmented reality application: A case study on its impact on knowledge acquisition and retention. *Turkish Online Journal of Educational Technology (TOJET)*, 12(4), 19–28.
- Peterson, R. T. (2005). An examination of the relative effectiveness of training in nonverbal communication: Personal selling implications. *Journal of Marketing Education*, 27(2), 143–150.
- Physiopedia. (2021). *Posture*. Retrieved from <https://www.physiopedia.com/index.php?title=Posture&oldid=272857>
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385–407.
- Pöchhacker, F. (1994). Quality assurance in simultaneous interpreting. In C. Dollerup & A. Lindegaard (Eds.), *Teaching translation and interpreting 2: Insights, aims and visions* (pp. 233–242). Amsterdam; Philadelphia: John Benjamins.
- Pöchhacker, F. (2001). Quality assessment in conference and community interpreting. *Meta*, 46(2), 410.
- Pöchhacker, F., & Shlesinger, M. (2002). *The interpreting studies reader*. Routledge.
- Pöchhacker, F. (2011). Conference interpreting. *Identity and Status in the Translational Professions*, 32, 49.
- Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona 1. *Journal of Sociolinguistics*, 11(4), 478–504.
- Pressley, M., Borkowski, J. G., & Schneider, W. (1987). *Cognitive strategies: Good strategy users coordinate metacognition and knowledge*. *Annals of Child Development*, 4, 89-129.
- Pressley, M., and McCormick, C. B. (1995). *Advanced Educational Psychology for Educators, Researchers, and Policymakers*, HarperCollins, New York.
- Putnam, A. R. (2001, December). *Problem-based teaching and learning in technology education*. Opinion Paper, presented at the Annual Conference of the Association for Career and Technical Education, New Orleans, LA.
- Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. NY: Oxford University Press.

- Reiman, T. (2019). *The human voice – Pitch*. Retrieved from <http://www.bodylanguageuniversity.com/public/203.cfm>
- Reina, T. (2018). *Types of vocal timbre*. Retrieved from <https://www.musictoyourhome.com/blog/types-of-vocal-timbre/#:~:text=Vocal%20Timbre%2C%20or%20as%20it,and%20how%20it%20is%20recognized>
- Roach, P. (2001). *Phonetics*. Oxford University Press.
- Roberts, R. (1994). Community interpreting today and tomorrow. In P. Krawutschke (Ed.), *35th Annual Conf. of the American Translators Association* (pp. 127–138). Medford, NJ: Learned Information.
- Rodenburg, P. (2015). *The right to speak: Working with the voice* (2nd ed.). London: Bloomsbury.
- Rodgers, J. (2002). *The complete voice and speech workout*. Montclair, NJ: Applause.
- Roelle, J., Berthold, K., & Fries, S. (2011). Effects of feedback on learning strategies in learning journals: Learner expertise matters. *International Journal of Cyber Behavior, Psychology and Learning, 1* (2), 16-30.
- Roloff, L. (1973). The roles of the interpreter and the actor. *The Speech Teacher, 22*(2), 144–147.
- Rosenthal, B. (2008). *Voice projection: Where presentation confidence starts*. Communispond. <https://communispond.com/insights/blog/742/voice-projection-where-presentation-confidence-starts/#:~:text=Although%20it%20may%20not%20seem,powerful%20presentation%20skills%20to%20learn.&text=Projecting%20your%20voice%20correctly%20can,your%20confidence%20as%20a%20speaker>
- Roy, C. B. (2000). *Interpreting as a discourse process*. Oxford: Oxford University Press.
- Russo, M. (1995). Self-evaluation: The awareness of one's own difficulties as a training tool for simultaneous interpretation. *Interpreters' Newsletter, 6*, 75–86.
- Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage Publications Ltd.
- Salomaa, E. (2019). *The development of speech rhythm and fluency of advanced English learners: A mixed methods study of the correlation between native speaker evaluations and acoustic*

- measures*. Retrieved from [https://www.utupub.fi/bitstream/handle/10024/147427/Salomaa\\_Emma\\_Thesis.pdf?isAllowed=y&sequence=1](https://www.utupub.fi/bitstream/handle/10024/147427/Salomaa_Emma_Thesis.pdf?isAllowed=y&sequence=1)
- Sanford, D. H. (1988). Begging the question as involving actual belief and inconceivable without it. *Metaphilosophy*, 19(1), 32–37.
- Sapir, E. (1927). Speech as a personality trait. *American Journal of Sociology*, 32(6), 892–905.
- Saxena, A. S. (2020). Impact of metacognitive strategies on self-regulated learning and intrinsic motivation. *Journal of Psychosocial Research*, 15(1), 35–46.
- Schack, K. (2000). *Comparison of intonation patterns in Mandarin and English for a particular speaker*. University of Rochester Working Papers in the Language Sciences (spring/summer).
- Schmeichel, B. J. (2007). Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *J Exp Psychol Gen*, 136, 241–255.
- Schmeichel, B. J., & Baumeister, R. F. (2004). Self-regulatory strength. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 84–98). NY: The Guilford Press.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158.
- Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review*, 11, 1126.
- Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human factors*, 27(3), 285-300.
- Schneider, C. M., Dennehy, C. A., & Saxon, K. G. (1997). Exercise physiology principles applied to vocal performance: The improvement of postural alignment. *Journal of Voice*, 11(3), 332–337.
- Schoenfeld, A. H. (1987). What’s All the Fuss about Metacognition? In A. H. Schoenfeld (Ed.), *Cognitive Science and Mathematics Education* (pp. 189-215). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Schraw, G. (1994). The effect of metacognitive knowledge on local and global monitoring. *Contemp Educ Psychol*, 19, 143–154.

- Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science*, 26, 113–125.
- Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in Science Education*, 36(1), 111–139.
- Schreier, M. (2012). *Qualitative content analysis in practice*. Thousand Oaks, CA: Sage.
- Schunk, D. H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology*, 7(2), 112–137.
- Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 23(1), 7–25.
- Segalowitz, N. (2010). *The cognitive bases of second language fluency*. NY: Routledge.
- Seleskovitch, D. (1968). *L'Interprète dans les conférences internationales. Problèmes de langage et de communication*. Paris: Minard. Trans. S. Dailey & N. McMillan (1978), *Interpreting for international conferences*. Arlington: Pen & Booth.
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4–14.
- Shewell, C. (2009). *Voice work – Art and science in changing voices*. UK: John Wiley & Sons
- Shewell, C. (2009). Peer Reviewed Article The Voice Skills Perceptual Profile: A Practical Tool for Voice Teachers. *Voice & Speech Review*, 6(1), 171–184.
- Shipley, K., & McAfee, J. (2016). *Assessment in speech-language pathology: A resource manual* (5th ed.). Clifton Park, NY: Delmar Cengage Learning.
- Shlesinger, M. (1994). Intonation in the production and perception of simultaneous interpretation. In S. Lambert & B. Moser-Mercer (Eds.), *Bridging the gap: Empirical research in simultaneous interpretation* (pp. 225–236). Amsterdam; Philadelphia: John Benjamins.
- Shlesinger, M. (1997). Quality in simultaneous interpreting. In Y. Gambier, D. Gile & C. Taylor (Eds.), *Conference interpreting: Current trends in research* (pp. 123–131). Amsterdam; Philadelphia: John Benjamins.
- Simmonds, A. J., Wise, R. J. S., & Leech, R. (2011). Two tongues, one brain: Imaging bilingual speech production. *Frontiers in Psychology*, 2, 166.

- Sipley, K. L. (1997). *Training the speaking voice through singing*. National Communication Association. Retrieved from [http://www.fva.net/wp-content/uploads/2011/02/ctr\\_Training-the-speaking-voice-through-singing.htm](http://www.fva.net/wp-content/uploads/2011/02/ctr_Training-the-speaking-voice-through-singing.htm)
- Skehan, P. (2014). Limited attentional capacity, second language performance, and taskbased pedagogy. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 211–260). Amsterdam: John Benjamins.
- Smith, B. L., Brown, B. L., Strong, W. J., & Rencher, A. C. (1975). Effects of speech rate on personality perception. *Language and Speech, 18*(2), 145–152.
- Smith, H. (2017). *Implicit vs. explicit instruction: Which is better for word learning?* Fast Forward. Retrieved from <https://www.scilearn.com/implicit-vs-explicit-instruction-which-better-word-learning/#:~:text=By%20implicit%20instruction%2C%20we%20refer,conceptual%20structures%20and%20assimilate%20the>
- Soar, T. (1999). *Define the Alexander Technique*. Retrieved from <https://www.alexander-technique.london/defining-the-alexander-technique/the-lying-down-procedure/>
- Stanton, J. D., Neider, X. N., Gallegos, I. J., & Clark, N. C. (2015). Differences in metacognitive regulation in introductory biology students: When prompts are not enough. *CBE – Life Sciences Education, 14*(2), ar15.
- Strong, M. & Rudser, S. F. (1992). The subjective assessment of sign language interpreters. In D. Cokely (Ed.), *Sign language interpreters and interpreting* (pp. 1–14). Burtonsville, MD: Linstok Press.
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education, 30*(4), 325–341.
- Stevens, K. N. (2000). *Acoustic phonetics*. Cambridge, MA: MIT Press.
- Stevenson, A., & Lindberg, C. (2010). *New Oxford American dictionary*. Oxford ; New York : Oxford University Press.
- Summer, L. (2009). Language acquisition for the bilingual child: A perspective on raising bilingual children in the US. In P. Fletcher & B. MacWhinney (Eds.), *The NCHAM ebook*. [www.infantheating.org/.../17-Chapter17LanguageAcquisition2014.pdf](http://www.infantheating.org/.../17-Chapter17LanguageAcquisition2014.pdf)

- Tavakoli, P., Campbell, C., & McCormack, J. (2016). Development of speech fluency over a short period of time: Effects of pedagogic intervention. *TESOL Quarterly*, *50*(2), 447–471.
- Thurman, L., & Welch, G. (2000). *Body, mind and voice: Foundations of voice education* (rev. ed., Vols. 1–3). Collegeville, MN: Voice Care Network, National Center for Voice and Speech, Fairview Voice Center.
- Thomas, D. R. (2006). A General Inductive Approach for Analyzing Qualitative Evaluation Data. *American Journal of Evaluation*, *27*(2), 237–246.
- Tissi, B. (2000). Silent pauses and disfluencies in simultaneous interpretation: A descriptive analysis. *The Interpreters' Newsletter*, *10*, 103-127.
- Tohidi, H., & Jabbari, M. M. (2012). The effects of motivation in education. *Procedia – Social and Behavioral Sciences*, *31*, 820–824.
- Tommola, J., and Helevä, M. (1998). “Language direction and source text complexity: effects on trainee performance in simultaneous interpreting,” in *Unity in Diversity?: Current Trends in Translation Studies*, eds L. C. Bowker, D. M. Kenny, and J. Pearson (Manchester: St Jerome), 177–186.
- Traub, J. (1997). Drive-thru U: Higher education for people who mean business (The next university). *New Yorker*, *73*, 114–122.
- Tsurutani, C. (2012). Evaluation of speakers with foreign-accented speech in Japan: The effect of accent produced by English native speakers. *Journal of Multilingual and Multicultural Development*, *33*(6), 589–603.
- Turan, S., Demirel, Ö., & Sayek, İ. (2009). Metacognitive awareness and self-regulated learning skills of medical students in different medical curricula. *Medical Teacher*, *31*(10), e477–e483.
- Turner, J. C. (2003). *Voice and speech in the theatre* (4th ed.). London: A & C Black.
- Tzelgov, J. (1999). Automaticity and processing without awareness. *Psyche*, *5*(3), 18–23.
- UDL. (2018). *UDL guidelines: Self-regulation*. Retrieved from <https://udlguidelines.cast.org/engagement/self-regulation>
- U. S. Department of Education, Office of Vocational and Adult Education. (2011). *Just write! guide*. Washington, DC: Author. Retrieved from <https://lincs.ed.gov/state-resources/federal-initiatives/teal/guide/metacognitive>



- Van Dam, I. M. (1989). Strategies of simultaneous interpretation. In *The theoretical and practical aspects of teaching conference interpretation* (pp. 167–176). Udine: Campanotto.
- van Egdom, G. W., Cadwell, P., Kockaert, H., & Segers, W. (2020). A turn to ergonomics in translator and interpreter training. *The Interpreter and Translator Trainer*, 14:4, 363-368,
- van Heuven, V. J. J. P. (1994). Introducing prosodic phonetics. In C. Odé & V.J. van Heuven (eds), *Experimental studies of Indonesian prosody*, 1-26. Leiden: Department of Languages and Cultures of South-East Asia and Oceania, Leiden University. (Semaian 9)
- Vandergrift, L & Goh, C. (2006). The metacognitive awareness listening questionnaire: Development and validation. *Language Learning*, 56(3), 431–462.
- Venditti, J. J. (2006). *21 tone and intonation*. Retrieved from [https://www.asc.ohio-state.edu/beckman.2/publications/BeckmanVenditti\\_HPhSci\\_finalVersion.pdf](https://www.asc.ohio-state.edu/beckman.2/publications/BeckmanVenditti_HPhSci_finalVersion.pdf)
- Viezzi M. (1996) *Aspetti della qualità in interpretazione (in English, “Aspects of Quality in Interpreting”)*, Università degli studi di Trieste, Scuola superior di lingue moderne per interpreti e traduttori.
- Virgil, A.A. (1967). *Training the Speaking Voice*. London: Oxford University.
- Voss, J. F. (1988). Problem solving and reasoning in ill-structured domains. In C. Antaki (Ed.), *Analysing everyday explanation: A casebook of methods* (pp. 74–93). Sage Publications, Inc.
- Vox Daily. (2009). *The dangers of whispering for your voice*. Retrieved from [http://blogs.voices.com/voxdaily/2009/07/the\\_dangers\\_of\\_whispering\\_for\\_your\\_voice.html](http://blogs.voices.com/voxdaily/2009/07/the_dangers_of_whispering_for_your_voice.html)
- Wagner, A., & Braun, A. (2003). Is voice quality language-dependent? Acoustic analyses based on speakers of three different languages. *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 651–654). Retrieved from [https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15\\_0651.pdf](https://www.internationalphoneticassociation.org/icphs-proceedings/ICPhS2003/papers/p15_0651.pdf)
- Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Newbury Park, CA: Sage.
- Weber, W. K. (1989). Improved ways of teaching consecutive interpretation. In L. Gran & J. Dodds (Eds.), *The theoretical and practical aspects of teaching conference interpretation* (pp. 161–166). Udine: Campanotto.
- Weiler, N. (2018). *Study reveals brain activity patterns underlying fluent speech*. Retrieved from <https://www.ucsf.edu/news/2018/05/410606/study-reveals-brain-activity-patterns-underlying-fluent-speech>

- Wenden, A. (1999). An introduction to metacognitive knowledge and beliefs in language learning: Beyond the basics [Special Issue]. *System*, 27, 435-441.
- Wessendarp, E. (2012). *Effects of the Linklater voice training technique on the voices of student actors* (doctoral dissertation). University of Toledo.
- White, B. Y. (1993). ThinkerTools: Causal models, conceptual change, and science education. *Cognition and Instruction*, 10(1), 1–100.
- William, L. B., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review*, 6(4), 345–375.
- Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *The Journal of the Acoustical Society of America*, 52(4B), 1238–1250.
- Williams, S. (1995). Observations on anomalous stress in interpreting. *The Translator*, 1(1), 47–64.
- Wilson, P. H. (2011). Act, Sing, Speak: Voice in the world of theatre, *Voice & Speech Review*, 7:1,298-304,
- Wisdom, J., & Creswell, J. W. (2013). Mixed Methods: Integrating Quantitative and Qualitative Data Collection and Analysis While Studying Patient-Centered Medical Home Models (pp. 1-5). *PCMH Research Methods Series 13*.
- Wolf K., Müller N.G. (2012) Attention and Implicit Learning. In: Seel N.M. (eds) *Encyclopedia of the Sciences of Learning*. Springer, Boston, MA.
- Wolfe, A. (1998). How a for-profit university can be invaluable to the traditional liberal arts. *Chronicles of Higher Education*, 44, B4–B5.
- Wolk, L., Abdelli-Beruh, N. B., & Slavin, D. (2012). Habitual use of vocal fry in young adult female speakers. *Journal of Voice*, 26(3), e111–e116.
- Wood, D. F. (2003). Problem-based learning. *BMJ*, 326(7384), 328–330.
- Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *Academy of Management Review*, 14(3), 361–384.
- Wright, B. A., & Sabin, A. T. (2007). Perceptual learning: How much daily training is enough? *Experimental Brain Research*, 180(4), 727–736.
- Yenkimaleki, M. (2018). Implicit vs. explicit prosody teaching in developing listening comprehension skills by interpreter trainees: An experimental study. *International Journal of English Language and Linguistics Research*, 6(1), 11–21.

- Yenkimaleki, M., & van Heuven, V. J. (2018). The effect of teaching prosody awareness on interpreting performance: an experimental study of consecutive interpreting from English into Farsi. *Perspectives: Studies in Translatology*, 26(1), 84–99.
- Yin, K. (2011). Disfluencies in consecutive interpreting among undergraduates in the language lab environment. *Proceedings of the 25th Pacific Asia Conference on Language, Information and Computation*, Singapore: Nanyang Technological University, 459–466.
- Yip, M. (2002). *Tone*. Cambridge: Cambridge University Press.
- Young, A., & Fry, J. D. (2008). Metacognitive awareness and academic achievement in college students. *Journal of the Scholarship of Teaching and Learning*, 8(2), 1–10.
- Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, 11(4), 307–313.
- Zimmerman, B. J. (1989). A Social Cognitive View of Self-Regulated Academic Learning. *Journal of Educational Psychology*, 81, 329-339.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25(1), 3–17.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70.
- Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 299–315). Routledge/Taylor & Francis Group.
- Zohar, A., & Barzilai, S. (2013). A review of research on metacognition in science education: Current and future directions. *Studies in Science Education*, 49(2), 121–169.



## Appendix A-1 Explanatory Statement and Consent Form

### Investigation into voice and speech training for interpreting students

This information sheet is for you to keep.

My name is Vera Gu. I am a PhD candidate in the Translation and Interpreting Studies Program, Faculty of Arts, Monash University. Dr Jim Hlavac is my supervisor. He is a lecturer in the Translation and Interpreting Studies Program, Faculty of Arts, Monash University.

As an accredited and experienced interpreter, I would like to invite you to participate in an interview which seeks to elicit questions about the place of voice and speech training in the training and testing of interpreters and how you as an interpreter take care of your voice continue to develop your spoken and rhetorical skills.

#### The aim of the research

The aim of this study is to examine the place and importance of voice and speech training for trainee interpreters. From the data collected, we hope to be able to inform our teaching of interpreting and training of interpreters.

#### Possible benefits

Possible eventual benefits from this research are the inclusion of specific modules on voice and speech training for trainee interpreters. This will advance not only future practitioners' interpreting performance but also contribute to our improvement in the pedagogy of interpreting in general.

#### What does the research involve?

The research involves the conducting of an interview with you. This interview will be semi-structured and it will contain questions which relate to your training or pathway to becoming an interpreter and questions related to your continued care of your voice and development of spoken skills. The interview will be audio-recorded by the student researcher, Ms Vera Gu.

#### How much time will the research take?

Participants will be asked to take part in a semi-structured interview that will take approximately 20–30 minutes.

Participation is voluntary. Participants may withdraw at any stage of the interview.

#### Confidentiality

No names or specific details will be collected from participants. However, should any information be collected that can identify a participant, this information will be changed and de-identified to ensure participants' anonymity.

#### Storage of data

Storage of the data collected will adhere to the University regulations and kept on University premises in a locked cupboard/filing cabinet for 5 years. An article of the study may be submitted for publication, but individual participants will not be identifiable in such a report.

#### Results

If you would like to be informed of the aggregate research finding, please contact Vera Gu at 0432802602 or by email or Jim Hlavac at (03) 9903 4575.

If you would like to contact the researchers about any aspect of this study, please contact Vera Gu.	If you have a complaint concerning the manner in which this research <insert your project number here> is being conducted, please contact:
Phone: 0432802602  Email: vera.gu@monash.edu	Executive Officer Monash University Human Research Ethics Committee (MUHREC) Building 3e Room 111 Research Office Monash University VIC 3800  Tel: +61 3 9905 2052    Fax: +61 3 9905 3831 Email: <a href="mailto:muhrec@monash.edu">muhrec@monash.edu</a>

Thank you.



## Explanatory Statement

This information sheet is for you to keep.

My name is Dr Jim Hlavac and I am a lecturer in the Translation and Interpreting Studies Program, Faculty of Arts, Monash University. I am conducting a research project with Ms Vera Yingzhi Gu also in the Translation and Interpreting Studies Program, Faculty of Arts, Monash University. The research is part of Ms Vera Yingzhi Gu's PhD research, which she is undertaking under my supervision.

As you are a student of the Interpreting Stream in the Masters of Translation and Interpreting, I would like to invite you to participate in an interview which seeks to elicit questions about the place of voice and speech training in interpreter training.

### **The aim of the research**

The aim of this study is to examine the place and importance of voice and speech training for interpreters. Information from a variety of informants – practising and accredited interpreters, trainee interpreters and voice/trainers and speech pathologists – will be the basis of this research project. From the data collected, we hope to be able to inform our teaching of interpreting and training of interpreters.

### **Possible benefits**

Possible eventual benefits from this research are the inclusion of specific modules on voice and speech training for trainee interpreters. This will advance not only future practitioners' interpreting performance but also contribute to our improvement in the pedagogy of interpreting in general.

### **What does the research involve?**

The research involves the conducting of an interview with you. This interview will be semi-structured and it will contain questions which relate to your training or pathway to becoming an interpreter and questions related to your continued care of your voice and development of spoken skills. The interview will be audio-recorded by the student researcher, Ms Vera Yingzhi Gu.

The research also includes gaining access to video-taped recordings of you from the language-specific workshops. Vera Gu will contact you about the arrangements to gain these recordings with your permission.

### **How much time will the research take?**

Participants will be asked to take part in a semi-structured interview that will take approximately 10–20 minutes.

Participation is voluntary. Participants may withdraw at any stage of the interview.

### **Confidentiality**

No names or specific details will be collected from participants. However, should any information be collected that can identify a participant, this information will be changed and de-identified to ensure participants' anonymity.

### **Storage of data**

Storage of the data collected will adhere to the University regulations and kept on University premises in a locked cupboard/filing cabinet for 5 years. An article of the study may be submitted for publication, but individual participants will not be identifiable in such a report.

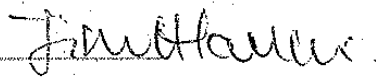
## Results

If you would like to be informed of the aggregate research finding, please contact Jim Hlavac on 9903 4575 or fax 9905 5437.

If you would like to contact the researchers about any aspect of this study, please contact the Chief Investigator: Dr Jim Hlavac	If you have a complaint concerning the manner in which this research <insert your project number here> is being conducted, please contact:
Phone: (03) 9903 4575 Email: Jim.Hlavac@monash.edu Fax: (03) 9905 5437	Executive Officer Monash University Human Research Ethics Committee (MUHREC) Building 3e Room 111 Research Office Monash University VIC 3800  Tel: +61 3 9905 2052 Fax: +61 3 9905 3831 Email: <a href="mailto:muhrec@monash.edu">muhrec@monash.edu</a>

Thank you.

Signature:



Jim Hlavac



Consent Form for Participants

**Investigation into voice and speech training for interpreters**

**NOTE:** This consent form will remain with the Monash University researcher for their records

---

**I agree to take part in the Monash University research project specified above. I have had the project explained to me, and I have read the Explanatory Statement, which I keep for my records. I understand that agreeing to take part means that:**

**I agree to be interviewed by the researcher**  Yes  No

**I agree to allow the interview to be audio-taped**  Yes  No

**I agree to complete a questionnaire**  Yes  No

**I agree to allow access to video recordings of myself when interpreting in the classroom, to be negotiated with Vera Gu.**  Yes  No

and

**I understand that my participation is voluntary, that I can choose not to participate in part or all of the project, and that I can withdraw at any stage of the project without being penalised or disadvantaged in any way.**

and

**I understand that any data that the researcher extracts from the interview in reports or published findings will not, under any circumstances, contain names or identifying characteristics.**

and

**I understand that any information I provide is confidential**

and

**I understand that data from the interview and audio-tape will be kept in a secure storage and accessible to the researcher. I also understand that the data will be destroyed after a 5 year period unless I consent to it being used in future research.**

Participant's name: \_\_\_\_\_

Signature: \_\_\_\_\_

Date: \_\_\_\_\_



## **Appendix A-2 Evaluation Instruction**

1. Please read the evaluation sheet carefully and clarify with the researcher if you have any questions and concerns regarding any questions and terminologies in evaluation sheet that are not clear to you.
2. You will be asked to evaluate four groups of students who are interpreting students from the Translation and Interpreting program at Monash University. The researcher has already collected two video recordings of each student in consecutive interpreting, one was when they were beginner interpreting students, and another was taken before they graduated. The scores of students' initial and final video recordings will be added up by the researcher and compared to evaluate the effectiveness of voice training and compare pre- and post-training results objectively. You were asked to evaluate their performance by viewing the video recordings. Please be aware that the purpose of the evaluation is to get your first impression of their performance rather than play the video back and forth to linger on the details.
3. The purpose of conducting the evaluation is not to check participants' interpreting techniques and linguistic transferring skills, nor is the acoustic voice analysis by voice and speech pathologists.
4. Please be as candid and neutral while assessing the participants' performance. Your evaluation will not affect participants' academic results. The participants would not be informed of your evaluation and assessment.

## Appendix A-3 Glossary for Interpreter Voice Skills Evaluation Sheet<sup>3</sup>

<b>Spoken voice categories /attributes</b>	<b>Explanation</b>
<b>Accent</b>	An accent is a manner of pronunciation peculiar to a particular individual, location, or nation.
<b>Alignment</b>	Alignment refers to how the head, shoulders, spine, hips, knees and ankles relate and line up with each other. Proper alignment of the body puts less stress on the spine and helps interpreters maintain a good posture.
<b>Articulation</b>	Articulation is an action of producing a sound or word clearly in speech.
<b>Audience connection/engagement</b>	To performing behaviours that increase feelings of liking, pleasure, and closeness in the minds of audience members
<b>Body awareness</b>	Awareness of the production of voiced sound by selecting appropriate use of body parts on stage, choosing to use the whole body, or isolated parts of the body.
<b>Breathing</b>	To allow a controlled use of voice, to convey a feeling e.g. a sigh or gasp.
<b>Breathing support and management</b>	Breathing support refers to how one stabilizes his/her body for air flow. Breathing management is how one regulate and coordinate airflow for speaking activity.
<b>Emphasis</b>	Choosing to stress particular words to convey meaning.
<b>Expressiveness</b>	Speech expressiveness is built from interactions between segmental (vowels and consonants) and prosodic (rhythm, intonation, voice quality, elocution rate, pauses, and stress) elements and between sound and meaning.
<b>Eye contact</b>	Establishing eye contact with speaker or the audience.
<b>Facial expression</b>	Using the face to show mood, emotion, feeling and Responses.
<b>Gesture</b>	Gesture means movement of any part of the body to express idea, feeling or mood.
<b>Inflection/Intonation</b>	The rise and fall of a voice.
<b>Pace and pause</b>	Pace refers to the rate at which words are spoken; Pause means where sound stops, how often and for how long (ability to adjust the speed and pausing according to the content of the speech).
<b>Pitch</b>	The highness or lowness of a sound.
<b>Pitch range</b>	An awareness of the optimum range of voice between the lowest and the highest and fundamental frequency.
<b>Physical tension balance</b>	Physical tension balance is awareness of the tension and stress in the body and ability to balance this tension with emotion.

<sup>3</sup> The glossary was compiled using the definitions in Section 2.2.5 as a guide.

<b>Posture</b>	Posture refers to stance, way of standing.
<b>Release of speech muscles</b>	Release of speech muscles refers to release the tight muscles around the larynx (voice box including organs of articulation such as the jaw, mouths, lips, tongue, soft palate.) during speaking that the voice box does not work efficiently.
<b>Resonance and timbre</b>	Resonance refers to the quality in a sound of being deep, full, and reverberating. Timbre describes the tone or unique quality of a sound.
<b>Spoken voice production</b>	Voice production corresponds to the physiological and physical processes by which vibration of the vocal fold is transformed into speech. Speech production is the process of uttering articulated sounds or words, i.e., how humans generate meaningful speech.
<b>Spoken voice quality</b>	Voice quality has been defined as the characteristic auditory colouring of an individual's voice, derived from a variety of laryngeal and supralaryngeal features and running continuously through the individual's speech. The distinctive tone of speech sounds produced by a particular person yields a particular voice.
<b>Spoken voice dynamism</b>	Reflect skills of using vocal variety to connect with the audience and increase credibility
<b>Speaking clarity (English)</b>	Clear articulation of English vowels and consonants; Pronunciation and accent that is unobstructive to the communication.
<b>Speech rhythm</b>	The sense of movement in speech, marked by the stress, timing, and quantity of syllables.
<b>Speaking confidence</b>	Speaking in a calm and deliberate manner and gives the audience impression that he or she is in control.
<b>Tone</b>	The feeling/emotion in the voice.
<b>Vocal modulation appropriate to the speech that is interpreting</b>	Voice convey or reveal of emotions of an intention of the original speaker, such as telling jokes or talking about serious topics.
<b>Volume/vocal projection/loudness</b>	Voice projection or loudness of spoken voice that is appropriate to the interpreting setting.

## Appendix B-1 Pre-Intervention Interpreting Speech

### 百尺竿头，还需一步 — 驻澳大利亚大使陈育明在澳中商会珀斯教育论坛上的演讲

女士们、先生们：

很高兴应邀出席澳中商会珀斯教育论坛。去年4月，我曾陪同全国政协主席贾庆林先生访问西澳大学。澳大利亚第一所孔子学院2015年就诞生在西澳大学。

40年来，中澳教育合作从无到有，从少到多，取得了令世人瞩目的成绩。1972年建交时，中国在澳留学生是零记载。1996年，澳大利亚只有不到5000名中国留学生。今天，‘留澳热’在中国成为时尚，有超过16万名中国学子在澳学习，澳每3个外国留学生中就有一个是中国面孔。中国连续多年是澳第一大留学生来源国。据澳政府统计，从2001年到2011年，总共有近150万中国学生来澳留学，占同期来澳留学生总数一半以上。他们不仅是澳教育产业的‘银行’，每年给澳带来近40亿澳元的可观收入，更是中澳人文交流的使者，为增进两国人民的了解和友谊发挥了重要桥梁作用。

中澳教育交流的内涵和外延还要宽广得多。两国政府签署了相互承认高等教育学历和学位的协议，教育部门建立了定期磋商机制。中国汉办与澳院校联合建立了12所孔子学院，在600多所中小学开设了汉语课程。普通话成为仅次于英语的澳第二大语言。中国九校与澳八校联盟建立了联席会议机制。中国国家留学基金委与澳国立大学、悉尼大学等高校开展联合培养博士生项目。莫纳什大学与中国东南大学在苏州建立研究生院，成为第一所在华设立校区的澳高等学府。

我最近思考的一个问题，是中澳教育合作能否在现有水平上实现可持续发展。相信这也是在座朋友关心的话题。在当前全球化纵深发展和‘亚洲世纪’敲门的背景下，这个问题显得紧迫。

我想带给大家两个消息。一是随着中国经济快速增长和民生水平不断提高，中国出国留学人数已占全球七分之一，居世界第一。澳有望从中国教育开放中分得更大红利。

二是美、欧、加、日等发达国家日益意识到中国留学生群体的经济价值和人文效应，纷纷出台优惠政策和便利措施。加之澳元高企等因素，澳在招揽中国留学生生源上面临空前激烈的国际竞争。

作为中国驻澳大使，我想为今后的中澳教育合作提出几条实在建议：

第一，希望澳政府和高校在吸引中国留学生生源方面加大投入。澳在入境、学费、生活、实习等方面可为中国留学生创造更好的条件及可靠的安全保障，使‘留澳热’保持温度。

第二，希望澳从政策层面支持双向留学交流。中澳互有留学生数量超过 50 比 1。美国政府最近推出‘十万强计划’，未来 4 年将向中国派遣 10 万留学生。澳能不能派一万学生呢？希望各界朋友提供支持。

第三，希望澳更多拥抱中文和中国文化，鼓励更多院校建立孔子学院、孔子课堂和中文课程，培养更多讲汉语、真正懂中国的友好传人。中方愿积极协助。

第四，探索中澳教育合作的新路径，比如推动中澳更多有条件的高等学府开展合作办学，交流互鉴，优势互补。谢谢大家！

**A long way to go further - Ambassador to Australia Chen Yuming's speech at the Australia  
China Chamber of Commerce Perth Education Forum**

[Reference Translation]

Ladies and Gentlemen,

I am delighted to be invited to attend the Australia–China Chamber of Commerce Perth Education Forum. In April last year, I accompanied Mr Jia Qinglin, Chairman of the National Committee of the Chinese People's Political Consultative Conference, to visit the University of Western Australia. The first Confucius Institute in Australia was established at the University of Western Australia in 2015.

In the past 40 years, China-Australia education cooperation has gone from strength to strength and from few to many and has achieved remarkable results. Today, the “study in Australia fever” has become fashionable in China, with more than 160,000 Chinese students studying in Australia and one in three foreign students in Australia having a Chinese face. China has been the number one source of international students in Australia for many years. According to the Australian government, from 2001 to 2011, nearly 1.5 million Chinese students came to study in Australia, accounting for more than half of all international students coming to Australia in the same period. They are not only the ‘bank’ of Australia's education industry, bringing nearly A\$4 billion to Australia every year, but also the messenger of China–Australia humanistic exchanges, playing an important role as a bridge to enhance the understanding and friendship between the two peoples.

The connotation and outreach of China–Australia education exchange is much broader. The two governments have signed an agreement on mutual recognition of higher education qualifications and degrees, and the education sector has established a regular consultation mechanism. The Confucius Institute Headquarters of China has jointly established 12 Confucius Institutes with Australian institutions, and Chinese language courses have been offered in more than 600 primary and secondary schools. Mandarin became the second most spoken language in Australia after English. A joint meeting mechanism was established between nine Chinese universities and eight Australian university consortia. China Scholarship Council, Australian National University, University of Sydney, and other universities have launched joint doctoral programs. Monash University and Southeast University in China established a graduate school in Suzhou, becoming the first Australian institution of higher learning to establish a campus in China.

One of the questions I have been thinking about recently is whether Sino-Australian educational cooperation can achieve sustainable development at the current level. I believe this is also a topic of

concern for our friends here. This issue seems urgent in the current context of deepening globalization and the “Asian Century” knocking at the door.

I would like to bring you two messages. One is that with China’s rapid economic growth and rising livelihood, the number of Chinese students studying abroad has accounted for one-seventh of the world, ranking first in the world. Australia is expected to get a bigger share of the dividends from the opening up of Chinese education.

Second, developed countries such as the US, Europe, Canada, and Japan are increasingly aware of the Chinese student population’s economic value and humanistic effects and have introduced preferential policies and facilitation measures. Coupled with the high Australian dollar and other factors, Australia is facing unprecedentedly fierce international competition in recruiting Chinese students.

As China’s ambassador to Australia, I would like to make some practical suggestions for future Sino-Australian education cooperation.

First, I hope the Australian government and universities will attract Chinese students. Australia can create better conditions and reliable security for Chinese students in terms of entry, tuition fees, living and internship to keep the temperature of “stay in Australia fever”.

Secondly, we hope Australia will support two-way study exchange from the policy level. The number of international students between China and Australia is more than 50 to 1. The US government has recently launched the 100,000 Strong Initiative, which will send 100,000 students to China in the next four years. Can Australia send 10,000 students? I hope friends from all walks of life will provide support.

Third, I hope Australia will embrace more Chinese language and Chinese culture, encourage more institutions to establish Confucius Institutes, Confucius Classrooms and Chinese language courses, and train more friendly passers-by who speak Chinese and understand China. China is willing to assist actively.

Fourth, we need to explore new paths for China-Australia education cooperation, such as promoting more qualified institutions of higher learning in China and Australia to conduct cooperative education, exchange and learn from each other, and to complement each other’s advantages.

Thank you all!

## Appendix B-2 Post-Intervention Interpreting Speech

### 岑建君在“中澳跨境教育与学生流动论坛”上的讲话

大家上午好!

我和同事很高兴参加今天的研讨会。首先我代表教育部国际合作与交流司对此次会议的召开表示热烈的祝贺。

我国历来重视与澳大利亚的教育交流，这是两国双边关系重要组成的部分。自我国改革开放以来，两国在科研合作、学生交流等各领域的交流不断拓展，涉及高等教育、基础教育、职业教育等各个层次，特别是学生交流规模不断扩大。在两国教育主管部门的努力下，中澳曾于 1986 年签订了里程碑式的合作协议，去年双方又重新签订了学历互认协议，极大地推动了双向学生交流，促进了中国学生赴澳留学以及澳大利亚学生来华留学。

中澳合作办学也在不断地扩展和推进。据统计，目前两国合作举办的机构和项目共 396 个。我认为，这个数字还有进一步增长的空间。目前中澳合作办学仍仅局限于高等院校，职业技术教育的合作还很有限。所以，我希望此次会议能够帮助双方院校进一步挖掘潜力，扩大职业教育领域的合作，让在转型建设中的中国应用技术大学能更好地了解澳大利亚院校的教学方法，学习其教学经验，为中国培养更多的实用性人才。

去年，中国召开了‘全国留学工作会议’。这是建国 65 年来，首次召开的出国留学和来华留学两个议题并重的全国性会议。会议提出了今后重点培养的 5 类人才，分别是创新人才、小语种人才、国际组织人才、国际区域研究人才和来华留学生。另一方面，我们也希望更多的外国优秀学生来华留学，通过学生交流，提升中国的影响力。2014 年，包括获得学位及短期访问在内的来华留学学生人数共有 37.7 万人。到 2020 年，我们希望这一数字能达到 50 万人，并且，其中能有 15 万人是来华攻读学位的。我们很高兴澳大利亚能够积极参与我们扩展来华留学生的计划。澳方创立的新科伦坡计划给其他西方国家开了一个好头，虽然来华人数虽然不多，但是可以被看做是一种信号。中方也会高度关注新科伦坡计划在中国的实施，并会努力使这些学生能够在中国平安留学、成功留学。



我很高兴听到澳大利亚朋友说中澳的教育交流发展得很好，我认为两国的合作拥有一个很好的过去，而且现在仍在不断进步，未来也必将会更加完美。谢谢大家！

## **Speech by Jianjun Cen at the ‘China-Australia Cross Border Education and Student Mobility Forum’**

[Reference translation]

Good morning, everyone!

My pleasure to attend today’s seminar. First of all, on behalf of the Department of International Cooperation and Exchange of the Ministry of Education, I’d like to extend my warm congratulation on the convening of this meeting.

China always attaches importance to educational exchanges with Australia, which is an integral part of the bilateral relations between the two countries. Since our country’s reform and opening up, exchanges between the two countries in various fields such as scientific research cooperation and student exchanges have been expanding. They have covered all levels of higher education, primary education and vocational education. In particular, the exchange of students has continuously grown. With the efforts of the education authorities of the two countries, China and Australia signed a landmark cooperation agreement in 1986. Last year, both sides renewed the agreement on mutual recognition of academic qualifications, which significantly promoted the exchange of two-way students and promoted the study of Chinese students in Australia and Australian students studying in China.

Sino-Australian cooperation in running schools is also continually expanding and promoting. According to statistics, there are currently 396 agencies and projects jointly organised by the two countries. I think there is room for further growth of this figure. At present, the cooperation between China and Australia in running schools is still limited to institutions of higher learning. The partnership in vocational and technical education is still limited. Therefore, I hope this meeting can help both nations to tap potentials further and expand cooperation in vocational training so that the China University of Technology in transition can better understand the teaching methods of Australian institutions, learn from their teaching experience, For China to train more practical talents.

Last year, China held a national study abroad conference. This is the first national conference in the last 65 years after founding the People’s Republic of China to study abroad and study in China. The conference proposed five types of talents to be specially cultivated in the future: innovative skills, talents of minority languages, international organizations, international, regional researchers and students studying in China. On the other hand, we hope that more international students will study in China and enhance their influence through student exchanges. In 2014, 377,000 students studied in China, including degree and short-term visits. By 2020, we hope that this figure will reach

500,000 people and that of these 150,000 people will come to China to study for a degree. We are delighted that Australia can actively participate in our plan to expand overseas students in China. The new Colombo Plan created by the Australian side has given other Western countries a good start. Although the number of arrivals in China is small, it can be regarded as a signal. China will also pay considerable attention to implementing the New Colombo Plan in China and will work hard to ensure that these students can study safely in China and study abroad successfully.

I am glad to hear Australian friends say that the education exchanges between China and Australia have developed very well. I think the cooperation between the two countries has a magnificent past and is still undergoing continuous improvement and will undoubtedly become even more perfect in the future.

Thank you all!

## Appendix C-1 Initial and Follow-Up Interview Questions

Initial Interview Questionnaire	Follow up Interview Questionnaire
<p><b>1. How important to you is your voice as a means of expressing yourself and in the field of interpreting? (Very important/ Important/ Not important/Not sure.) Why?</b></p> <p><i>Probe: Let the interviewees tell what they understand about their voice to be to their life and profession</i></p> <p><i>prompt: to set a scene for the following questions and to raise interviewees' spoken voice awareness</i></p>	<p><b>1. How important to you is your voice as a means of expressing yourself and in the field of interpreting? (Very important/ Important/ Not important/Not sure.) Why?</b></p> <p><i>Probe: Let the interviewees tell what they understand about their voice to be to their life and profession</i></p> <p><i>Prompt: to compare interviewees' understanding of spoken voice to their profession with the initial interview questionnaire</i></p>
<p><b>2. Do you often record your interpreting and listen afterwards? If so, do you enjoy your spoken voice?</b></p> <p><i>Probe: let the interviewees reflect on their voice</i></p> <p><i>Prompt: To raise awareness of voice consciousness</i></p>	<p><b>2. Do you enjoy listening to your interpreting recording? Why or why not?</b></p> <p><i>Probe: let the interviewees reflect on their voice</i></p>
<p><b>3. Which do you believe are the most important spoken voice skills within an interpreting setting? (Intonation, fluency, pitch, dynamism, voice-volume, care of voice, elocution, articulation, tone, appropriate pause and speed, connection with audience, confident, poise.) Why?</b></p> <p><i>Probe: Let the interviewees tell what they understand about voice skills to their interpreting study and work</i></p> <p><i>Prompt: to raise interviewees' awareness spoken voice of skills and self-reflection</i></p>	<p><b>3. Which do you believe are the most important spoken voice skills within an interpreting setting? (Intonation, fluency, pitch, dynamism, voice-volume, care of voice, elocution, articulation, tone, appropriate pause and speed, connection with audience, confident, poise.) Why?</b></p> <p><i>Probe: Let the interviewees tell what they understand about voice skills to their interpreting study and work</i></p> <p><i>Prompt: to compare interviewees' understanding of spoken voice skills with their initial interview questionnaire</i></p>
<p><b>4. Have you ever taken any spoken-voice related training or lesson or activities before?</b></p> <p><i>Probe: -When and where did it happen? -How did you think about the training? -Does the training have any effect on your</i></p>	<p><b>4. What have you done to your spoken voice after the first interview, such as undertaking a spoken-voice related training or lesson or initiate a self-practice?</b></p> <p><i>Probe: -When and where did it happen?</i></p>

<p><i>spoken voice skills?</i></p> <p><i>Prompt: to understand interviewee's experience of voice training</i></p>	<p><i>-How did you think about the training?</i></p> <p><i>-Does the training have any effect on your spoken voice skills?</i></p> <p><i>Prompt: any change in action?</i></p>
<p><b>5. How would you rate your spoken voice quality like now?</b> <i>(Excellent/Good/Average/Poor/Not sure)</i></p> <p><i>Probe: Would you like to elaborate on that?</i></p> <p><i>Prompt: self- reflection of voice quality; the interviewer is to translate the concept of 'VOICE MANAGEMENT'</i></p>	<p><b>5. How would you rate your spoken voice quality like now after the first interview?</b> <i>(Excellent/Good/Average/Poor/Not sure)</i></p> <p><i>Probe: Is there any improvement since the last interview? Would you like to elaborate on that?</i></p> <p><i>Prompt: any change in action?</i></p>
<p><b>6. Does your voice fatigue sometimes whilst interpreting practising?</b></p> <p><i>Probe: -If yes, when was the last time that happened?</i></p> <p><i>-What did you do when that happened?</i></p> <p><i>-How do you look after your voice as an interpreting student?</i></p> <p><i>Prompt: reflection of voice care; the interviewer is to translate the concept of 'voice hygiene'</i></p>	<p><b>6. Do you still experience vocal fatigue or discomfort after or during an interpreting practise? How would you look after your voice so it would remain in good condition after a long interpreting practising?</b></p> <p><i>Probe: -If yes, when was the last time that happened?</i></p> <p><i>-What did you do when that happened?</i></p> <p><i>Prompt: reflection of voice care; any change practice of vocal care?</i></p>
<p><b>7. Do you think that you must force the volume of your voice to be heard in interpreting practising situation?</b></p> <p><i>Probe: If yes, where did that happen? Could you give me an example?</i></p> <p><i>Prompt: Reflect on voice projection</i></p>	<p><b>7. How is your voice projection right now? Do you still feel like you need to use force to push the volume of your voice to be heard in interpreting practising situation?</b></p> <p><i>Probe: If yes, where did that happen? Could you give me an example?</i></p> <p><i>Prompt: Reflect on voice projection; any change practice of voice projection?</i></p>

<p><b>8. As an interpreting student, do you think you communicate well with your interpreting audience (your classmate or your instructor) by using spoken voice skills such using as pitch, speed, tone, pause, vocal variety, eye contact etc. (not just linguistically)?</b></p> <p><i>Probe: Could you elaborate on these skills? Which do you think you are doing well now? Which skills you need to improve?</i></p> <p><i>prompt: Reflect on vocal dynamism /variety</i></p>	<p><b>8. Is there any improvement in your communicating with your interpreting audience (your classmate or your instructor) by purposefully using spoken voice skills such using as pitch, speed, tone, pause, vocal variety, eye contact etc. (not just linguistically)?</b></p> <p><i>Probe: Could you elaborate on these skills? Which do you think you are doing better now? Which skills do you still need to improve?</i></p> <p><i>prompt: Reflect on vocal dynamism /variety Any change?</i></p>
<p><b>9. Are you comfortable delivering interpreting to a big audience in a public place?</b></p> <p><i>Probe: Yes/No, why? Is there any form of interpreting that you prefer and feel most comfortable with? I.e. business, medical, court interpreting, conference? Why is this?</i></p> <p><i>Prompt: Reflect on confidence of voice skills</i></p>	<p><b>9. Do you feel that you are more comfortable delivering interpreting to a big audience in a public place now?</b></p> <p><i>Probe: Yes/No, why? Is there any change of form of interpreting that you prefer and feel most comfortable with? I.e. business, medical, court interpreting, conference? Why is this?</i></p> <p><i>Prompt: Reflect on change of confidence of voice skills</i></p>
<p><b>10. Do you often feel nervous when interpreting for your instructor and classmates or other audience? If so, how do your control your physical tension and keep interpreting in a calm manner?</b></p> <p><i>Probe: when you got nervous, how did you feel about it? How did that come about?</i></p> <p><i>Prompt: Reflect on physical tension</i></p>	<p><b>10. Do you still feel nervous when interpreting for your instructor and classmates or other audience? If so, how do your control your physical tension and keep interpreting in a calm manner?</b></p> <p><i>Probe: when you got nervous, how did you feel about it? How did that come about?</i></p> <p><i>Prompt: Reflect on management of physical tension</i></p>
<p><b>11. How is your breathing while speaking or in an interpreting setting?</b></p> <p><i>Probe: How did you notice your breath? How did that feel to you?</i></p> <p><i>Prompt: reflect on breath management</i></p>	<p><b>11. How is your breathing while speaking or in an interpreting setting?</b></p> <p><i>Probe: if so, what would you do now?</i></p> <p><i>Prompt: reflect on breath management</i></p>

<p><b>12. As an interpreting audience, have you witnessed any interpreter who appear to be impressive or disappointing in terms of their spoken voice skills? If so, what attributes did they display? (For example: confident, physical movement/poise/voice quality/volume/rate of speech/fluency/capacity to convey speaker's role/clear articulation/good communication/good speaking skills)</b></p> <p><i>Probe: Could you say some more about that? Any example you can give?</i></p> <p><i>Prompt: reflect on other's interpreting practise</i></p>	<p><b>12. Do you feel you have improved since first interview in terms of your voice projection, intonation, pitch, fluency, communication etc? Do you believe this could have been improved even more if you were provided with voice training in your poSTGraduate course? Yes/no. Why?</b></p> <p><i>Probe: could you elaborate on what kind of spoken voice training are most important for your interpreting practise?</i></p> <p><i>Prompt: reflect on change of self-practice</i></p>
<p><b>13. Do you think that spoken voice training should or should not be provided to interpreting students in conjunction with interpreting training? Yes/No. Why?</b></p> <p><i>Probe: to understand interviewees' attitude of voice training for interpreting students</i></p> <p><i>Prompt: The interviewer is to impart the voice training concept to interviewee</i></p>	<p><b>13. Do you still believe/or not that spoken voice training should be provided to interpreting students in conjunction with interpreting training? If yes, do you think spoken voice training should be taken a selective course or integrated into a main interpreting course?Yes/No. Why?</b></p> <p><i>Probe: to understand interviewees' attitude of voice training for interpreting students</i></p> <p><i>Prompt: any change of attitude?</i></p>
<p><b>14. If there is voice training provided in the interpreting course or outside of the course, are you willing to participate the training and/or to practice at your own time? Yes/no. Why?</b></p> <p><i>Probe: could you elaborate on what kind of spoken voice training are most important for your interpreting practise?</i></p> <p><i>Prompt: reflect on self-practice</i></p>	<p><b>14. If there is voice training provided in the interpreting course or outside of the course, are you willing to participate the training and/or to practice at your own time? Yes/no. Why?</b></p> <p><i>Probe: could you elaborate on what kind of spoken voice training are most important for your interpreting practise?</i></p> <p><i>Prompt: change of attitude</i></p>

## Appendix C-2 Self-Evaluation Sheet

Performance criteria	Rating scale (initial)	Comments (initial)	Rating scale (follow-up)	Comments (follow-up)
Good posture and movement in standing, sitting or way of holding notebook?	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Present with reasonable tension and generally with ease	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Appropriate volume	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Smooth breath while delivery	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Not too much tension shown in your face and jaw	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Fluency (less hesitation)	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Clear articulation	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	



Your pronunciation and accent are easy to understand and do not hinder audience's understanding	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Pleasant voice quality	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Dynamic and expressive (intonation and emphasis)	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Vocal variety	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Pitch range	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Appropriate delivery speed and using of pause (instead of fillers)	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Audience communication (eye contact, smile, facial expression, etc)	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Confident and professional looking	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	

## Appendix D-1 Voice Training Schedule for LTG

SESSION DATE	TRAINING TOPIC	TRAINING ACTIVITES AND EXERCISES	LEARNING OUTCOME
<p style="text-align: center;"><b>Session 1</b> <b>April 16, 2015</b></p>	<p>Voice and speech production</p> <ul style="list-style-type: none"> <li>- Posture and alignment</li> <li>- Movement</li> <li>- Speaking skills</li> </ul>	<ol style="list-style-type: none"> <li>1. Introduction of the voice training workshops and the research project (5 min)</li> <li>2. Get to know your spoken voice (25 min)</li> <li>3. Body stretch and vocal care (30min)</li> <li>4. Self - Reflection and Group Discussion (20min)</li> </ol> <p>Exercises:</p> <ol style="list-style-type: none"> <li>1) Ice breaking activity;</li> <li>2) Impromptu speech (1-3 min) in English and feedback;</li> <li>3) Mini-seminar: spoken voice apparata and speech organs;</li> <li>4) Gentle body stretch and vocal warming up exercises.</li> </ol>	<ul style="list-style-type: none"> <li>- Explore spoken voice and raising vocal awareness.</li> <li>- Understand basic vocal production from - - physical side and vocal warm-up activities.</li> <li>- Training improvisation and public speaking ability</li> <li>- Improved self-monitoring and self-reflective skills</li> </ul>
<p style="text-align: center;"><b>Session 2</b> <b>MAY 8, 2015</b></p>	<p>Voice and speech production</p> <ul style="list-style-type: none"> <li>- posture and alignment</li> <li>- Movement</li> <li>- Breathing skills</li> <li>- Speech muscles</li> <li>- tension reduction</li> </ul> <p>Voice and speech quality</p> <ul style="list-style-type: none"> <li>-Voice projection</li> <li>- Resonance</li> <li>- Vocal care</li> </ul>	<ol style="list-style-type: none"> <li>1. Integrating spoken voice, breath, and movement (40min)</li> <li>2. Emotional release for interpreter(30min)</li> <li>3. Voice projection and resonance (30min )</li> <li>4. Reflection and group discussion (20min)</li> <li>5. Interpreters' vocal care and endurance (30min)</li> </ol> <p><u>Exercises:</u></p> <ol style="list-style-type: none"> <li>1) balance and alignment (siting and standing posture) instruction; neck, jaws and shoulder stretches/side stretches/Breath exercises</li> <li>2) Mini seminar: vocal care</li> <li>3) Breathing support and tension release exercises</li> </ol>	<ul style="list-style-type: none"> <li>- Raising vocal care awareness</li> <li>- Understand basic vocal production from physical side</li> <li>- Tension release activity</li> <li>- Improved self-monitoring and reflective skills</li> </ul>

		4) Throat and neck release on sound exercises	
<b>SESSION 3</b> <b>MAY 22, 2014</b>	Voice and speech quality and voice and speech dynamism - Articulation - Audience connection - Professional confidence	1. Vocal dynamism (20min) 2. Audience connection, confidence and persuasion (40 min) 3. Articulation exercise and English accent reduction and neutralisation for interpreters from Chinese ethnic background (40min) 4. Reflection and group discussion (20min)  <u>Exercise:</u> 1) Fluency, intonation patterns, rhythm, pace, pauses, emphasis, expressiveness 2) Overcome self-consciousness confidence, eye contact, authoritative voice exercises 3) Exercises of speak with clarity English; Chinese accent reduction and English accent neutralisation	- Raising awareness of vocal variety - Improve vocal quality - Improve public speaking ability - Improve audience awareness - Improved self-monitoring and self-reflective skills
<b>SESSION 4</b> <b>JUNE 4, 2014</b>	Voice and speech quality and voice and speech dynamism - Articulation - Audience connection - Professional confidence - Integrating voice training and interpreting sessions	1. Audience connection, confidence and persuasion (30 min) 2. Articulation exercise and English accent reduction and neutralisation for interpreters from Chinese ethnic background (30min) 3. Integrate voice techniques and interpreting practices (30min) 4. Group discussion (60min) 5. Reflection and feedback	- Raising vocal awareness and applying vocal skills into interpreting practice - Raising awareness of vocal variety - Improve vocal quality - Improve audience awareness - Improved self-monitoring and self-reflective skills

## Appendix D-2 Voice Training Schedule for STG

SESSION/DATE	TRAINING TOPIC	TRAINING ACTIVITES AND EXERCISES OUTLINE	LEARNING OUTCOME
Session 1 June 15, 2015	Voice and speech production Introducing voice research project	- Introduction of the voice training workshops and the research project (5 mins) - Get to know your spoken voice (10 mins)/ Ice breaking activities - Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Understand basic vocal production from physical side and vocal warm-up activities. - Improved self-monitoring and reflective skills.
Session 2 June 16, 2015	Voice and speech production Introducing vocal and physical warm up	- Warming up exercise (3 mins) - Impromptu speech (12 min) - Group discussion (5 min)	- Explore spoken voice and raise vocal awareness. -Improve improvisation and public speaking ability -Improve self-monitoring and reflective skills.
Session 3 June 17, 2015	Introducing the structure of chest and throat and the passage of the breath.	- Voice and speech production seminar (10mins) - Breathing activities (5min) - Group discussion (5 min)	- Explore spoken voice and raise vocal awareness. - Improve breathing techniques - Improve self-monitoring and reflective skills.
Session 4 June 22, 2015	Introducing muscular support for spoken voice	- Balance and alignment instruction (siting and standing posture) (10mins) -posture and alignment exercises for interpreters- stretching, releasing and alignment (5 mins) - Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve Posture and alignment techniques - Improve self-monitoring and reflective skills.
Session 5 June 23, 2015	Introducing speech organs and muscles	- Speech production organs instruction (10 mins) -Speech production organs such as lips, mouth vocal folds, lips, tongue, larynx, pharynx (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve speech production skills - Improve self-monitoring and reflective skills.

		- Group discussion (5 mins)	
Session 6 June 24, 2015	Resonators (neck, mouth and throat)	- Resonator instruction and exercises including neck, jaws and shoulder stretches/side stretches (15 mins) -Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve speech production skills - Improve self-monitoring and reflective skills.
Session 7 June 29, 2015	Interpreters' vocal care and endurance	-Vocal care instruction and warming up exercises for interpreters (15 mins) - Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve vocal care awareness - Improve self-monitoring and reflective skills.
Session 8 June 30, 2015	Interpreters' vocal care and endurance	-Vocal warming activities (15 mins) - Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve vocal care awareness - Improve self-monitoring and reflective skills.
Session 9 July 1, 2015	Breathing supporting interpreting work	- Breathing support and physical tension release exercises (15 mins) -Group discussion (5 mins)	- Explore spoken voice and raising vocal awareness. - Improve breathing techniques. - Improve self-monitoring and reflective skills.
Session 10 July 6, 2015	Musculoskeletal supporting interpreting work	- Stretching big muscle and releasing and aligning exercises (15 mins) - Group discussion (5 mins)	- Explore spoken voice and raising vocal awareness. - Improve breathing techniques. - Improve self-monitoring and reflective skills.
Session 11 July 7, 2015	Speech organs support interpreter's clear speech.	- Stretching and releasing lips, mouth, vocal folds, lips, tongue exercises (15 mins) - Group discussion (5 min)	- Explore spoken voice and raising vocal awareness. - Improve speaking clarity techniques. - Improve self-monitoring and reflective skills.
Session 12 July 8, 2015	Resonators such as neck, mouth and throat support better voice quality	- Resonator release and stretches including neck, jaws and shoulder, side stretches, throat and neck (15 mins) - Group discuss (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve body awareness. - Improve self-monitoring and reflective skills.
Session 13 July 13, 2015	Explore interpreter's emotional challenge and emotion release	- Stress releasing activities (10 mins) - Integrating breathing support and tension release exercises (5 mins) - Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Improving breathing and tension reduction techniques. - Improve self-monitoring and reflective skills.
Session 14 July 14, 2015	Voice and speech production: Overcome stage freight and	- Integrating breathing support and tension release exercises (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve tension reduction techniques.

	performance anxiety/ Interpreter's psychological tension reduction and adjustment.	- Overcome self-conciseness and stage freight and negative emotion (10 mins) - Group discussion (5 mins)	- Improve self-monitoring and reflective skills.
Session 15 July 15, 2015	Modulation of pitch range	Tension reduction exercises (5 mins) - Pitch range exercise (10 mins) - Group discussion (5 mins)	- Explore spoken voice and raise vocal awareness. - Improve pitch modulation skills - Improve self-monitoring and reflective skills.
Session 16 July 20, 2015	Resonating and voice projecting	- Chest resonating exercises (5 mins) - Voice projection exercises (10 mins) - Group discussion (5 mins)	- Improve voice quality - Improve voice projection skills - Improve self-monitoring and reflective skills.
Session 17 July 21, 2015	Speaking pace and pause	- Pace and pause with text (15 mins) - Group discussion (5 mins)	- Improve voice quality - Improve speech pace and pause - Improve self-monitoring and reflective skills.
Session 18 July 22, 2015	Integrating voice production skills and interpreting skills	- Students presentation of speaking (15 mins) - Group discussion (5 mins)	- Improve voice quality - Apply voice and speech skills into interpreting - Improve self-monitoring and reflective skills.
Session 19 July 27, 2015	Intonation and rhythm exercise	- Exercises of intonation, speech rate, rhythm and pauses with reading speech text aloud (15 mins) - Group discussion (5 mins)	- Improve voice dynamism - Improve intonation, rhythm, speech rate and pauses - Improve self-monitoring and reflective skills.
Session 20 July 28, 2015	Audience connection and showing confidence and persuasion through voice and speech	- Public speaking exercise (overcome self-conciseness when speaking in public showing confidence, eye contact, authoritative voice and expressiveness) (15 mins) - Group discussion (5 mins)	- Improve voice dynamism - Improve audience awareness and confidence - Improve self-monitoring and reflective skills.
Session 21 July 29, 2015	Articulation and accent neutralisation for Chinese interpreters	- Speaking clarity (vows and consonants) (5 mins) - Speaking with clarity exercises with text (10 mins) - Group discussion (5 mins)	- Improve speech clarity - Improve speaking skills - Improve self-monitoring and reflective skills.
Session 22	Articulation and accent	- English accent neutralisation (15 mins)	- Improve speech clarity

Aug 3, 2015	neutralisation for Chinese interpreters	- Group discussion (5 mins)	- Accent reduction - Improve self-monitoring and reflective skills.
Session 23 Aug 4, 2015	Integrate vocal dynamism and clear speaking into text	- Speaking clarity with text and interpreting exercise (15 mins) - Group discussion (5 mins)	- Improve speech clarity - Accent reduction - Improve self-monitoring and reflective skills.
Session 24 Aug 5, 2015	Integrate voice techniques and interpreting practices/videos/ Recap	- Speaking and interpreting role play (15 mins) - Group discussion (5 mins)	- Apply voice and speech skills into interpreting - Improve self-monitoring and reflective skills.

## Appendix D-3 Spoken Voice Exercises for Interpreting Students

### **Exercise 1: Ice-breaking activity**

Purpose of the Exercise: to help students connect emotion, body and voice when interacting with people around

Students introduce themselves to someone in the room with whom they do not associate very often. The instructor reminds the students to pay attention to their physical sensations and voice when speaking. In a group discussion, the students reflect and discuss their experiences.

### **Exercise 2: Drawing body image (otherwise referred as body mapping)**

Purpose of the Exercise: to help students raise their awareness about physical aspects of voice production. This activity may divide into several steps, as follows:

**Step 1:** Students work in pairs. They are instructed to draw a line and the lateral view of their own body, such as spinal curve, head and neck position on the line. After that, they draw for each other. They may find that the spinal curve and neck position is different from what they think of themselves. Then, the instructor shows them a model full skeleton front view and side view or picture of the skeleton to explain the head and neck body, joints, hip, knees alignment and explain their importance to voice.

**Step 2:** Students are still in pairs. One student is observing while the other walks around the room. The observing student takes notes of his/her partners' posture and the position of chest, pelvis, head, and where tension seems to be held in their body. Then the two students swap.

**Step 3:** Students sit in a circle and discuss what they have observed about themselves and their partner's body movement.

### **Exercise 3: Posture and alignment and vocalization exercises**

Purpose of the Exercise: to help students be aware of their posture while standing and seating  
to help students correct their posture  
to help students experience the connection of posture and spoken voice

**Reading aloud with a standing position:** the students stand tall with a long spine and neutrally aligned. Their shoulders blades are naturally slopping down their back, but not to thrust chest forward purposefully, imagine their head a hot balloon balanced at the top of their spine.

Students read aloud a text/speech with the standing position.

**Reading aloud with a seated position:** The instructor guides the students to sit on their sitting bones, fully supported by their muscles and lumbar spine but not their lower back. Their knees be a little apart below the pelvis. The students imagine they grow tall when they speak to create more space in the rib cage for the lungs to inflate and give them all the power they need. With this position, students read aloud a text/speech. In a group discussion, the students reflect on their sensation of reading the text aloud in both standing and seated positions.



#### **Exercise 4: Kinaesthesia Exercise**

Purpose of the Exercise: to help students be aware of posture while moving around  
to help students experience the connection of movement and spoken voice

First, students work in pairs. One student walks across the diagonal of the workshop room, and the other observe their partner' walking position, including the shape of the spines, the habitual head and neck alignment, the way their arms hang from their shoulders, the way they use their pelvis, the weight balance in their hips, the tension in their knees and the placing of their feet, the way their use their body when self-conscious, the way they stand when they are nervous, and afterwards when they are relaxed.

Then, students discuss changing negative patterns and build up strength, balance and flexibility. So, their body could deal with any demand in the future.

Finally, students hold a text in hand and walked while reading the text aloud and paying attention to their sensation and movement. Then the students discuss in pairs their observations and reflection.

#### **Exercise 5: Semi-Supine position and vocalization**

Purpose of the Exercise: to help students be aware of posture  
to help students experience the connection of body and spoken voice  
to help reduce tension and performance anxiety

Students lie on their back on a hard surface with one or two legs bent without habitually arching their lower back. They may use one or two paperback books to support the head. The instructor guides the students to scan different parts of the torso and see any tension in any part of their body or limbs. Students keep their torso very still and relaxed.

The instructor guides students with a few steps: first, they breathe abdominally and exhale slowly through their mouths. Second, they concentrate on relaxing their lips and jaw by allowing them to hang open. Third, they relax their throat. Each time they exhale, they may feel the passage of air through their throat and mouth is smooth and unobstructed. Then, in the same position, the students exhale with a breathy sigh and say 'ah'. The students repeat and open their mouths wide as if they are yawning. With their next series of exhalations, they keep sighing, but this time count slowly up to five. The students repeat the exercise until they can pronounce each word without strain.

Then, the instructor gives directions to the students to do different tasks with their limbs. For example, "let the head go forward and up," "let the back lengthen and widen" and so on. In the meantime, the instructor asks the students to vocalise, for example, whispering "Ah", or read a text, recite a poem, sing a tune or count numbers. Students are directed to be aware of the sensation of their limbs and body when they vocalise.

Lying down can be an advantageous position to practise initially, as students will get feedback from the hard surface so that they can notice more quickly if they have habitual physical attention such as tightening their necks and tilting their head back when they vocalise, which poses an increased pressure on their neck and head. Students are encouraged to practise the semi-supine position daily to strengthen the conscious awareness of body-mind integration. The students were encouraged to recite a little poem or a passage while in the semi-supine position and observe their sensations.

After the semi-supine position, students go back to a seated position, repeat the above exercises. Then they count to five using their full voice. If they sense constriction in their throat, repeat the exercises while lying on their back.

Students do the above exercises in a standing position. Eventually, the students could produce open phonation while counting to 100 and taking in a new breath after every five numbers.

### **Exercise 6: Physical and vocal warming up and vocalising exercises**

Purpose of the Exercise: to help the students warm up physically and vocally to prepare for challenging interpreting tasks  
to help students reduce tension and performance anxiety

1. Throat massage: students knead the throat muscles lightly with their fingers to eliminate tightness.
2. Shake out tension: students wiggle their wrists and gradually move their whole body and swing arms.
3. Releasing of speech muscle:
  - Jaw tension release: students place their fingertips just by and in front of their ears on both sides of their face and open their mouth. Massage these joints using their fingertips or the palms of their hands. Release their jaw further and massage deeper with each exhale. Students massage jaw muscles while saying ‘ya, ya, ya,’ and ‘yeah, yeah, yeah’.
  - Tongue waggle and funny faces: students stand back to each other and stretch their tongue out of their mouth as far as they can in all directions. Make as many funny faces as possible to exercise all of their facial muscles and stretch their cheeks. (the reason to instruct the students to stand back to each other is some of them may feel uncomfortable and embarrassed)
  - relaxing tongue by playing with pitch and lip flex. students do tongue dart/furrow/stretch on their own.
  - Face stretch: instruct the students to tighten the muscles on a frown as they say “Oh no!” and then release them on a big smile, to “Oh yes!” . The students pretend to chew sticky toffee to ease a dry mouth.
4. Shoulder and neck massage and yawn a few times to open up the back of the throat.
5. Peach picking exercise: students stretch intercostal and open ribcage for free breath and sound.
6. Pelvic stir exercise: students wiggle their pelvis and engage the lower body.
7. Releasing of vocal folds (This exercise warms the muscles and the vocal folds safely and helps increase resonance):
  - Students hum gently to ‘mm’, feel the vibration in the upper chest, throat, nose and chest. Start very quietly, gradually increasing volume.
  - Students hum and hold any pitch for ten or more seconds until feeling a tingling sensation around their lips and nose.
8. Buzz lips to make a “Brr” sound: students start on any pitch. Then go up and down their vocal range.
9. Vocalize on an open ‘Ahh’ sound: students go up and down their vocal range
10. Open the pharynx, relax the throat and increase resonance.
  - Instruct the students to smile and yawn loudly
  - Students look in the mirror and say ‘Ah, there you are!’ in warm tones.
11. Students make an impromptu speech aloud on any topics that is familiar to them.

## **Exercise 7: Breath and vocalise exercises**

Purpose of the Exercise: to ground and deepen the breath

to improve alignment

to increase body flexibility and enhance sense of empowerment

1. Students sit comfortably in the chair, with hands in lap and feet on the floor. The instructor guides them to relax back into the chair. The students become aware of the back resting comfortably against the back of the chair and experience the weight of buttocks resting easily on the surface of the chair. Their lips are lightly closed—tongue flat. The tip of the tongue is resting behind the bottom front teeth.
2. The instructor guides the students to focus on the breath. The students breathe in through nostrils and into abdomen and back. Imagine inhalation filling up the abdomen and the ribcage. The instructor reminds the students to be aware of the full 360 of the rib cages. The students imagine the diaphragm moving with each inhalation and exhalation. The students begin to feel increasingly “grounded” as they experience breath dropping into a deeper part of the body. As they continue to breathe in through nostrils, they experience an increasing sense of autonomy and strength.
3. The instructor guides the students to exhale on voiceless fricative “f” or “s” while engaging abdominal muscles to support exhalation. Repeat three times.
4. Students exhale on voiced fricative “v” or “z”. Repeat three times.
5. Students breathe into the lower back.
6. Students repeat each line after the instructor. The students take a quick, deep, and silent breath after each line to ensure enough breath to sustain the thought so that the volume will not fade out. The instructor adds words to make the sentence longer and longer to build up the students’ breath control. Below is an exercise is adapted from Rodgers 2002, p.33)

E.g., “Hello!”

“Hello, I am a feminist!”

“Hello, I am a feminist because it bothers me!”

“Hello, I am a feminist because it bothers me that a woman gets killed!”

“Hello, I am a feminist because it bothers me that a woman gets killed by her male partner!”

“Hello, I am a feminist because it bothers me that a woman gets killed by her male partner every single week!”

“Hello, I am a feminist because it bothers me that a woman gets killed by her male partner every single week, and one in four Australian women!”

“Hello, I am a feminist because it bothers me that a woman gets killed by her male partner every single week, and one in four Australian women has experienced emotional abuse!”

“Hello, I am a feminist because it bothers me that a woman gets killed by her male partner every single week, and one in four Australian women has experienced emotional abuse by a current or former partner!”

(This exercise is adapted from Rodgers 2002, p.33)

(Notes: The training content is chosen from a speech against family violence. The reason to choose such a topic as “killing and family violence” is that violence against women is a topic that interpreters would often encounter in their work. In such an interpreting context, the interpreters are required to express the)

## **Exercise 8: Story telling**

Purpose of the Exercise: to help students improvise on speech

to help students understand structure of speech

Students sit in a circle and tell a story or make a short speech around. The aim is for each person to fully ‘receive what he/she has been given from the last person’ (Berry 1987:149) and then pass on his/her thought to the next person. Each student forms part of speaking and contributes to the whole text. The students move through the text aloud and focus on the sound instead of meaning. This provided a good stimulus for students improvising on the speech. The instructor guides the students’ awareness about the lengths of thoughts, the rhythms of short and long words, and the lengths of sounds in vowels and weight in consonants. The last aspect of this exercise is for the group to sit or stand still and speak the text quietly with clear intentions, paying attention to the phrasing structure. This consolidates the meaning of the text. (Adapted from Cicely Berry 1987)

### **Exercise 9: Resonance exercise**

Purpose of this exercise: to help students develop chest resonance in speaking

The instructor starts by reminding students of the places of resonance - head, nasal cavity, mouth, larynx, chest (upper ribs), lower ribs and the relations between resonance and release of tongue and jaws. Then, the instructor guides the students to put one hand flat and spread on their upper chest and gently hum, aiming to isolate the places of chest resonance by sending the sound to the chest and feeling the vibrations. “me me me me meh meh meh meh mai mai mai”.

The instructor guided the students to activate the resonator cavities - head, nasal cavity, mouth, larynx, chest (upper ribs), lower ribs - with gentle humming. Isolate the places of resonance by sending sound to them individually and feeling the vibrations.

“me me me me meh meh meh meh mai mai mai”

“hey, hi, how are you”

“nah nah nah nah nah”

EE – head

EH - mouth

AH - larynx

AW – chest/ upper ribs

OH – lower ribs and back

OO – pelvis

The instructor asks students to feel the vibrations under their hand and voice the numbers 1 to 10 with increased chest resonance. The students read a prepared speech in that chest resonant quality. Next, the instructor guides the students to put their hands on their lower ribs to feel the vibration at their back while they read the text. The students practise interpreting the speech in pairs with the same sensation of chest resonance and vibration in lower ribs and back. The students discuss in group their sensation of practice. The instructor asks questions, e.g. why have they not felt the resonance. Is it because they held breath while speaking or interpreting?

### **Exercise 10: Voice projection exercise**

Purpose of this exercise: to help students project and sustain sound to clarity speech

Varying the loudness using vowel sounds (maintain correct posture, alignment and not to strain the voice):

Step 1: Take a breath and say the vowel sound ah as follows:

1. ah..... (soft → loud) ↗
2. ah..... (loud → soft) ↘
3. ah..... (soft → loud → soft) ↗ ↘
4. ah..... (soft → loud → soft → loud → soft ..... ) ↗ ↘ ↗ ↘

Repeat with different vowel sounds

Step 2: Take a breath and say the vowel sound ah as follows:

- oo as in the word bood
- oh as in bode
- aw as in bawd
- ay as in bayed
- ee as in bead

Step 3: Varying the loudness using numbers:

1. Count and gradually increase the loudness  
**1 2 3 4 5 6 7 8 9 10**
2. Count and gradually decrease the loudness  
**1 2 3 4 5 6 7 8 9 10**
3. Count and increase the loudness on every 2nd number  
**1 2 3 4 5 6 7 8 9 10**
4. ... on every 3rd number  
**1 2 3 4 5 6 7 8 9**
5. ... on every 4th number  
**1 2 3 4 5 6 7 8 9 10 11 12**
6. ... on every 5th number  
**1 2 3 4 5 6 7 8 9 10**

Step 4: Voice projection using words and phrases:

Instruction: The students are instructed to exaggerate the vowel sounds' vocalisation by lengthening each vowel and continuing the voicing through the whole word. The instructor reminds the students that they are not increasing the muscle tension in the head and neck area or their general posture.

Single words (example)

1. moon – moan – mourn – mane – mean
2. mooning – moaning – morning – mining – meaning
3. memory – minimum – marmalade – manoeuvre – manager

### Phrases (example)

When speaking phrases, the students imagine that they are speaking just one long word and keep the vocalization going from the start of the phrase until its end. There should be an almost continuous voicing with no breaks between the words.

1. many men munch many melons
2. Mandy made marinade in May
3. Major Mickey's malt makes me merry
4. my mom's marvellous modern manicure
5. Mervin Maclean's mess marred my marmalade

### Sentences/speech

The instructor reminds the students to use the same intoning technique recite a sentence and a speech. The instructor also remind the students as to when they will take a breath.

### Loudness or volume variety

The instructor reminds the students to use diaphragmatic respiration to project volume. The students interpret a speech while paying attention to what words should be spoken loudly and which ones softly.

### Example of speech:

"I am not unmindful that some of you have come here out of great trials and tribulations. Some of you have come fresh from narrow jail cells. Some of you have come from areas where your quest for freedom left you battered by the storms of persecution and staggered by the winds of police brutality ... Go back to Mississippi, go back to Alabama, go back to South Carolina, go back to Georgia, go back to Louisiana, go back to the slums and ghettos of our Northern cities, knowing that somehow this situation can and will be changed. Let us not wallow in the valley of despair. I say to you today, my friends, that in spite of the difficulties and frustrations of the moment, I still have a dream. It is a dream deeply rooted in the American dream." (Excerpt from I Have A Dream)

## **Exercise 11: Articulation exercise**

Purpose of this exercise: to help improve speech clarity

to help students practice using strong, clear articulation of consonant and vowel sounds

Students sit in a circle and read aloud the following text. The instructor guides the students to exaggerate vowel and consonant for the first time and read in the average speed for the second time. The text is adapted from Cicley Berry's muscularity and word articulation exercises:

PA BA TA DA KA GA  
PAY BAY TAY DAY KAY GAY  
PEE BEE TEE DEE KEE GEE  
PAY BAY TAY DAY KAY GAY  
PA BA TA DA KA GA  
POR BOR TOR DOR KOR GOR  
POO BOO TOO DOO KOO GOO  
POR BOR TOR DOR KOR GOR  
PA BA TA DA KA GA  
MA MA LA BA BA LA  
MAY MAY LAY BAY BAY LAY

MEE MEE LEE BEE BEE LEE  
 MAY MAY LAY BAY BAY LAY  
 MA MA LA BA BA LA  
 MOR MOR LOR BOR BOR LOR  
 MOO MOO LOO BOO BOO LOO  
 MOR MOR LOR BOR BOR LOR  
 MA MA LA BA BA LA  
 HOOSKT HOHSKT HORSKT HAHSKT HAYSKT HEESKT  
 HOOKT HOHKT HORKT HAHKT HAYKT HEEKT  
 HOOSPT HOHSPT HORSPT HAHSPT HAYSPT HEESPT  
 HOOPT HOHPT HORPT HAHPT HAYPT HEEPT  
 HOOZBD HOHZBD HORZBD HAHZBD HAYZBD HEEZBD  
 HOOBD HOHBD HORBD HAHBD HAYBD HEEBD

Stretching tongue and lips:

Vowel sounds. EEE ERR AWW OOH AAH - 'Jean burnt all Sue's cards. Talk about the 'Vowel tree'. 'up up up into the clear blue sky' 'down down down into the depths of the deep dark pool.'

Lips stretch and pucker:

'Will you wait for Willie and winnie Williams?'  
 'Why are you walking when the weather's so wild?'

Open up the back of the mouth (Talk on the edge of a yawn):

'Goo gah gee gah gay gah guy gah go gah.'

Lip trills: (Release lip tension and connects breathing and speaking)

Instruct students to place their lips loosely together and release air in a steady stream to create a trill sound; The students hold the sound steady and keep the air moving past the lips. Next, encourage the students to repeat the b-trill slowly up and down the scales, but remind them not to go beyond what is comfortable at the top or bottom of the scale.

Tongue twisters

The instructor explains to the students that the aim of doing tongue twisters is to practise speaking without extra tension in the tongue, lips, jaw, or throat. Encourage the students to keep a good posture and full breath support when going through the exercises below.

Peter Piper picked a peck of pickled peppers.  
 A peck of pickled peppers Peter Piper picked.  
 If Peter Piper picked a peck of pickled peppers,  
 Where's the peck of pickled peppers Peter Piper picked?  
 Denise sees the fleece,  
 Denise sees the fleas.  
 At least Denise could sneeze  
 and feed and freeze the fleas.  
 One-one was a racehorse.  
 Two-two was one too.  
 One-one won one race.  
 Two-two won one too.

Vocalize vowels and consonants to various parts of body:

EE - head  
EH - mouth  
AH - larynx  
AW - chest/ upper ribs  
OH - lower ribs and ck  
OO - pelvis

Students read aloud a speech. For the first time, they only verbalize the vowel sounds in words and feel the lengths and difference between long and short ones. Then, they return to speaking the text naturally (adapted from Berry 1987; Hofmeyr, 2016).

### Exercise 12: English speech accent reduction exercise

Purpose of this exercise: to help reduce English speech accent  
to help improve speech clarity

1. Strengthening Word Stress & Emphasis,  
e.g.: product /'prɒdʌkt/  
colleague /'kɒli:g/
2. Improving English Weak Vowels (otherwise referred as Schwa sound),  
e.g.: carrot /'kærət/  
support /sə'pɔ:t/  
purchase /'pɜ:tʃəs/  
onion /'ʌnjən/
3. Distinguish English Vowel Sounds,  
\*Minimal pairs word exercises:  
e.g. sleep /i:/ slip /I/  
gem /e/ jam /æ/  
hut /ʌ/ heart /ɑ:/  
cushion /ʊ/ fusion /u:/  
\*Minimal pairs sentence exercises:  
e.g.: I don't like these spots.  
I don't like these sports
4. Commonly mispronounced words  
e.g.: recipe, salmon, says, suit (rhymes with boot), suite (rhymes with suite), chaos,  
women, nature
5. Learning speech sounds that aren't used in Mandarin,  
e.g.:  
'v' as in value or five;  
'z' as in authorise;  
'sh' as in share and punish;  
'zh' as in vision or measure;  
'ch' as in chairman or cheque;  
'j' as in judge or gentleman;  
unvoiced 'th' as in thing; and  
voiced 'th' as in this or that.
6. Learning combinations or 'clusters' of sounds found in English, but not used in Mandarin,



e.g.: words containing sl, sn, sm, br, cr, fr, bl, cl, spr, sq, nt.

## 7. Speech Melody and Intonation

Learning English sentence stress patterns to distinguish content words (like nouns and verbs) from function words (like prepositions and articles) and rules of intonation (e.g. to signal statements or yes/no questions).

e.g.: Mary 'went.

Mary would 'go

May would have 'gone.

### **Exercise 13: Pitch exercise**

Purpose of this exercise: to help explore and extend pitch range.

#### Pitch management exercise:

##### **Step 1**

Students sing the sound ah at an average volume. Gradually the students increase their volume until they feel themselves straining or losing clarity of sound. Then they repeat this process several times in smaller segments until they reach the maximum loudness level comfortable for them.

##### **Step 2**

The students sing the sound "ah" as in the preceding exercise, but this time do it at various higher and lower pitch levels. The students do this exercise for more than a few minutes and stop whenever they feel the strain. The instructor guides the students to periodically rest their voice by performing breathing and throat relaxation exercises.

##### **Step 3**

The students repeated the two preceding exercises, but this time recite letters of the alphabet, numbers, days of the week.

#### Pitch Range Exercise:

##### **Step 1**

Students inhale and exhale deeply five times before they begin their presentation. The instructor guides the students to take a deep breath and exhale vigorously into their first words. By breathing deeply between ideas, the students create a more energetic sound and more voice projection.

##### **Step 2**

The students read one sentence from a short speech with their natural pitch. Then, they read a second sentence with their highest pitch. Finally, they read the third sentence with their lowest pitch. The goal here is to explore the range of their voice and unlock their vocal variety's power. It may seem uncomfortable or unnatural, but this will help the students understand their voice's true potential.

##### **Step 3**

The students reread the speech, but they say the first sentence in a high pitch, the next in a low pitch. Then, they repeat.

It will sound unnatural, but they keep repeating in this rhythm to finish reading the whole speech. Repeat the exercise by saying a few words with a high pitch and the following few phrases with a low pitch.

##### **Step 4**

The students now read the whole speech with significant pitch variability to express the emotions. Then, the instructor encourages the students to add a sentiment to their

presentation; whether it is to inspire, convince, or inform, the students will naturally vary their pitch and sound more engaging. The instructor reminds them to add emotion naturally, not forced. The instructor encourages the students to change their high and low pitch to emphasize critical points during their presentation.

#### **Exercise 14: Pace and Pausing Exercise**

Purpose of exercise: to help improve pace and pause of speaking.

##### **Pace Variety Exercise:**

Students to read a passage in their everyday speaking style. A slash (/) occurs at 50-word intervals. They use a stopwatch or a watch with a second hand to time themselves and determine their speaking rate. Then the students read the speech in a fast speed (e.g. 150 words per minute). They take a selection from a book, magazine, or newspaper and do the same as self-practice.

Example of speech excerpt:

THE U.S. PRESIDENT'S MESSAGE ON EDUCATION TO CONGRESS, JANUARY 29, 1963

“Education is the keystone in the area of freedom and progress. Nothing has contributed more to the enlargement of this nation’s strength and opportunities than our traditional system of free, universal elementary and secondary education, coupled with widespread availability of college education. For the individual, the doors to the schoolhouse, to the library, and to the college lead to the richest treasures of our open society: to the power of knowledge – to the training and skills necessary for productive employment – to the wisdom, the ideals, and the culture which enrich life – and to the creative, self-disciplined understanding of / society needed for good citizenship in today’s changing and challenging world. For the nation, increasing the quality and availability of education is vital to both our national security and our domestic well-being. A free nation can rise no higher than the stan- dard of excellence set in its schools and colleges. / Ignorance and illiteracy, unskilled workers and school dropouts – these and other failures of our educational system breed failures in our social and economic system: delinquency, unemployment, chronic dependence, a waste of human resources, a loss of productive power and purchasing power, and an increase in tax-supported benefits.”

(By John F. Kennedy)

The students take a full two seconds to pronounce each of the following words:

*droopy roar lovely glimmer rustle lonely  
crash grumble snowflake sluggish stroll fluffy  
gloomy luscious merry passion hush happy  
sputter murmur stretch moan lover delicious*

The instructor guides a slow speaker (less than 120 words per minute) to pronounce each of the following words rapidly:

*flick glib tip snap clip nit  
flip chip snit pip blip fib  
flit dip glint bit snit click*

Then the students read the above speech aloud again by varying the speed.

##### **Pause Variety Exercise:**

The students read the following speech and pause at each dash. The students take a selection from a book, magazine, or newspaper and do the same as self-practice.

*Example of speech excerpt:*

“The whole fury and might – of the enemy – must very soon be turned on us. Hitler knows – that he will have to break us in this island – or lose the war. If we – can stand up to him – all Europe – may be free – and the life of the world may move forward – into broad sunlit uplands. – But if we fail – then the whole world – including the United States – including all that we have known and cared for – will sink into the abyss – of a new dark age – made more sinister – and perhaps more protracted – by the lights of perverted science. Let us therefore – brace ourselves to our duties – and so bear ourselves that – if the British Empire and its Common wealth last for a thousand years – men will say –This – was their finest hour.” (By Winston S. Churchill)

### **Exercise 15 Speech shadowing**

Purpose of this exercise: to help students improve natural pronunciation and intonation.

to help improve fluency.

to help produce sounds accurately.

The instructor selects a short piece of audio (max. 5 minutes) with a monologue. Students listen to the audio recording at least once to get used to the speaker’s accent, rhythm, and intonation before they start shadowing. Students closely follow the speaker’s pace and repeat what the speaker says as quickly and as accurately as possible. The instructor reminds the students not too concerned with meaning, but to mimic the speaker’s intonation and tone. The instructor encourages the students to record themselves and listen back to compare their speech to the original audio.

### **Exercise 16: Voice personas (also known as character traits)**

Purpose of exercise: to help connect voice and emotion.

to help improve speaking fluency and expressiveness.

to help connect emotion with words.

to help students connect with audience

To prepare for this activity, the instructor does the following steps:

1. Cut apart a list of character traits and emotions.
2. Select a few scripts with two to four characters.
3. Put students into groups of two or three.

During the workshop, students sit with their partners along with a script given, and each student randomly selects a character trait/emotion from character traits paper bag. Students read their script using the character trait they have selected and get into their roles. Below is the example of a list of character traits:

Adventurous Concerned Cheerful Optimistic Persuasive Warm/friendly Aggressive Annoyed Argumentative Humble Authoritative Boastful Depressed Bossy Pessimistic Awkward Shy Loving Carefree Cautionary
---

### Exercise 17: Connecting voice and body

Purpose of exercise: to help connect voice and body.

to help students understand better how speech constructs.

Students hold their speech in their hand and start reading it aloud in a simple and natural manner without thinking an interpretation of the text, or rhythm or accent. They speak the words in a natural pace and speaking volume, not fast or slowly. At the same time, they start to walk in the room. The students listen to their body and feel their body react when reading aloud. When they feel that they encounter another speech units, they stop walking and talking, turn on the spot to face another direction and then start walking and talking again from where they left off in the text. Their voice and the feet shall start at the same time, and they both stop together. They were not allowed to walk without talking or continuing to talk once their feet have stopped. They are looking for their own body's way of producing speech, using their personal knowledge and experience, not some externally imposed programme or any "right way" to create speech units. They might find that they stop and turn after one word, or after five or six. It is up to their body to decide when to turn direction.

(Adapted from Cicely Berry 1988)

### Exercise 18 Intonation and tone exercises

Purpose of exercise: to help change tone of voice to express emotion.

Practice Different Tones with Common Words, e.g., "Yeah, No, Ok, Right, Maybe":

Example: Intonation Exercise: 'Yeah'

Yeah (neutral),

Yeah. (falling)

Yeah? (rising)

Yeah! (enthusiastic)

Yeah!! (excited)

Yeah!!! (super excited)

Yeah?? (uncertain)

Yeah??? (surprise)

Yeah... (disinterested)

Intonation Exercise: Maybe

Maybe (neutral)

Maybe. (falling)

Maybe. (unconvinced)

Maybe? (rising)

Maybe?? (uncertain)

Maybe... (noncommittal)

Practice inflection with sentences:

I was born in Australia (You, on the other hand, were born somewhere else.)

I **was** born in Australia (How dare you imply that I wasn't?)

I was **born** in Australia (I'm a native – not a newcomer.)

I was born **in** Australia (Not outside Australia.)

I was born in **Australia** (Not in New Zealand.)

(adapted from Toastmasters 1996)

### **Exercise 19: Rhythm and fluency exercise**

Purpose of Exercise: to help improve fluency, intonation, stress and rhythm.

Students were given a couple of limericks to practise word stress and sentence stress. Limericks are five-line poems written with one couplet and one triplet. The last line of a good limerick contains the punch line or heart of the joke. The rhythm of limericks is essential, especially when working to reduce a foreign accent. The students were encouraged to clap the rhythm as they speak to them. When they start to focus on rhythm rather than pronunciation, the rhythm will help the students come to grips with stress, which distinguishes English from many other languages. Below is an example of limericks (the stress in the underlined part):

There once was an old man of Esser,  
Whose knowledge grew lessor and lesser,  
It, at last, grew so small  
He knew nothing at all  
And now he's a college professor.

The instructor guides the students to analyse word stress and sentence stress in a speech and practice in pairs.

### **Exercise 20 “Me”: Self-Consciousness Reduction Exercise**

Purpose of the Exercise: to help students reduce tension and excessive self-consciousness.

The students lined up to walk around the classroom. One student call out “Look at me” (or any other verse) with a clear and loud voice, and everyone else stop walking and gaze at the student for a few second and then move on until the next student call out “me”.

### **Exercise 21 Role playing: Impromptus speaking and Interpreting**

Purpose of the Exercise: to help students improvise on speech  
to help students understand structure of speech  
to help students experience roles of as speaker and interpreter

Students hold their speech in their hand and practise reading it aloud in a natural manner without thinking of an interpretation of the text, or rhythm or accent. Then, students make an impromptus speech in their A language for about two minutes. They are given 30 seconds to prepare for the speech. The speech topics are given to the students on the spot, chosen from those used in interpreting situation, for example, education, environment, and modern technology. Another student was picked up to interpret the speech either from A language to B language or from B language to A language. When the exercise is done, both students are asked to reflect on their own physical and emotional sensation and give feedback on each other's performance (role as a speaker and an interpreter). In that way, both students learnt their role as a speaker and an interpreter and their expectation of each other.

### **Exercise 22 Role playing: Prepared speech, Interpreting and Evaluating**

Purpose of the Exercise: to help students to develop a pre-performance vocal warm up routine  
to help students experience roles of as speaker and interpreter

Students are instructed to prepare a three-minute speech at their own time on certain give topics (e.g. education, medical, social welfare, etc). At the workshop, the students are given a few minutes to rehearse their prepared speeches. The instructor reminds the students a few things as follows:

1. The instructor encourages the students to move around and relax their vocal organs without inhibitions or interruptions;
2. The instructor encourages the students to apply a wide variety of vocal styles in presenting their speech and let the motivation for voice changes coming from the context of their speech;
3. The instructor reminds the students to build appropriate pauses into their presentations to give listeners a chance to mentally digest what they have said and catch up with them;
4. The instructor prompts the students to pay attention to their body alignment in standing position and not to be tense and rigid so as to give their lungs room to expand, to take moderately deep breaths as they speak and to vary their voice to match their words in the speaking context, and to speak clearly and project their voice so the entire audience can hear them.

The instructor points a student to interpret for the speaker applying vocal modulation skills appropriate to the speech and render the same impact as the speaker does to their audience. The students take turns to play a role as a speaker or an interpreter. A students speech Students Speech and Interpreting Self-Evaluation Sheet was provided to the students.

## STUDENTS SPEECH AND INTERPETING SELF- EVALUATION SHEET

	Negative	1	2	3	4	5	6	Positive
Volume	Too quiet							Good projection
	Inaudible							Easily heard
	Flat							Vibrant and dynamic
Pitch	High							Low
	Shrill (Head resonance)							Full (Chest resonance)
	Monotonous							Varied
Voice Quality	Nasal							Open
	Breathy							Clear
	Harsh, raspy							Mellow
	Lifeless							Enthusiastic
Articulation	Mumbling							Clear
	Lazy lips							Crisp
	Tangled tongue							Controlled
	Tight jaw							Open mouth
	Heavy accent							Accentless
	Mispronunciation							Correct Pronunciation
Rate and pause	Jerky							Smooth
	Slow, plodding							Fluent
	Unvaried							Varied, exciting
	Hesitant							Deliberate
Vocal Variety	Emotionless							Conveys emotion
	Unfriendly							Genial
	Strained							Natural
	Dull							Vital

### Exercise 23 Role-playing: Speaker and audience

Purpose of the Exercise: to help build rapport and friendliness with audience  
to help improve confidence

Students make a short speech. The instructor encourages the students to make eye contact with as many audiences as possible but avoids making mechanical and excessive head movement from side to side. The instructor reminds the students not to focus on the audience's face or one particular audience member as this can be embarrassing. The instructor encourages the student to adopt a pleasant facial

approach, with a smile where appropriate. The instructor also inspires the “audience” to smile back whenever they feel they receive eye contact and attention from the speaker. Students practise in turn.



## Appendix D-4 Sample of Voice Training Handout

### Mini-seminar: Spoken Voice Apparatus and Speech Organs

**Voice** is the sound produced using the **lungs** and the **vocal folds** in the **larynx**. Voice is the stepping-stone to speech. It is generated by airflow from the lungs as the vocal folds are brought close together causing them to vibrate. The air then passes into the **pharynx** and out through the mouth and nose. This process is called **phonation**.

**Speech** – the process of turning the sound produced (voice) into words and phrases by activating the **organs of articulation**.

**Breath** is the foundation of voice and speech. It is the support and the energy source. A controlled, stable breath, and an understanding of the breathing apparatus, is crucial to maintaining vocal quality and vocal strength.

#### Five basic vocal characteristics

**Pitch** – the highness or lowness of vocal tone.

**Pace** – the speed in which one speaks.

**Volume** – the loudness or softness of the voice.

**Inflection** – the rising and falling tune of the voice.

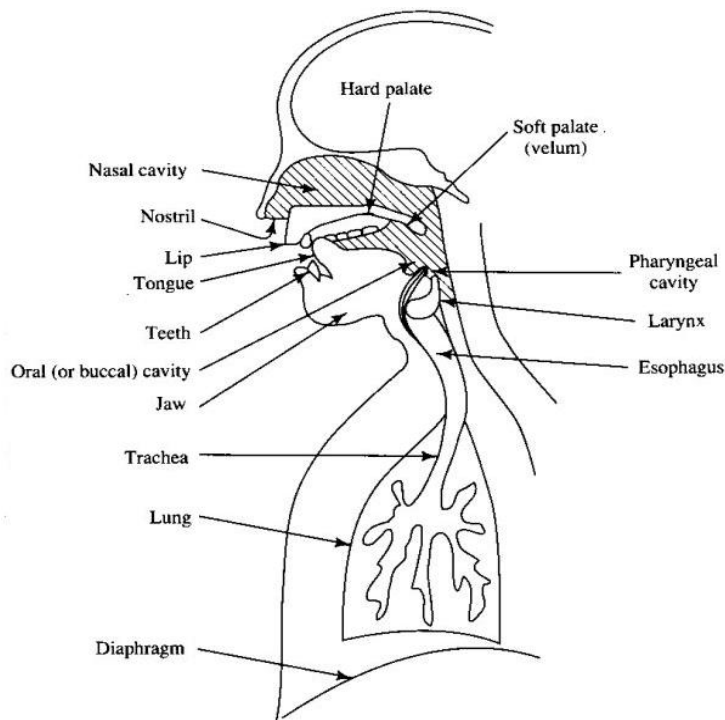
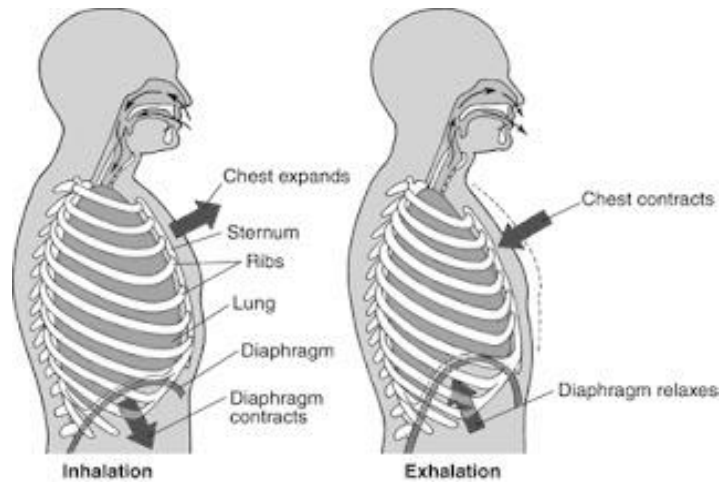
**Resonance / timbre** – the way in which basic sound finds resonance in one of the main resonator cavities - nasal cavity, oral cavity, pharyngeal cavity, chest.

A good speaking voice is dynamically balanced between pitch, pace, volume, inflection and resonance. Speaking from middle zones helps listeners to focus on the message without being distracted. Another foundation of a good speaking voice is clarity, which is achieved by precise articulation.

**Articulation** – the formation of clear and distinct sounds in speech, shaped and controlled by the **organs of articulation** – lips, cheeks, teeth, tongue, alveolar ridge (upper gum ridge), hard palate, soft palate, lower jaw. The more these muscles are exercised, the clearer the speech becomes with much less of an effort. It enables you and the audience to concentrate more on the content of what you are saying rather than how you are saying it.

**Vowels** – unimpeded voiced stream of air. It is important that all vowel sounds are open and resonant.

**Consonants** – voiced or unvoiced stream of air that is somehow interrupted, impeded or stopped. It is important that all consonant sounds are clear and crisp. Aim to eliminate the ‘splash’ – for example, many Australians say ‘chune’ and ‘chuna fish’ instead of ‘tune’ and ‘tuna fish’.



**Tension – habitual, mental, physical.**

Everyone holds tension in his or her body! These tensions evolve from the moment we are born as a result of our general anatomy or our physical and emotional experiences.

- It could be that you sustained an injury, resulting in you having to readjust your alignment, causing muscle tension in the shoulders.
- Or perhaps you grew up with someone telling you that your voice was too loud, causing you to be hyper aware of the way you sound, therefore struggling to be present and connected when speaking.
- All tensions impact the body in ways that challenge vocal delivery and the ability to communicate confidently and effectively.

Can you think of a physical tension that relates to you? Can you think of an emotional tension that relates to you?



**7) Resonance and timbre**

— +  
|.....|.....|.....|.....|.....|.....|

**8) Pace and pausing**

— +  
|.....|.....|.....|.....|.....|.....|

**9) Articulation and clarity**

(Consonants and vowels & unobstructive pronunciation and accent)

— +  
|.....|.....|.....|.....|.....|.....|

**10) Modulation of pitch range**

— +  
|.....|.....|.....|.....|.....|.....|

**11) Intonation and tone patterns**

— +  
|.....|.....|.....|.....|.....|.....|

**12) Rhythmic and fluent rendition**

— +  
|.....|.....|.....|.....|.....|.....|

**13) Expressiveness/ Emphasis**

— +  
|.....|.....|.....|.....|.....|.....|

**14) Professional confidence**

— +  
|.....|.....|.....|.....|.....|.....|





**8) Pace and pausing**

– +  
|.....|.....|.....|.....|.....|.....|

**9) Articulation and clarity**

(Consonants and vowels & unobstructive pronunciation and accent)

– +  
|.....|.....|.....|.....|.....|.....|

**10) Modulation of pitch range**

– +  
|.....|.....|.....|.....|.....|.....|

**11) Intonation and tone patterns**

– +  
|.....|.....|.....|.....|.....|.....|

**12) Rhythmic and fluent rendition**

– +  
|.....|.....|.....|.....|.....|.....|

**13) Expressiveness/ Emphasis**

– +  
|.....|.....|.....|.....|.....|.....|

**14) Professional confidence**

– +  
|.....|.....|.....|.....|.....|.....|

**15) Audience connection**

– +  
|.....|.....|.....|.....|.....|.....|

**B) What is your general impression of the second interpretation? Could you elaborate on this?**

**Excellent (90–105)**

**Good (70–90)**

**Average (60–75)**

**Below average (45–60)**

**Poor (<45)**

**C) Which performance do you think is better, video one or video two? What is the reason for your choice?**

---

---

---

**D) Do you have any specific Comment/feedback/recommendation for the second interpretation?**

---

---

---





**7) Did interpreter's resonance and timbre sound warm, mellow and bright to you?**

— +  
|.....|.....|.....|.....|.....|.....|

**8) Did the interpreter deliver with appropriate pause and comfortable speed?**

— +  
|.....|.....|.....|.....|.....|.....|

**9) Did the interpreter articulate clear and his/her accent did not prevent you from understanding the source speaker's message?**

— +  
|.....|.....|.....|.....|.....|.....|

**10) Were you comfortable with the interpreter's pitch range?**

— +  
|.....|.....|.....|.....|.....|.....|

**11) Did the interpreter use inflection/intonation to make his/her interpretation sound appealing to you?**

— +  
|.....|.....|.....|.....|.....|.....|

**12) Was the interpretation presented in a rhythmical and fluent way?**

— +  
|.....|.....|.....|.....|.....|.....|

**13) Did the interpreter sound expressive?**

— +  
|.....|.....|.....|.....|.....|.....|

**14) Did the interpreter present with professional confidence and sound convincing?**

— +  
|.....|.....|.....|.....|.....|.....|

**15) Did the interpreter interact with audience through appropriate eye contact, facial expression and modulated voice?**

— +  
|.....|.....|.....|.....|.....|.....|

**B) What is your general impression of the First interpretation? Could you elaborate on your comment?**

- Excellent (90–105)**
- Good (70–90)**
- Average (60–75)**
- Below average (45–60)**
- Poor (<45)**

**C) Do you have any comment/feedback/recommendation for the first interpretation?**

---

---

---

---

[Second Task Evaluation]

1) Do you think the interpreter presented with a good alignment and posture in his/her standing, sitting or way of holding interpreting notebook?

– +  
|.....|.....|.....|.....|.....|.....|

2) Did the interpreter manage his/her physical tension and present in a calm and ease manner?

– +  
|.....|.....|.....|.....|.....|.....|

3) Did the interpreter manage to breath smoothly while speaking?

– +  
|.....|.....|.....|.....|.....|.....|

4) Did the interpreter show discernible tension in face, lips, jaw, tongue and vocal cord while speaking?

– +  
|.....|.....|.....|.....|.....|.....|

5) Did the interpreter use appropriate movement to support his/her speaking?

– +  
|.....|.....|.....|.....|.....|.....|

6) Did the interpreter's volume sound appropriate to you?

– +  
|.....|.....|.....|.....|.....|.....|

7) Did interpreter's resonance and timbre sound warm, mellow and bright to you?

– +  
|.....|.....|.....|.....|.....|.....|

**8) Did the interpreter deliver with appropriate pause and comfortable speed?**

— +  
|.....|.....|.....|.....|.....|.....|

**9) Did the interpreter articulate clear and his/her accent did not prevent you from understanding the source speaker's message?**

— +  
|.....|.....|.....|.....|.....|.....|

**10) Were you comfortable with the interpreter's pitch range?**

— +  
|.....|.....|.....|.....|.....|.....|

**11) Did the interpreter use inflection/intonation to make his/her interpretation sound appealing to you?**

— +  
|.....|.....|.....|.....|.....|.....|

**12) Was the interpretation presented in a rhythmical and fluent way?**

— +  
|.....|.....|.....|.....|.....|.....|

**13) Did the interpreter sound expressive?**

— +  
|.....|.....|.....|.....|.....|.....|

**14) Did the interpreter present with professional confidence and sound convincing?**

— +  
|.....|.....|.....|.....|.....|.....|

**15) Did the interpreter interact with audience through appropriate eye contact, facial expression and modulated voice?**

— +  
|.....|.....|.....|.....|.....|.....|

**B) What is your general impression of the first interpretation? Could you elaborate on your comment?**

**Excellent (90–105)**     

**Good (70–90)**     

**Average (60–75)**     

**Below average (45–60)**     

**Poor (<45)**     

**C) Do you have any comment/feedback/recommendation for the second interpretation?**

---

---

---

---

---

**D) Which performance do you think is better, video one or video two? What is the reason for your choice?**

---

---

---

---

## Appendix F Samples of Interview Responses

### Part 1: Response to questions (transcripts and translation from interview audio recordings)

Initial Questions	Interviewee Response	Follow-up Questions	Interviewee Response
<p><b>1. How important to you is your voice as a means of expressing yourself, and to you as a would-be interpreter?</b> (Very important/ Important/ Not important/Not sure.) <b>Why?</b></p>	<p><i>To be honest, I never thought it before. If you ask me now, I would say voice is important. Yes, voice is of course important. Without voice, I wouldn't be able to speak, let alone be an interpreter.</i></p>	<p><b>1. How important to you is your voice as a means of expressing yourself, and to you as a would-be interpreter?</b> (Very important/ Important/ Not important/Not sure.) <b>Why?</b></p>	<p><i>Very very important. Speed, pronunciation, accent, confidence... all important for me to take and do the job in the future. To present an idea to an audience both in English and Chinese. Crucial for me to have a nice and pleasant voice while speaking.</i></p>
<p><b>2 Do you often record your interpreting and listen afterwards? If so, do you enjoy your spoken voice?</b></p>	<p><i>Yes, I recorded but I never listened in either Chinese or English. Because I don't like my voice.</i></p>	<p><b>2. Do you enjoy listening to your interpreting recording? Why or why not?</b></p>	<p><i>Still not much like. But I know to change my voice, I must listen to my recording often. I will keep record my interpreting and see if there is any way that I can improve. After that I will listen to my recording.</i></p>
<p><b>3. Which do you believe are the most important spoken voice skills within an interpreting setting? (Intonation, fluency, pitch, dynamism, voice-volume, care of voice, elocution, articulation, tone, appropriate pause and speed, connection with audience, confident, poise.) Why?</b></p>	<p><i>If I must choose one, I would say volume. Volume is a basic thing. It is like not having an interpreter at all if the lays/audience can't hear what the interpreter is saying. If an interpreter speaks too softly and the lays must prick up their ears, which is stressful and annoy. Luckily, I speak loud enough. I don't have to worry about my volume, but I do see my classmates struggling with their volume.</i></p>	<p><b>3. Which do you believe are the most important spoken voice skills within an interpreting setting? (Intonation, fluency, pitch, dynamism, voice-volume, care of voice, elocution, articulation, tone, appropriate pause and speed, connection with audience, confident, poise.) Why?</b></p>	<p><i>Well, I think control of voice and volume is important. I used to be happy with my volume, but I realize now through watching my interpreting video recording last time at your place and feedback from my classmates, I could be too loud sometimes and sound a bit aggressive to my audience. I must pay attention to that. So, I think it is the most important to control our voice and adjust our voice according to different situation.</i></p>
<p><b>4. Have you ever taken any spoken-voice related training or lesson or activities before?</b></p>	<p><i>No. Never even heard about voice training before. I never knew there was a course on voice training. But I did attend a local public speaking club when I just arrived in Australia. I went there once a week. I also got a certificate for joining the club. I learn from there how to get an idea quickly and express myself effectively and let people know what I promote. I must convince my idea within a short time and make sure people are</i></p>	<p><b>4. What have you done to your spoken voice after the first interview, such as take a spoken-voice related training or lesson or initiate a self-practise?</b></p>	<p><i>I did some practice at home. Since I talked to you last time, I understand that I can manage my voice. So, I watched YouTube about voice training and practiced alone at home. I don't have too much time because I must practise interpreting and do assignment. But I did watch YouTube and practice pronunciation sometimes. My pronunciation is not good. Got very heavy accent. I think I will practice more, especially about my pronunciation and</i></p>

	<i>interested in my idea. So, to get people interested, I have to be humorous. But I stop going because I have heavy study load and work part time. No time to go. Couldn't carry on any more.</i>		<i>accent. I often read speeches in both A and B languages and record them, just to see how I improve to read them more naturally and fluently.</i>
<b>5.How would you rate your spoken voice quality like now?</b> (Excellent/Good/Average/Poor/Not sure)	<i>Poor. I have some concern of my voice. According to the feedback from my interpreting tutor and classmates, I' am loud. When I get nervous, I would speak up and speak very fast. So, people do not understand me in that way and ask me to keep volume down and speak slowly. I also have a strong accent. And, I have high pitch when I am nervous. I thought it was all about personality: What and how I want to be and how I speak to people. I thought I need to practise on my personality and be gentler. I don't know if there is other way to help me.</i>	<b>5.How would you rate your spoken voice quality like now after the first interview?</b> (Excellent/Good/Average/Poor/Not sure)  <i>Probe: Is there any improvement since last interview? Would you like to elaborate on that?</i>	<i>I would rate my voice as fair or poor plus. I used to struggle a lot with my accent. I often struggle if I should continue my study of interpreting course. I thought maybe I do not suit to be an interpreter and should quit and study something else. I'm excited that you showed me there might be some way to change my voice and reduce accent. I think if I practice, I will improve.</i>
<b>6. Does your voice fatigue sometimes whilst interpreting practising?</b>	<i>I have a strong voice and I like speaking. It never occurred to me that voice could get tired.</i>	<b>6. Do you still experience vocal fatigue or discomfort after or during an interpreting practise? How would you look after your voice, so it would remain in good condition after a long interpreting practising?</b>	<i>Since we talked last time, I started to pay attention to my voice. I do feel sometimes my voice got cracked when I speak a lot. I think my voice is my tool for future work, so I will pay more attention to look after it.  <i>I think I need to drink more water and speak less. Of course, I need to practise a lot interpreting and I must speak. Anyway, I may go to YouTube to find ways to deal with voice fatigue.</i></i>
<b>7. Do you think that you must force the volume of your voice to be heard in interpreting practicing situation?</b>  <i>Probe: If yes, where did that happen? Could you give me an example?</i>	<i>No. I don't have this problem. On the contrary, I am thinking how I could make my voice soft. As I said, I speak too loud when I get nervous, especially when a lot of people there. I don't know how to control the volume. My tutor said I sound aggressive sometimes. That's embarrassing and frustrating. I used to be so proud of my loud voice.</i>	<b>7. How is your voice projection right now? Do you still force the volume of your voice to be heard in interpreting practicing situation?</b>  <i>Probe: If yes, where did that happen? Could you give me an example?</i>	<i>Well. Two weeks ago, we had a mock interpreting conference, I was asked to act a speaker deliver a speech for an advanced interpreting student. I decided to have a go to manage my voice volume. I found I did well with a bit effort. My tutor said I sound professional. I found if I could manage my voice, I could deliver my speech more effectively. Yes, I was wondering why I did</i>



			<i>not think like that before. Voice could be just with a bit of effort. But I mean making conscious efforts.</i>
<b>8. As an interpreting student, do you think you communicate well with your interpreting audience (your classmate or your instructor) by using spoken voice skills such using as pitch, speed, tone, pause, vocal variety, eye contact etc. (not just linguistically)?</b>	<i>I don't think I communicate well with my audience. Sometimes I was so carried away in interpreting that I forgot that I should have an eye contact with my audience. So, there is less communication there. I guess I am not confident enough because my accent in English, my high pitch, high speed. My tutor said my interpreting into English is monotonous. I don't know if I have vocal variety or not. But my tutor also said I am an energetic person. I sometimes bring energy to interpreting with my loud voice. I sometimes have a lot of fillers in my interpreting especially when I meet difficulties and I always meet difficulties in terminology and comprehension. So, I have a lot of pause and do not sound fluent enough.</i>	<b>8. Is there any improvement of you communicate with your interpreting audience (your classmate or your instructor) by purposefully using spoken voice skills such using as pitch, speed, tone, pause, vocal variety, eye contact etc. (not just linguistically)?</b>	<i>Now I try to remember to have an eye contact with my audience but sometimes it was difficult. I would forget what I should say after I look at them. It does take practice. I also purposefully control pitch and try to speak slowly and use a bit variety in tone and intonation. I also practice a lot of shadowing, speaking after the native English speaker. I feel I got much fluent after I do that. I think I need to improve eye contact and have more vocal variety.</i>
<b>9. Are you comfortable delivering interpreting to a big audience in a public place?</b>	<i>I got excited to speak to a big audience and at the same time, I got nervous. I do want to be a good public speaker. Everybody wants to be. But it takes a lot to speak well in a public place such as ideas, confidence and be able to manage the whole situation. He will constantly observe whether people follow with him or not. So as an interpreter, always be observant and know if his audience could understand his interpreting. I enjoy speaking in a public place, probably in a big conference. I think I want to work more in a conference situation. I got excited.</i>	<b>9. Do you feel that you are more comfortable delivering interpreting to a big audience in a public place now?</b>	<i>I always want to be a conference interpreter delivering to a big audience. But I don't have confidence. I know I have a lot of problem in delivery and my voice is too tough. Now I know I need to do a lot of practice to change my voice habit. Manage voice and make audience understand me and like me. Voice gives first impression. Very important for a conference interpreter. I now feel I could practice and control and voice and, so I have much more confidence to be a conference interpreter. But I found that many skills that I must build up including voice skills you mentioned last interview.</i>
<b>10. Do you often feel nervous when interpreting for your instructor and classmates or other audience? If so, how do you control your physical</b>	<i>Always. People think I am an outgoing person, but I am very easy to get nervous. When I get nervous, I would laugh and speak loudly to hide my nerve. So my interpreting</i>	<b>10. Do you still feel nervous when interpreting for your instructor and classmates or other audience? If so, how do you control your physical tension and</b>	<i>Still nervous, but much better. Now that I will tell myself that's ok. I have many skills to use, for example, to breathe deeply, think of a good interpreter as a model, etc. Tell</i>

<p><b>tension and keep interpreting in a calm manner?</b></p>	<p><i>always sound high tone and never calm. When I feel nervous, I could feel my face get very hot and I could stutter and speak very fast. I would low my head to deliver very fast and finish asap.</i></p>	<p><b>keep interpreting in a calm manner?</b></p>	<p><i>myself to calm down and start to speak softly and slow down my speed of delivery. It is effective. Indeed. It's all in my head, just a matter of control my mind, and then when I heard my voice calms down, my nerve is also calms down. Marvellous. My voice tells a lot, not only to my audience also to myself.</i></p>
<p><b>11. Do you find yourself often hold your breath while speaking or in an interpreting setting?</b></p>	<p><i>Sometimes. But when I concentrate on interpreting, I may not notice my breath too much.</i></p>	<p><b>11. Do you still hold your breath while speaking or interpreting in the classroom?</b></p>	<p><i>Much better. I found by control and release breath naturally, my nerve calms down and ease. Now I often take a deep breath if I meet difficulties in intercepting.</i></p>
<p><b>12. As an interpreting audience, have you witnessed any interpreter who appear to be impressive or disappointing in terms of their spoken voice skills? If so, what attributes did they display? (for example: confident, physical movement/poise/voice quality/volume/rate of speech/fluency/capacity to convey speaker's role/clear articulation/good communication/good speaking skills)</b></p>	<p><i>Yes, I saw an interpreter at Magistrate court who is very impressive. He strictly followed the professional ethics and performed his role. His interpreting was accurate. I did not pay too much attention about his vocal performance at that time.</i></p> <p><i>In terms of voice performance, I remember the interpreter looked very confident and spoke with good clarity.</i></p> <p><i>I would think exaggeration of idea and too much omission would be disappointing. Sometimes, I have that problem myself. (So, you don't think vocal performance such as voice projection, confidence, etc, is as important as accuracy?)</i></p> <p><i>Yes, for the time being, I know if I want to pass exam, accuracy is very important than vocal performance. But, accuracy may become less important in real life. May be after graduate from this course, I will change my opinion, at that time, vocal performance will become more important to me in real work place if you asked me by then.</i></p>	<p><b>12. Without voice training in the interpreting course, do you feel you have improved since first interview in terms of your voice projection, intonation, pitch, fluency, communication etc? Do you believe this could have been improved even more if you were provided with voice training in your poSTGraduate course? Yes/no. Why?</b></p>	<p><i>Since we met last time, I started to pay more attention to my voice projection and control of pitch. I tried to have more eye contact with my audience. I am more alert than before. I'm not sure how my tutor and classmates look at me, but I think I have sort of training, I will do better. At least, I am confident that I can improve now.</i></p>
<p><b>13. Do you think that spoken voice</b></p>	<p><i>Yes, voice training is a part of progress that I</i></p>	<p><b>13. Do you still believe/or not that spoken</b></p>	<p><i>Yes, I do. Would be great if voice training</i></p>

<p><b>training should or should not be provided to interpreting students in conjunction with interpreting training?</b></p>	<p><i>want to make in my study. Voice training is only small aspect of interpreting study, but it will contribute to our overall performance.</i></p>	<p><b>voice training should be provided to interpreting students in conjunction with interpreting training? If yes, do you think spoken voice training should be taken a selective course or integrated into a main interpreting course?</b></p>	<p><i>would be a part of the course so we could find out what needs to be improved with our voice.</i>  <i>Selective at the current stage. I think voice training is a new training area. Better to do selective course first to see how students' response. If a lot of students think voice training is important, then voice training could be put into the core subject. Otherwise, some student may get frustrated because they don't want to waste time on something that does not worth their time. Better to organize some workshops as a trial and do a bit of research on that and see students' perception on voice training and then decide.</i></p>
<p><b>14. If there is voice training provided in the interpreting course or outside of the course, are you willing to participate the training and/or to practice at your own time? Yes/no. Why?</b></p>	<p><i>I'll think about it. You know my study is very busy. So I prefer on line training rather than spend too much time attending the course. I don't know I need to think about it. It never occurred to me.</i></p>	<p><b>14. If there is voice training provided in the interpreting course or outside of the course, are you willing to participate the training and/or to practice at your own time? Yes/no. Why?</b></p>	<p>Yes. I want to participate in voice training.</p>

Part 2: View a video recording taken in the mock interpreting setting and please give a self-evaluation. How well do you think your performed? Please provide a response about your performance by circling one of the four criteria as following and provide comments.

Performance Criteria	Rating scale (Initial)	Comments (Initial)	Rating scale (Follow-up)	Comments (Follow-up)
Good posture and movement in standing, sitting or way of holding notebook?	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input checked="" type="radio"/> Poor <input type="radio"/> Very poor	Stiff and clumsy. I was not standing straight. Lurch to one side. I don't like me hear which coving one of my eyes.	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	Better. Standing straightly, but not steadily. Moved a lot back and forth.
Present with reasonable tension and generally with ease	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input checked="" type="radio"/> Poor <input type="radio"/> Very poor	Very nervous. No expression while speaking	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	Still nervous. Between Average and poor.
Appropriate volume	<input type="radio"/> Very good <input checked="" type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	Loud	<input type="radio"/> Very good <input checked="" type="radio"/> Good <input type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Smooth breath while delivery	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	I'm not sure. Looks ok.	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Not too much tension shown in your face and jaw	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average	I can see my facial muscle twitch sometimes. Too nervous.	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average	

	<input checked="" type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Poor <input type="radio"/> Very poor	
Fluency (less hesitation)	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input checked="" type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	Better fluency
Clear articulation	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input checked="" type="radio"/> Poor <input type="radio"/> Very poor	Sometimes, I spoke too fast so not clear.	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Your pronunciation and accent are easy to understand and do not hinder audience's understanding	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input checked="" type="radio"/> Poor <input type="radio"/> Very poor	I couldn't understand myself sometimes.	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	Sounds better
Pleasant voice quality	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	
Dynamic and expressive (intonation and emphasis)	<input type="radio"/> Very good <input type="radio"/> Good <input type="radio"/> Average <input checked="" type="radio"/> Poor <input type="radio"/> Very poor	I sound very flat.	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	I think if I was not nervous, I would do better.
Vocal variety	<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor		<input type="radio"/> Very good <input type="radio"/> Good <input checked="" type="radio"/> Average <input type="radio"/> Poor <input type="radio"/> Very poor	

Pitch range	<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input checked="" type="radio"/> Average</li> <li><input type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>		<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input checked="" type="radio"/> Average</li> <li><input type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>	
Appropriate delivery speed and using of pause (instead of fillers)	<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input type="radio"/> Average</li> <li><input checked="" type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>	a lot of fillers. Not aware of this.	<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input checked="" type="radio"/> Average</li> <li><input type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>	
Audience communication (eye contact, smile, facial expression, etc)	<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input type="radio"/> Average</li> <li><input type="radio"/> Poor</li> <li><input checked="" type="radio"/> Very poor</li> </ul>	No eye contact, no smile, no facial expression. Nothing.	<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input type="radio"/> Average</li> <li><input checked="" type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>	Less eye contact.
Confident and professional looking	<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input type="radio"/> Average</li> <li><input type="radio"/> Poor</li> <li><input checked="" type="radio"/> Very poor</li> </ul>		<ul style="list-style-type: none"> <li><input type="radio"/> Very good</li> <li><input type="radio"/> Good</li> <li><input checked="" type="radio"/> Average</li> <li><input type="radio"/> Poor</li> <li><input type="radio"/> Very poor</li> </ul>	Much better than last time.
For any of the above skills, are there any that you have needed to work on them? If yes, how are you going to do this?	It's eye opening to learn so much voice skills. I think I have a lot to improve specially to control my nerve and speed. And I think I also need to control my breath which might be helpful. I will voice record or video record of my practise and review later of practise these skills with my	For any of the above skills, are there any that you think you've made improvement and you think some skills that you need to continue to work on them?	I still looked very nervous in this video but seemed composure and more professional. I did not have much eye contact. I need to work on that with my classmates. I spoke much clearer than before and fluent. I need to practice more about pitch and vocal variety to make my interpreting sound dynamic and interesting. I will also practise reading aloud and how to break down sentences and syntax, etc.	

	friends. Most importantly, I will self-monitor from my daily practice and get more feedback from my friends.		
What else would you like to add/comment on?	No. That's it.	What will you seek for further improvement regarding better vocal performance in interpreting work or any other comments?	Yes. I think spoken voice is extremely important. It shows our confidence and it is the first impression of an interpreter to their audience. I will work on that continually now and future. I think both language proficiency and voice training are important. Just like a scale having two sides. Both gives people confidence. For me, confidence and fluency are what I aim to get a ' <i>delivery confidence</i> '

## Appendix G Samples of Reflective Journals

I really appreciate the program of voice training.

The recent voice training class impresses me a lot. In the introduction section of the function of voice, the tutor said that as far as voice training is concerned, there was no difference whether it focuses on production in English or in Chinese – the same principles and applications apply. All voice can present the speaker's emotions and conditions. For me, I was very nervous when I was asked to introduce myself in English, but that does not mean that I would be nervous when I use Chinese to introduce myself. I find that I cannot perform in English as well as in Chinese because I am not satisfied with my English pronunciation. There are many Chinese elements in my voice. I cannot speak as a real English person. This made me feel less confident when I speak English. So, the self-diagnosis workshop was very useful. I got to reflect on the reason for my lack of confidence.

In one workshop of the voice training, I was video recorded. I had prepared two speeches. One was in Chinese and another was in English. What surprised me most was that I needed to interpret for somebody once and gave the speech once. When it was my turn, I was very nervous. I need to interpret from Chinese to English first. When I heard the topic of my partner's speech, I felt that I cannot do it well because I had no knowledge of that. In addition, I forgot to bring a pen and a paper to take notes of the long list of the names of the tax laws. I could say that I did not perform well as an interpreter. And when I delivered a speech in Chinese, I felt more confident because I could remember all the things and showed them out. Later, I was given the video recording and asked to reflect on my practice. I was very shocked at my poor performance.

I found that during this workshop, there were several elements that affected my voice. My pitch was low when I interpreted because I was nervous and unconfident. I could not speak louder, and the voice was blocked in my mouth. I had to think of the contents, take notes and deliver them, which distracted me, and I could not pay much attention of my voice. But when I gave a speech, I felt much freer to speak and this was indicated in my voice.

I have learnt a lot from other the voice training workshop and from other students.

The first and the most importance thing is how to make a good voice. We learnt that we can make sound from our head, throat and chest. The three parts of our body can create difference voices. I find that it is most easy for me to make sound from throat. I seldom use head and chest to make sound.

We did many warming-up activities to activate our voice before we practice interpreting. The following activities that I found most useful to me: First is to train the flexibility of our tongue and mouth. Second, we need to release ourselves so as to produce sound easier. We also did some exercises on our neck, back, shoulders, waist and so on, which can improve our health condition and help us to make a natural sound. Thirdly, we were told that we need to practice every day and this will allow our diaphragms to develop well in the future.

Another thing that I have learnt from this class is about stand gesture. People stand differently, and the stance can imply the emotion of the speaker as well. It is essential to have a good stance as an interpreter, because we need to stand for a long time in some situations. It is better to have some distances between the two feet. Standing straightly is good but it will increase the pressure on the knees. We need to find a proper stand gesture of our own.

What's more, eye contact and hand gestures are two more aspects that we need to pay attention to. Both interpreters and speakers need to interact with the audience, and eye contact can help that. If



the speaker or interpreter do not have eye contact with the audience, then they will lose the audience. Hand gesture is good sometimes to deliver the speaker's emotion and help the audience understand the topic. However, improper hand gesture will mislead the audience and make them distracted. As an interpreter, we will have a notebook and a pen in our hands all the time, so there is no such problem, I think.

In conclusion, voice plays a very important role in especially an interpreter's life. I need to learn how to take control of my voice. Such exercises that we have done in the class everyday help me a lot. In fact, I have kept exercising my mouth every day, such as speaking vowels and moving my tongue as quickly as possible. And I have tried to imitate the sound from the recordings. The pronunciation and the pitch I can imitate well, but I still have problem of following the tone of the recordings.

I think that I have improved a lot after the training. Firstly, I feel more confident in speaking English and doing Chinese to English direction interpretation. Voice training gave me a lot of confidence. Voice helps me to express myself, so I can be confident in speaking English as if I speak Chinese. And I believe, with more exercises, I will do even better. I no longer feel fearful when I interpret from Chinese and English. I know it is not because my English level is bad, all is about my confidence. Voice training in English helps me find my weak point and get rid of fear.

Second, I learnt to project my voice better although I still have a flat tone and low pitch. I learnt to make sound from my chest rather than from my throat only. Through many activities that we took in the workshop and exercise done at home, I find my voice projection is easier than before. I also practice the flexibility of the tongue and mouth and do exercises on neck, back, shoulders, waist often to make a natural sound. I feel great after these exercises and I feel energetic and ease.

We recently had another interpreting exercise; the tutor and students gave me much positive feedback. They said that I had a better posture and eye contact with the audience. I looked confident and professional. No one cares about my interpreting skills! I secretly think that I missed many interpreting points. But everyone seemed enjoying my interpreting and my fake confidence. I will keep practice my interpreting skills so that I will be really confidence and a good interpreter with good voice.

## Appendix H Reliability Statistics

<b>Cronbach's alpha</b>	<b>N of items</b>
0.826	34

## Appendix I-1 Test of Within-Subject Effects for Four Groups (Voice Training Assessor)

**Paired samples statistics**

Group			Mean	N	Std. deviation	Std. error mean
Training group (STG group)	Pair 1	pretotal	58.4000	5	9.14740	4.09084
		posttotal	81.6000	5	12.66590	5.66436
Training group (LTG group)	Pair 1	pretotal	67.1000	5	12.81796	5.73236
		posttotal	83.6000	5	13.03553	5.82967
ARG group	Pair 1	pretotal	65.4000	5	14.46720	6.46993
		posttotal	68.7000	5	19.02170	8.50676
Control group	Pair 1	pretotal	69.3000	5	13.05565	5.83866
		posttotal	72.5000	5	12.89864	5.76845

**Paired samples test**

Group			Paired differences					t	df	Sig. (2-tailed)
			Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
						Lower	Upper			
Training group (STG group)	Pair 1	pretotal - posttotal	-23.20000	15.40536	6.88948	-42.32828	-4.07172	-3.367	4	<b>.028</b>
Training group (LTG group)	Pair 1	pretotal - posttotal	-16.50000	12.86954	5.75543	-32.47964	-.52036	-2.867	4	<b>.046</b>
ARG group	Pair 1	pretotal - posttotal	-3.30000	10.34771	4.62763	-16.14837	9.54837	-.713	4	.515
Control group	Pair 1	pretotal - posttotal	-3.20000	12.00833	5.37029	-18.11031	11.71031	-.596	4	.583

## Appendix I-2 Test of Within-Subject Effects for Four Groups (Lay Assessor)

**Paired samples statistics**

Group			Mean	N	Std. deviation	Std. error mean
Training group (STG group)	Pair 1	posttotal	88.3000	5	11.89853	5.32118
		pretotal	68.2000	5	6.73238	3.01081
Training group (LTG group)	Pair 1	posttotal	77.9000	5	20.25586	9.05870
		pretotal	74.8000	5	27.21580	12.17128
ARG group	Pair 1	posttotal	81.2000	5	16.57408	7.41215
		pretotal	59.2000	5	9.57601	4.28252
Control group	Pair 1	posttotal	71.0000	5	22.71563	10.15874
		pretotal	82.0000	5	11.76860	5.26308

**Paired samples test**

group			Paired differences					t	df	Sig. (2-tailed)
			Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
						Lower	Upper			
Training group (STG group)	Pair 1	posttotal - pretotal	20.10000	10.10198	4.51774	7.55674	32.64326	4.449	4	<b>.011</b>
Training group (LTG group)	Pair 1	posttotal - pretotal	3.10000	18.46754	8.25893	19.83048	26.03048	.375	4	.726
ARG group	Pair 1	posttotal - pretotal	22.00000	20.62765	9.22497	-3.61261	47.61261	2.385	4	.076
Control group	Pair 1	posttotal - pretotal	11.00000	23.80126	10.64425	40.55317	18.55317	-1.033	4	.360

## Appendix I-3 Test of Within-Subject Effects for Four Groups (Interpreting Instructor Assessor)

**Paired samples statistics**

Group			Mean	N	Std. deviation	Std. error mean
Training group (STG group)	Pair 1	posttotal	86.2000	5	10.56882	4.72652
		pretotal	60.2000	5	15.57883	6.96707
Training group (LTG group)	Pair 1	posttotal	69.2000	5	17.83816	7.97747
		pretotal	68.8000	5	16.82855	7.52596
ARG group	Pair 1	posttotal	71.6000	5	21.05469	9.41594
		pretotal	58.4000	5	10.71448	4.79166
Control group	Pair 1	posttotal	58.4000	5	16.81666	7.52064
		pretotal	59.6000	5	12.54193	5.60892

**Paired samples test**

Group			Paired differences					t	df	Sig. (2-tailed)
			Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
						Lower	Upper			
Training group (STG group)	Pair 1	posttotal - pretotal	26.00000	10.46422	4.67974	13.00695	38.99305	5.556	4	<b>.005</b>
Training group (LTG group)	Pair 1	posttotal - pretotal	.40000	12.03329	5.38145	-14.54130	15.34130	.074	4	.944
ARG group	Pair 1	posttotal - pretotal	13.20000	14.54991	6.50692	-4.86610	31.26610	2.029	4	.112
Control group	Pair 1	posttotal - pretotal	-1.20000	8.40833	3.76032	-11.64032	9.24032	-.319	4	.766

## Appendix J-1 Test of Within-Subject Effects for Three Categories of Four Groups (Voice Trainer Assessors)

### STG voice training group

**Paired samples statistics<sup>a</sup>**

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	14.3000	5	3.80132	1.70000
	post_cat1	19.9000	5	3.89551	1.74213
Pair 2	pre_cat2	20.2000	5	4.43847	1.98494
	post_cat2	25.8000	5	3.70135	1.65529
Pair 3	pre_cat3	21.0000	5	4.94975	2.21359
	post_cat3	32.2000	5	5.71839	2.55734

a. group = Training group (STG group)

**Paired samples test<sup>a</sup>**

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-5.60000	3.83080	1.71318	-10.35656	-0.84344	-3.269	4	0.031
Pair 2	pre_cat2 - post_cat2	-5.60000	4.66905	2.08806	-11.39739	0.19739	-2.682	4	0.055
Pair 3	pre_cat3 - post_cat3	11.20000	9.60208	4.29418	-23.12256	0.72256	-2.608	4	0.060

a. group = Training group (STG group)

## LTG voice training group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	16.3000	5	2.28035	1.01980
	post_cat1	21.3000	5	2.07966	0.93005
Pair 2	pre_cat2	22.4000	5	6.02495	2.69444
	post_cat2	28.6000	5	3.71484	1.66132
Pair 3	pre_cat3	25.4000	5	6.73053	3.00998
	post_cat3	30.0000	5	7.64853	3.42053

a. group = Training group (LTG)

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-5.00000	2.78388	1.24499	-8.45665	-1.54335	-4.016	4	0.016
Pair 2	pre_cat2 - post_cat2	-6.20000	6.14003	2.74591	-13.82386	1.42386	-2.258	4	0.087
Pair 3	pre_cat3 - post_cat3	-4.60000	8.50294	3.80263	-15.15780	5.95780	-1.210	4	0.293

a. group = Training group (LTG group)

## ARG group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	19.0000	5	2.91548	1.30384
	post_cat1	19.2000	5	4.68508	2.09523
Pair 2	pre_cat2	21.0000	5	6.12372	2.73861
	post_cat2	22.4000	5	6.84105	3.05941
Pair 3	pre_cat3	22.2000	5	7.85493	3.51283
	post_cat3	23.6000	5	10.31019	4.61086

a. group = ARG

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-0.20000	4.73814	2.11896	-6.08318	5.68318	-0.094	4	0.929
Pair 2	pre_cat2 - post_cat2	-1.40000	4.50555	2.01494	-6.99438	4.19438	-0.695	4	0.525
Pair 3	pre_cat3 - post_cat3	-1.40000	6.58027	2.94279	-9.57049	6.77049	-0.476	4	0.659

a. group = Training group (ARG group)



## Control group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	17.0000	5	4.41588	1.97484
	post_cat1	20.1000	5	5.90974	2.64292
Pair 2	pre_cat2	20.4000	5	2.30217	1.02956
	post_cat2	21.0000	5	4.69042	2.09762
Pair 3	pre_cat3	29.0000	5	7.44983	3.33167
	post_cat3	27.6000	5	4.72229	2.11187

a. group = Control group

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-3.10000	6.45562	2.88704	-11.11571	4.91571	-1.074	4	0.343
Pair 2	pre_cat2 - post_cat2	-0.60000	4.61519	2.06398	-6.33052	5.13052	-0.291	4	0.786
Pair 3	pre_cat3 - post_cat3	1.40000	4.09878	1.83303	-3.68931	6.48931	0.764	4	0.488

a. group = Training group (Control group)

## Appendix J-2 Test of Within-Subject Effects for Three Categories of Four Groups (Lay Assessors)

### STG voice training group

**Paired samples statistics<sup>a</sup>**

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	18.2000	5	2.48998	1.11355
	post_cat1	22.8000	5	3.89872	1.74356
Pair 2	pre_cat2	21.8000	5	2.58844	1.15758
	post_cat2	28.4000	5	3.28634	1.46969
Pair 3	pre_cat3	24.8000	5	2.68328	1.20000
	post_cat3	33.0000	5	4.58258	2.04939

a. group = Training group (STG group)

**Paired samples test<sup>a</sup>**

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-4.60000	3.84708	1.72047	-9.37678	0.17678	-2.674	4	0.056
Pair 2	pre_cat2 - post_cat2	-6.60000	2.70185	1.20830	-9.95479	-3.24521	-5.462	4	<b>0.005</b>
Pair 3	pre_cat3 - post_cat3	-8.20000	4.60435	2.05913	13.91705	-2.48295	-3.982	4	<b>0.016</b>

a. group = Training group (STG group)

## LTG voice training group

**Paired samples statistics<sup>a</sup>**

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	19.2000	5	6.72309	3.00666
	post_cat1	19.8000	5	4.91935	2.20000
Pair 2	pre_cat2	23.4000	5	8.73499	3.90640
	post_cat2	24.4000	5	6.76757	3.02655
Pair 3	pre_cat3	29.0000	5	10.79352	4.82701
	post_cat3	29.6000	5	7.92465	3.54401

a. group = Training group (LTG)

**Paired samples test<sup>a</sup>**

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-0.60000	4.97996	2.22711	-6.78344	5.58344	-0.269	4	0.801
Pair 2	pre_cat2 - post_cat2	-1.00000	5.14782	2.30217	-7.39186	5.39186	-0.434	4	0.686
Pair 3	pre_cat3 - post_cat3	-0.60000	7.09225	3.17175	-9.40619	8.20619	-0.189	4	0.859

a. group = Training group (LTG group)

## ARG group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	16.0000	5	2.54951	1.14018
	post_cat1	20.8000	5	3.76829	1.68523
Pair 2	pre_cat2	18.8000	5	2.77489	1.24097
	post_cat2	26.2000	5	5.31037	2.37487
Pair 3	pre_cat3	21.4000	5	4.92950	2.20454
	post_cat3	30.2000	5	6.76018	3.02324

a. group = ARG

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-4.80000	4.96991	2.22261	-10.97096	1.37096	-2.160	4	0.097
Pair 2	pre_cat2 - post_cat2	-7.40000	5.94138	2.65707	-14.77720	-0.02280	-2.785	4	<b>0.050</b>
Pair 3	pre_cat3 - post_cat3	-8.80000	9.78264	4.37493	-20.94675	3.34675	-2.011	4	0.115

a. group = Training group (ARG)

## Control group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	19.6000	5	3.20936	1.43527
	post_cat1	17.8000	5	5.97495	2.67208
Pair 2	pre_cat2	27.0000	5	3.00000	1.34164
	post_cat2	24.8000	5	5.97495	2.67208
Pair 3	pre_cat3	31.2000	5	4.76445	2.13073
	post_cat3	25.4000	5	9.71082	4.34281

a. group = Control group

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	1.80000	5.76194	2.57682	-5.35440	8.95440	0.699	4	0.523
Pair 2	pre_cat2 - post_cat2	2.20000	6.53452	2.92233	-5.91368	10.31368	0.753	4	0.493
Pair 3	pre_cat3 - post_cat3	5.80000	10.18332	4.55412	-6.84426	18.44426	1.274	4	0.272

a. group = Training group (Control group)

## Appendix J-3 Test of Within-Subject Effects for Three Categories of Four Groups (Interpreting Instructor Assessor)

### STG voice training group

**Paired samples statistics<sup>a</sup>**

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	14.8000	5	3.03315	1.35647
	post_cat1	22.0000	5	2.00000	0.89443
Pair 2	pre_cat2	20.6000	5	4.82701	2.15870
	post_cat2	28.2000	5	4.02492	1.80000
Pair 3	pre_cat3	21.8000	5	7.15542	3.20000
	post_cat3	31.8000	5	4.14729	1.85472

a. group = Training group (STG group)

**Paired samples test<sup>a</sup>**

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-7.20000	2.28035	1.01980	-10.03143	-4.36857	-7.060	4	0.002
Pair 2	pre_cat2 - post_cat2	-7.60000	2.96648	1.32665	-11.28337	-3.91663	-5.729	4	0.005
Pair 3	pre_cat3 - post_cat3	10.00000	4.69042	2.09762	-15.82392	-4.17608	-4.767	4	0.009

a. group = Training group (STG group)

## LTG voice training group

**Paired samples statistics<sup>a</sup>**

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	18.4000	5	3.64692	1.63095
	post_cat1	21.6000	5	4.72229	2.11187
Pair 2	pre_cat2	21.8000	5	5.63028	2.51794
	post_cat2	23.2000	5	6.37966	2.85307
Pair 3	pre_cat3	25.4000	5	7.89303	3.52987
	post_cat3	21.2000	5	7.29383	3.26190

a. group = Training group (LTG)

**Paired samples test<sup>a</sup>**

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-3.20000	1.92354	0.86023	-5.58839	-0.81161	-3.720	4	<b>0.020</b>
Pair 2	pre_cat2 - post_cat2	-1.40000	3.57771	1.60000	-5.84231	3.04231	-0.875	4	0.431
Pair 3	pre_cat3 - post_cat3	4.20000	7.46324	3.33766	-5.06684	13.46684	1.258	4	0.277

a. group = Training group (LTG)

## ARG group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	15.2000	5	3.70135	1.65529
	post_cat1	18.0000	5	4.74342	2.12132
Pair 2	pre_cat2	19.4000	5	3.28634	1.46969
	post_cat2	23.0000	5	7.87401	3.52136
Pair 3	pre_cat3	21.2000	5	4.02492	1.80000
	post_cat3	26.6000	5	9.01665	4.03237

a. group = ARG

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-2.80000	4.54973	2.03470	-8.44923	2.84923	-1.376	4	0.241
Pair 2	pre_cat2 - post_cat2	-3.60000	5.50454	2.46171	-10.43479	3.23479	-1.462	4	0.217
Pair 3	pre_cat3 - post_cat3	-5.40000	7.30068	3.26497	-14.46500	3.66500	-1.654	4	0.173

a. group = Training group (ARG)



## CONTROL group

### Paired samples statistics<sup>a</sup>

		Mean	N	Std. deviation	Std. error mean
Pair 1	pre_cat1	14.6000	5	3.28634	1.46969
	post_cat1	15.6000	5	3.57771	1.60000
Pair 2	pre_cat2	19.6000	5	3.84708	1.72047
	post_cat2	19.6000	5	5.68331	2.54165
Pair 3	pre_cat3	22.4000	5	5.68331	2.54165
	post_cat3	20.6000	5	7.63544	3.41467

a. group = Control group

### Paired samples test<sup>a</sup>

		Paired differences					t	df	Sig. (2-tailed)
		Mean	Std. deviation	Std. error mean	95% confidence interval of the difference				
					Lower	Upper			
Pair 1	pre_cat1 - post_cat1	-1.00000	2.12132	0.94868	-3.63397	1.63397	-1.054	4	0.351
Pair 2	pre_cat2 - post_cat2	0.00000	3.67423	1.64317	-4.56216	4.56216	0.000	4	1.000
Pair 3	pre_cat3 - post_cat3	1.80000	4.71169	2.10713	-4.05033	7.65033	0.854	4	0.441

a. group = Training group (Control group)

## Appendix K-1 Between-Group Test Results for Four Groups from Three Assessors

### One-way ANCOVA test and pairwise comparisons of four groups from voice training assessor

#### Tests of between-subject effects

Dependent variable: posttotal

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	1868.983 <sup>a</sup>	4	467.246	3.006	0.052	0.445
Intercept	646.276	1	646.276	4.158	0.059	0.217
Pretotal	1102.883	1	1102.883	7.096	0.018	0.321
Group	1110.554	3	370.185	2.382	0.110	0.323
Error	2331.317	15	155.421			
Total	121551.500	20				
Corrected total	4200.300	19				

a. R squared = 0.445 (adjusted R squared = 0.297)

#### Estimated marginal means

Dependent variable	posttotal			
	Group	Mean	Std. error	95% confidence interval
				Lower bound
Training group (SEU group)	86.007a	5.816	73.612	98.403
Training group (Monash Group)	82.241a	5.599	70.308	94.175
Interview group	68.468a	5.576	56.583	80.353
Control group	69.683a	5.675	57.588	81.779

a. Covariates appearing in the model are evaluated at the following values: pretotal = 65.0500.

## Pairwise comparisons

Dependent variable: posttotal

(I) Group		Mean difference (I-J)	Std. error	Sig. <sup>b</sup>	95% confidence interval for difference <sup>b</sup>	
					Lower bound	Upper bound
<b>Training group (STG)</b>	Training group (LTG)	<b>3.766</b>	8.176	<b>0.652</b>	-13.662	21.194
	ARG	<b>17.539*</b>	8.075	<b>0.046</b>	0.328	34.750
	Control group	<b>16.324</b>	8.338	<b>0.069</b>	-1.448	34.096
<b>Training group (LTG)</b>	Training group (STG group)	<b>-3.766</b>	8.176	<b>0.652</b>	-21.194	13.662
	ARG	<b>13.773</b>	7.896	<b>0.102</b>	-3.057	30.603
	Control group	<b>12.558</b>	7.904	<b>0.133</b>	-4.288	29.404
<b>ARG group</b>	Training group (STG group)	<b>-17.539*</b>	8.075	<b>0.046</b>	-34.750	-0.328
	Training group (LTG)	<b>-13.773</b>	7.896	<b>0.102</b>	-30.603	3.057
	Control group	<b>-1.215</b>	7.944	<b>0.880</b>	-18.148	15.717
<b>Control group</b>	Training group (STG group)	<b>-16.324</b>	8.338	<b>0.069</b>	-34.096	1.448
	Training group (LTG)	<b>-12.558</b>	7.904	<b>0.133</b>	-29.404	4.288
	ARG	<b>1.215</b>	7.944	<b>0.880</b>	-15.717	18.148

Based on estimated marginal means

\*a. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

## One-way ANCOVA test and pairwise comparisons of four groups from lay assessor

### Tests of between-subject effects

Dependent variable: posttotal

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	1615.099 <sup>a</sup>	4	403.775	1.337	0.302	0.263
Intercept	1732.508	1	1732.508	5.736	0.030	0.277
Pre-total	839.599	1	839.599	2.780	0.116	0.156
Group	1383.668	3	461.223	1.527	0.248	0.234
Error	4530.701	15	302.047			
Total	132869.000	20				
Corrected total	6145.800	19				

a. R squared = 0.263 (adjusted R squared = 0.066)

### Estimated marginal means

Dependent variable	posttotal			
	Mean	Std. error	95% confidence interval	
			Lower bound	Upper bound
group				
Training group (SEU group)	89.595a	7.811	72.946	106.244
Training group (Monash group)	76.196a	7.839	59.487	92.905
Interview group	86.586a	8.417	68.645	104.526
Control group	66.023a	8.326	48.277	83.769

a. Covariates appearing in the model are evaluated at the following values: pretotal = 71.0500.

## Pairwise comparisons

**Dependent variable: posttotal**

(I) Group	Mean difference (I-J)		Std. error	Sig. <sup>a</sup>	95% confidence interval for difference <sup>a</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	13.400	11.138	0.248	-10.341	37.140
	ARG	3.010	11.262	0.793	-20.995	27.015
	Control group	23.572	11.618	0.061	-1.191	48.334
Training group (LTG)	Training group (STG group)	-13.400	11.138	0.248	-37.140	10.341
	ARG	-10.390	11.786	0.392	-35.510	14.731
	Control group	10.172	11.166	0.377	-13.627	33.971
ARG	Training group (STG group)	-3.010	11.262	0.793	-27.015	20.995
	Training group (LTG)	10.390	11.786	0.392	-14.731	35.510
	Control group	20.562	12.627	0.124	-6.352	47.476
Control group	Training group (STG group)	-23.572	11.618	0.061	-48.334	1.191
	Training group (LTG)	-10.172	11.166	0.377	-33.971	13.627
	ARG	-20.562	12.627	0.124	-47.476	6.352

Based on estimated marginal means

a. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of four groups from interpreting instructor assessor**

**Tests of between-subject effects**

**Dependent variable: posttotal**

Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	4481.767 <sup>a</sup>	4	1120.442	7.977	0.001	0.680
Intercept	219.122	1	219.122	1.560	0.231	0.094
pretotal	2517.217	1	2517.217	17.922	0.001	0.544
group	2295.471	3	765.157	5.448	0.010	0.521
Error	2106.783	15	140.452			
Total	108405.000	20				
Corrected total	6588.550	19				

a. R squared = 0.680 (adjusted R squared = 0.595)

**Estimated marginal means**

Dependent variable	posttotal				
	Group	Mean	Std. error	95% confidence interval	
				Lower bound	Upper bound
Training group (SEU group)	87.576a	5.310	76.258	98.894	
Training group (Monash group)	62.939a	5.502	51.211	74.668	
Interview group	74.575a	5.346	63.179	85.971	
Control group	60.309a	5.319	48.972	71.647	

a. Covariates appearing in the model are evaluated at the following values: pretotal = 61.7500.

## Pairwise comparisons

**Dependent variable: posttotal**

(I) group		Mean difference (I-J)	Std. error	Sig. <sup>b</sup>	95% confidence interval for difference <sup>b</sup>	
					Lower bound	Upper bound
<b>Training group (STG group)</b>	Training group (LTG)	<b>24.637*</b>	7.709	<b>0.006</b>	8.205	41.069
	ARG	13.002	7.505	0.104	-2.995	28.998
	Control group	<b>27.267*</b>	7.496	<b>0.002</b>	11.289	43.245
<b>Training group (LTG)</b>	Training group (STG group)	<b>-24.637*</b>	7.709	<b>0.006</b>	-41.069	-8.205
	ARG	-11.636	7.806	0.157	-28.275	5.003
	Control group	2.630	7.740	0.739	-13.867	19.127
<b>ARG</b>	Training group (STG group)	-13.002	7.505	0.104	-28.998	2.995
	Training group (LTG)	11.636	7.806	0.157	-5.003	28.275
	Control group	14.266	7.500	0.077	-1.719	30.251
<b>Control group</b>	Training group (STG group)	<b>-27.267*</b>	7.496	<b>0.002</b>	-43.245	-11.289
	Training group (LTG)	-2.630	7.740	0.739	-19.127	13.867
	ARG	-14.266	7.500	0.077	-30.251	1.719

Based on estimated marginal means

\*. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

## Appendix K-2 Between-Group Test Results of Three Categories

One-way ANCOVA test and pairwise comparisons of “spoken voice production” (cat1) from the voice training assessor

### Tests of between-subjects effects

Dependent variable:	post_cat1					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	40.910 <sup>a</sup>	4	10.228	0.556	0.698	0.129
Intercept	122.537	1	122.537	6.659	0.021	0.307
pre_cat1	29.473	1	29.473	1.602	0.225	0.096
group	24.005	3	8.002	0.435	0.731	0.080
Error	276.027	15	18.402			
Total	8417.250	20				
Corrected total	316.938	19				

a.R squared = 0.129 (adjusted R squared = -0.103)

### Pairwise comparisons

Dependent variable:		post_cat1				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>a</sup>	95% confidence interval for difference <sup>a</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	-0.614	2.783	0.829	-6.546	5.319
	ARG	2.548	3.081	0.421	-4.019	9.115
	Control group	0.862	2.840	0.766	-5.191	6.915
Training group (LTG)	Training group (STG group)	0.614	2.783	0.829	-5.319	6.546
	ARG	3.162	2.840	0.283	-2.891	9.215
	Control group	1.475	2.722	0.596	-4.326	7.277



ARG	Training group (STG group)	-2.548	3.081	0.421	-9.115	4.019
	Training group (LTG)	-3.162	2.840	0.283	-9.215	2.891
	Control group	-1.686	2.783	0.554	-7.619	4.246
Control group	Training group (STG group)	-0.862	2.840	0.766	-6.915	5.191
	Training group (LTG)	-1.475	2.722	0.596	-7.277	4.326
	ARG	1.686	2.783	0.554	-4.246	7.619

Based on estimated marginal means

a. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments)

**One-way ANCOVA test and pairwise comparisons of “spoken voice quality” (cat2) from the voice training assessor**

**Tests of between-subjects effects**

Dependent variable:	post_cat2					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	265.183 <sup>a</sup>	4	66.296	3.362	0.037	0.473
Intercept	179.344	1	179.344	9.096	0.009	0.377
pre_cat2	89.433	1	89.433	4.536	0.050	0.232
group	144.978	3	48.326	2.451	0.104	0.329
Error	295.767	15	19.718			
Total	12517.000	20				
Corrected total	560.950	19				

a. R squared = 0.473 (adjusted R squared = 0.332)

**Pairwise comparisons**

Dependent variable:		post_cat2				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>b</sup>	95% confidence interval for difference <sup>b</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	-1.753	2.851	0.548	-7.830	4.323
	ARG	3.781	2.814	0.199	-2.218	9.779
	Control group	4.895	2.809	0.102	-1.092	10.882
Training group (LTG)	Training group (STG group)	1.753	2.851	0.548	-4.323	7.830
	ARG	5.534	2.826	0.069	-0.489	11.557
	Control group	6.649*	2.844	<b>0.034</b>	0.587	12.710
ARG	Training group (STG group)	-3.781	2.814	0.199	-9.779	2.218

	Training group (LTG)	-5.534	2.826	0.069	-11.557	0.489
	Control group	1.115	2.812	0.697	-4.878	7.107
Control group	Training group (STG group)	-4.895	2.809	0.102	-10.882	1.092
	Training group (LTG)	-6.649*	2.844	<b>0.034</b>	-12.710	-0.587
	ARG	-1.115	2.812	0.697	-7.107	4.878

Based on estimated marginal means

\*. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of “spoken voice dynamism” (cat3)  
from the voice training assessor**

**Tests of between-subjects effects**

Dependent variable:	post_cat3					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	380.873 <sup>a</sup>	4	95.218	2.036	0.141	0.352
Intercept	320.397	1	320.397	6.849	0.019	0.313
pre_cat3	177.523	1	177.523	3.795	0.070	0.202
Group	256.404	3	85.468	1.827	0.186	0.268
Error	701.677	15	46.778			
Total	17157.000	20				
Corrected total	1082.550	19				

a. R squared = 0.352 (adjusted R squared = 0.179)

**Pairwise comparisons**

Dependent variable:		post_cat3				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>a</sup>	95% confidence interval for difference <sup>a</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	4.344	4.463	0.346	-5.170	13.857
	ARG	9.185	4.336	0.051	-0.057	18.427
	Control group	8.497	4.766	0.095	-1.661	18.656
Training group (LTG)	Training group (STG group)	-4.344	4.463	0.346	-13.857	5.170
	ARG	4.841	4.399	0.288	-4.535	14.217
	Control group	4.154	4.418	0.362	-5.264	13.571
ARG	Training group (STG group)	-9.185	4.336	0.051	-18.427	0.057
	Training group (LTG)	-4.841	4.399	0.288	-14.217	4.535

	Control group	-0.687	4.648	0.884	-10.594	9.220
Control group	Training group (STG group)	-8.497	4.766	0.095	-18.656	1.661
	Training group (LTG)	-4.154	4.418	0.362	-13.571	5.264
	ARG	0.687	4.648	0.884	-9.220	10.594

Based on estimated marginal means

a. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of “spoken voice production” (cat1) from the lay assessor**

**Tests of between-subjects effects**

Dependent variable:	post_cat1					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	117.786 <sup>a</sup>	4	29.446	1.451	0.266	0.279
Intercept	118.503	1	118.503	5.839	0.029	0.280
pre_cat1	52.786	1	52.786	2.601	0.128	0.148
Group	89.140	3	29.713	1.464	0.264	0.227
Error	304.414	15	20.294			
Total	8664.000	20				
Corrected total	422.200	19				

a. R squared = 0.279 (adjusted R squared = 0.087)

**Pairwise comparisons**

Dependent variable:		post_cat1				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>a</sup>	95% confidence interval for difference <sup>a</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	3.440	2.862	0.248	-2.661	9.540
	ARG	1.032	2.912	0.728	-5.174	7.238
	Control group	5.616	2.875	0.070	-0.511	11.743
Training group (LTG)	Training group (STG group)	-3.440	2.862	0.248	-9.540	2.661
	ARG	-2.408	2.980	0.432	-8.759	3.944
	Control group	2.176	2.851	0.457	-3.901	8.253
ARG	Training group (STG group)	-1.032	2.912	0.728	-7.238	5.174

	Training group (LTG)	2.408	2.980	0.432	-3.944	8.759
	Control group	4.584	3.014	0.149	-1.840	11.007
Control group	Training group (STG group)	-5.616	2.875	0.070	-11.743	0.511
	Training group (LTG)	-2.176	2.851	0.457	-8.253	3.901
	ARG	-4.584	3.014	0.149	-11.007	1.840

Based on estimated marginal means

a. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of “spoken voice quality” (cat2) from the lay assessor**

**Tests of between-subject effects**

Dependent variable:	post_cat2					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	166.809 <sup>a</sup>	4	41.702	1.718	0.198	0.314
Intercept	136.876	1	136.876	5.638	0.031	0.273
pre_cat2	117.859	1	117.859	4.855	0.044	0.245
Group	115.993	3	38.664	1.593	0.233	0.242
Error	364.141	15	24.276			
Total	13999.000	20				
Corrected total	530.950	19				

a. R squared = 0.314 (adjusted R squared = 0.131)

**Pairwise comparisons**

Dependent variable:		post_cat2				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>a</sup>	95% confidence interval for difference <sup>a</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	4.870	3.141	0.142	-1.825	11.565
	ARG	0.569	3.203	0.861	-6.258	7.396
	Control group	6.427	3.370	0.076	-0.756	13.610
Training group (LTG)	Training group (STG group)	-4.870	3.141	0.142	-11.565	1.825
	ARG	-4.301	3.316	0.214	-11.369	2.768
	Control group	1.557	3.240	0.638	-5.349	8.464



ARG	Training group (STG group)	-0.569	3.203	0.861	-7.396	6.258
	Training group (LTG)	4.301	3.316	0.214	-2.768	11.369
	Control group	5.858	3.715	0.136	-2.061	13.777
Control group	Training group (STG group)	-6.427	3.370	0.076	-13.610	0.756
	Training group (LTG)	-1.557	3.240	0.638	-8.464	5.349
	ARG	-5.858	3.715	0.136	-13.777	2.061

Based on estimated marginal means

a. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

### One-way ANCOVA test and pairwise comparisons of “spoken voice dynamism” (cat3) from the lay assessor

#### Tests of between-subject effects

Dependent variable:	post_cat3					
	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	237.246 <sup>a</sup>	4	59.311	1.104	0.391	0.227
Intercept	365.287	1	365.287	6.801	0.020	0.312
pre_cat3	89.496	1	89.496	1.666	0.216	0.100
Group	225.996	3	75.332	1.402	0.281	0.219
Error	805.704	15	53.714			
Total	18507.000	20				
Corrected total	1042.950	19				

a. R squared = 0.227 (adjusted R squared = 0.021)

### Pairwise comparisons

Dependent variable:		post_cat3				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>a</sup>	95% confidence interval for difference <sup>a</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	4.921	4.783	0.320	-5.273	15.114
	ARG	1.569	4.732	0.745	-8.518	11.656
	Control group	9.917	4.971	0.065	-0.678	20.512
Training group (LTG)	Training group (STG group)	-4.921	4.783	0.320	-15.114	5.273
	ARG	-3.351	5.102	0.521	-14.226	7.523
	Control group	4.996	4.676	0.302	-4.970	14.963
ARG	Training group (STG group)	-1.569	4.732	0.745	-11.656	8.518
	Training group (LTG)	3.351	5.102	0.521	-7.523	14.226
	Control group	8.348	5.389	0.142	-3.138	19.834
Control group	Training group (STG group)	-9.917	4.971	0.065	-20.512	0.678
	Training group (LTG)	-4.996	4.676	0.302	-14.963	4.970
	ARG	-8.348	5.389	0.142	-19.834	3.138

Based on estimated marginal means

a. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of “spoken voice production” (cat1) from the interpreting instructor assessor**

**Tests of between-subject effects**

Dependent variable:	post_cat1					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	257.890 <sup>a</sup>	4	64.473	7.537	0.002	0.668
Intercept	33.937	1	33.937	3.967	0.065	0.209
pre_cat1	118.090	1	118.090	13.805	0.002	0.479
Group	102.335	3	34.112	3.988	0.028	0.444
Error	128.310	15	8.554			
Total	7836.000	20				
Corrected total	386.200	19				

a. R squared = 0.668 (adjusted R squared = 0.579)

**Pairwise comparisons**

Dependent variable:		post_cat1				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>b</sup>	95% confidence interval for difference <sup>b</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	3.253	2.003	0.125	-1.016	7.522
	ARG	4.317*	1.852	0.034	0.370	8.264
	Control group	6.241*	1.850	0.004	2.298	10.185
Training group (LTG)	Training group (STG group)	-3.253	2.003	0.125	-7.522	1.016
	ARG	1.064	1.972	0.597	-3.139	5.266
	Control group	2.988	2.020	0.160	-1.316	7.293
ARG	Training group (STG group)	-4.317*	1.852	0.034	-8.264	-0.370

	Training group (LTG)	-1.064	1.972	0.597	-5.266	3.139
	Control group	1.924	1.854	0.316	-2.028	5.877
Control group	Training group (STG group)	-6.241 *	1.850	0.004	-10.185	-2.298
	Training group (LTG)	-2.988	2.020	0.160	-7.293	1.316
	ARG	-1.924	1.854	0.316	-5.877	2.028

Based on estimated marginal means

\*. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of “spoken voice quality” (cat2) from the interpreting instructor assessor**

**Tests of between-subject effects**

Dependent variable:	post_cat2					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	531.735 <sup>a</sup>	4	132.934	7.632	0.001	0.671
Intercept	4.659	1	4.659	0.268	0.613	0.018
pre_cat2	343.535	1	343.535	19.723	0.000	0.568
Group	164.793	3	54.931	3.154	0.056	0.387
Error	261.265	15	17.418			
Total	11838.000	20				
Corrected total	793.000	19				

a. R squared = 0.671 (adjusted R squared = 0.583)

**Pairwise comparisons**

Dependent variable:		post_cat2				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>b</sup>	95% confidence interval for difference <sup>b</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	6.239*	2.654	0.033	0.581	11.896
	ARG	3.961	2.654	0.156	-1.696	9.619
	Control group	7.568*	2.650	0.012	1.920	13.216
Training group (LTG)	Training group (STG group)	-6.239*	2.654	0.033	-11.896	-0.581
	ARG	-2.277	2.698	0.412	-8.028	3.473
	Control group	1.329	2.689	0.628	-4.402	7.060
ARG	Training group (STG group)	-3.961	2.654	0.156	-9.619	1.696

	Training group (LTG)	2.277	2.698	0.412	-3.473	8.028
	Control group	3.606	2.640	0.192	-2.020	9.233
Control group	Training group (STG group)	-7.568*	2.650	0.012	-13.216	-1.920
	Training group (LTG)	-1.329	2.689	0.628	-7.060	4.402
	ARG	-3.606	2.640	0.192	-9.233	2.020

Based on estimated marginal means

\*. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).

**One-way ANCOVA test and pairwise comparisons of “spoken voice dynamism” (cat3) from the interpreting instructor assessor**

**Tests of between-subject effects**

Tests of between-subject effects						
Dependent variable:	post_cat3					
Source	Type III sum of squares	df	Mean square	F	Sig.	Partial Eta squared
Corrected model	708.465 <sup>a</sup>	4	177.116	4.879	0.010	0.565
Intercept	111.794	1	111.794	3.080	0.100	0.170
pre_cat3	295.515	1	295.515	8.141	0.012	0.352
Group	533.215	3	177.738	4.897	0.014	0.495
Error	544.485	15	36.299			
Total	13803.000	20				
Corrected total	1252.950	19				

a. R squared = 0.565 (adjusted R squared = 0.450)

**Pairwise comparisons**

Dependent variable:		post_cat3				
(I) group		Mean difference (I-J)	Std. error	Sig. <sup>b</sup>	95% confidence interval for difference <sup>b</sup>	
					Lower bound	Upper bound
Training group (STG group)	Training group (LTG)	13.031*	3.905	0.005	4.709	21.353
	ARG	4.795	3.813	0.228	-3.333	12.922
	Control group	11.605*	3.813	0.008	3.478	19.733
Training group (LTG)	Training group (STG group)	-13.031*	3.905	0.005	-21.353	-4.709
	ARG	-8.236	3.938	0.054	-16.630	0.157
	Control group	-1.426	3.876	0.718	-9.688	6.836

ARG	Training group (STG group)	-4.795	3.813	0.228	-12.922	3.333
	Training group (LTG)	8.236	3.938	0.054	-0.157	16.630
	Control group	6.810	3.821	0.095	-1.334	14.955
Control group	Training group (STG group)	-11.605*	3.813	0.008	-19.733	-3.478
	Training group (LTG)	1.426	3.876	0.718	-6.836	9.688
	ARG	-6.810	3.821	0.095	-14.955	1.334

Based on estimated marginal means

\*. The mean difference is significant at the 0.05 level.

b. Adjustment for multiple comparisons: least significant difference (equivalent to no adjustments).



## Appendix L-1 Correlational Analysis Interpreting Performance and Vocal Performance (Pre and Post)

<b>Correlations</b>					
		Voice-related pre-to	Pre test	Voice-related post-to	Post test
Voice-related pre-to	Pearson Correlation	1	.638**	.408**	0.197
	Sig. (2-tailed)		0.000	0.001	0.132
	N	60	60	60	60
Pre test	Pearson Correlation	.638**	1	.363**	.350**
	Sig. (2-tailed)	0.000		0.004	0.006
	N	60	60	60	60
Voice-related post-to	Pearson Correlation	.408**	.363**	1	.789**
	Sig. (2-tailed)	0.001	0.004		0.000
	N	60	60	60	60
Post test	Pearson Correlation	0.197	.350**	.789**	1
	Sig. (2-tailed)	0.132	0.006	0.000	
	N	60	60	60	60
** . Correlation is significant at the 0.01 level (2-tailed).					
<b>bivariate correlation</b>					
By this bivariate correlational analysis, we intended to examine if the voice-related score will predict the interpreting performance					
<b>Results</b>					
Voice-related pre-score is highly correlated with pre interpreting test score ( $r=.638, p.<.001$ )					
The findings indicate voice-related score is highly correlated with interpreting performance					

## Appendix L-2 Correlational Analysis Attributes and Overall Interpreting Performance

CORRELATIONS  
 /VARIABLES=Pre1 PreInterpto  
 /PRINT=TWOTAIL NOSIG  
 /MISSING=PAIRWISE.

### Correlations

Notes		
Output Created		24-JUL-2020 17:51:31
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre1 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Pre1	Pre-Interp to
Pre1	Pearson Correlation	1	.371
	Sig. (2-tailed)		.004
	N	60	60
Pre-Interp to	Pearson Correlation	.371	1
	Sig. (2-tailed)	.004	
	N	60	60

#### CORRELATIONS

```

/VARIABLES=Post1 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

## Correlations

### Notes

Output Created	24-JUL-2020 17:53:48	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post1 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Correlations

		Post1	post-Interp to
Post1	Pearson Correlation	1	.479
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.479	1
	Sig. (2-tailed)	.000	
	N	60	60

#### CORRELATIONS

```

/VARIABLES=Pre2 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

### Correlations

#### Notes

Output Created	24-JUL-2020 17:54:51	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre2 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00

**Correlations**

		Pre2	Pre-Interp to
Pre2	Pearson Correlation	1	.483
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.483	1
	Sig. (2-tailed)	.000	
	N	60	60

**CORRELATIONS**

```

/VARIABLES=Post2 postInterp to
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

**Correlations****Notes**

Output Created	24-JUL-2020 17:55:22	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.

Syntax	CORRELATIONS /VARIABLES=Post2 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.03
	Elapsed Time	00:00:00.01

### Correlations

		Post2	post-Interp to
Post2	Pearson Correlation	1	.646
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.646	1
	Sig. (2-tailed)	.000	
	N	60	60

CORRELATIONS  
/VARIABLES=Pre3 PreInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

### Correlations

#### Notes

Output Created	24-JUL-2020 17:59:26	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.

Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre3 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time 00:00:00.00
	Elapsed Time 00:00:00.01

### Correlations

		Pre3	Pre-Interp to
Pre3	Pearson Correlation	1	.506
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.506	1
	Sig. (2-tailed)	.000	
	N	60	60

#### CORRELATIONS

/VARIABLES=Post3 postInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

### Correlations

#### Notes

Output Created	24-JUL-2020 17:59:47
Comments	
Input	Active Dataset DataSet1
	Filter <none>
	Weight <none>
	Split File <none>
	N of Rows in Working Data File 60

Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post3 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post3	post-Interp to
Post3	Pearson Correlation	1	.446
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.446	1
	Sig. (2-tailed)	.000	
	N	60	60

```
CORRELATIONS
/VARIABLES=Pre4 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

### Correlations

#### Notes

Output Created		24-JUL-2020 18:00:13
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>



	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre4 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.06
	Elapsed Time	00:00:00.01

### Correlations

		Pre4	Pre-Interp to
Pre4	Pearson Correlation	1	.458
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.458	1
	Sig. (2-tailed)	.000	
	N	60	60

### CORRELATIONS

```

/VARIABLES=Post4 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

### Correlations

#### Notes

Output Created		24-JUL-2020 18:22:25
Comments		
Input	Active Dataset	DataSet1

	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post4 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post4	post-Interp to
Post4	Pearson Correlation	1	.563
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.563	1
	Sig. (2-tailed)	.000	
	N	60	60

CORRELATIONS  
/VARIABLES=Pre5 PreInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

### Correlations

#### Notes

Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre5 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Pre5	Pre-Interp to
Pre5	Pearson Correlation	1	.414
	Sig. (2-tailed)		.001
	N	60	60
Pre-Interp to	Pearson Correlation	.414	1
	Sig. (2-tailed)	.001	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post5 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

### Correlations

## Notes

Output Created		24-JUL-2020 18:24:17
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post5 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

## Correlations

		Post5	post-Interp to
Post5	Pearson Correlation	1	.406
	Sig. (2-tailed)		.001
	N	60	60
post-Interp to	Pearson Correlation	.406	1
	Sig. (2-tailed)	.001	
	N	60	60

```

CORRELATIONS
/VARIABLES=Pre6 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
  
```

## Correlations

### Notes

Output Created		24-JUL-2020 18:25:09
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre6 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Pre6	Pre-Interp to
Pre6	Pearson Correlation	1	.572
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.572	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post6 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

## Correlations

### Notes

Output Created	24-JUL-2020 18:25:39	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post6 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post6	post-Interp to
Post6	Pearson Correlation	1	.648
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.648	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Pre7 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

## Correlations

### Notes

Output Created	24-JUL-2020 18:27:01	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre7 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Pre7	Pre-Interp to
Pre7	Pearson Correlation	1	.476
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.476	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post7 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
  
```

## Correlations

### Notes

Output Created	24-JUL-2020 18:27:54	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post7 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post7	post-Interp to
Post7	Pearson Correlation	1	.520
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.520	1
	Sig. (2-tailed)	.000	
	N	60	60

CORRELATIONS



```

/VARIABLES=Pre8 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

## Correlations

<b>Notes</b>		
Output Created		24-JUL-2020 19:10:46
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre8 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

## Correlations

		Pre8	Pre-Interp to
Pre8	Pearson Correlation	1	.353
	Sig. (2-tailed)		.006
	N	60	60
Pre-Interp to	Pearson Correlation	.353	1
	Sig. (2-tailed)	.006	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post8 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

## Correlations

<b>Notes</b>		
Output Created		24-JUL-2020 21:31:46
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post8 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.02

## Correlations

		Post8	post-Interp to
Post8	Pearson Correlation	1	.646
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.646	1

Sig. (2-tailed)	.000	
N	60	60

CORRELATIONS  
/VARIABLES=Pre9 PreInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

## Correlations

Notes		
Output Created		24-JUL-2020 21:32:19
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre9 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

## Correlations

		Pre9	Pre-Interp to
Pre9	Pearson Correlation	1	.288
	Sig. (2-tailed)		.026

	N	60	60
Pre-Interp to	Pearson Correlation	.288	1
	Sig. (2-tailed)	.026	
	N	60	60

**CORRELATIONS**

/VARIABLES=Post9 postInterpto  
 /PRINT=TWOTAIL NOSIG  
 /MISSING=PAIRWISE.

**Correlations**

**Notes**

Output Created		24-JUL-2020 21:33:03
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post9 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

**Correlations**

Post9	post-Interp to
-------	----------------

Post9	Pearson Correlation	1	.588
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.588	1
	Sig. (2-tailed)	.000	
	N	60	60

### Notes

Output Created		24-JUL-2020 21:43:15
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre10 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Notes

Output Created		24-JUL-2020 21:43:50
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>

	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post10 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

```
CORRELATIONS
/VARIABLES=Pre10 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

## Correlations

### Notes

Output Created		24-JUL-2020 21:47:20
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.

Syntax	CORRELATIONS /VARIABLES=Pre10 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Pre10	Pre-Interp to
Pre10	Pearson Correlation	1	.310
	Sig. (2-tailed)		.016
	N	60	60
Pre-Interp to	Pearson Correlation	.310	1
	Sig. (2-tailed)	.016	
	N	60	60

CORRELATIONS  
/VARIABLES=Post10 postInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

### Correlations

#### Notes

Output Created	24-JUL-2020 21:47:48	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.

Cases Used		Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post10 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Correlations

		Post10	post-Interp to
Post10	Pearson Correlation	1	.685
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.685	1
	Sig. (2-tailed)	.000	
	N	60	60

CORRELATIONS  
/VARIABLES=Pre11 PreInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

### Correlations

#### Notes

Output Created	24-JUL-2020 21:48:21	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60



Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre11 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Correlations

		Pre11	Pre-Interp to
Pre11	Pearson Correlation	1	.329
	Sig. (2-tailed)		.010
	N	60	60
Pre-Interp to	Pearson Correlation	.329	1
	Sig. (2-tailed)	.010	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post11 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

### Correlations

#### Notes

Output Created		24-JUL-2020 21:48:53
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>

	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post11 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post11	post-Interp to
Post11	Pearson Correlation	1	.626
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.626	1
	Sig. (2-tailed)	.000	
	N	60	60

CORRELATIONS  
/VARIABLES=Pre12 PreInterpto  
/PRINT=TWOTAIL NOSIG  
/MISSING=PAIRWISE.

### Correlations

#### Notes

Output Created		24-JUL-2020 21:49:17
Comments		
Input	Active Dataset	DataSet1

	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Pre12 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Pre12	Pre-Interp to
Pre12	Pearson Correlation	1	.546
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.546	1
	Sig. (2-tailed)	.000	
	N	60	60

```
CORRELATIONS
/VARIABLES=Post12 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
```

### Correlations

#### Notes

Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post12 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post12	post-Interp to
Post12	Pearson Correlation	1	.572
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.572	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Pre13 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

### Correlations

### Notes

Output Created		24-JUL-2020 21:50:15
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre13 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Correlations

		Pre13	Pre-Interp to
Pre13	Pearson Correlation	1	.534
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.534	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post13 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

### Correlations

### Notes

Output Created		24-JUL-2020 21:50:46
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post13 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.06
	Elapsed Time	00:00:00.01

### Correlations

		Post13	post-Interp to
Post13	Pearson Correlation	1	.712
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.712	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Pre14 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
  
```

## Correlations

### Notes

Output Created	24-JUL-2020 21:51:17	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre14 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Correlations

		Pre14	Pre-Interp to
Pre14	Pearson Correlation	1	.463
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.463	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Post14 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.
  
```

## Correlations

### Notes

Output Created	24-JUL-2020 21:51:42	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Post14 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.00
	Elapsed Time	00:00:00.01

### Correlations

		Post14	post-Interp to
Post14	Pearson Correlation	1	.683
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.683	1
	Sig. (2-tailed)	.000	
	N	60	60

```

CORRELATIONS
/VARIABLES=Pre15 PreInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```



## Correlations

### Notes

Output Created	24-JUL-2020 21:52:04	
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax	CORRELATIONS /VARIABLES=Pre15 PreInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.	
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

### Correlations

		Pre15	Pre-Interp to
Pre15	Pearson Correlation	1	.530
	Sig. (2-tailed)		.000
	N	60	60
Pre-Interp to	Pearson Correlation	.530	1
	Sig. (2-tailed)	.000	
	N	60	60

CORRELATIONS

```

/VARIABLES=Post15 postInterpto
/PRINT=TWOTAIL NOSIG
/MISSING=PAIRWISE.

```

## Correlations

Notes		
Output Created		24-JUL-2020 21:52:26
Comments		
Input	Active Dataset	DataSet1
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	60
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics for each pair of variables are based on all the cases with valid data for that pair.
Syntax		CORRELATIONS /VARIABLES=Post15 postInterpto /PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
Resources	Processor Time	00:00:00.02
	Elapsed Time	00:00:00.01

## Correlations

		Post15	post-Interp to
Post15	Pearson Correlation	1	.809
	Sig. (2-tailed)		.000
	N	60	60
post-Interp to	Pearson Correlation	.809	1
	Sig. (2-tailed)	.000	
	N	60	60

## Appendix M Sample of transcribing of assessor's comments

### Voice trainer assessor's comments – STG group (Participant A)

#### Video 1

<sup>1</sup>She has a clear accent, I should say. I could clearly understand her English. Good consonant and vowel clarity. <sup>2</sup>But apart from that, she was very non-expressive, and <sup>3</sup>not confident and <sup>4</sup>not connected at all with the audience. <sup>5</sup> She spoke in a low volume that I could hardly hear, and <sup>6</sup> she had quite a low modulation of pitch range, very low expressiveness. <sup>7</sup> The pace was a little fast. <sup>2</sup> She was sort of skimming over things. Therefore, she just lacked emphasis. <sup>8</sup> And her alignment and posture were very bent forward. She had her hips jutted forward. She seemed to be sort of deflated. <sup>9</sup>Physical tension balance and <sup>10</sup>breath...I'd say Average. But, she had really good accent or clear accent, I should say.

<sup>1</sup> Articulation: 3  
<sup>2</sup> Expressiveness: 1  
<sup>3</sup> Confidence: 1  
<sup>4</sup> Audience connection: 1  
<sup>5</sup> Voice projection: 1  
<sup>6</sup> Modulation of pitch range: 1  
<sup>7</sup> Pace and pause: 1  
<sup>8</sup> Posture and alignment: 1  
<sup>9</sup> Physical tension balance: 2  
<sup>10</sup> Breath: 2

#### Video 2

It is a wow! What an improvement! Second video of this lady. <sup>1</sup>She had really great emphasis, I wrote in the first evaluation that she lacked emphasis. <sup>2</sup>Her pace is ok and <sup>3</sup>her volume is much louder. <sup>1</sup>I found her expressiveness was very strong, and <sup>4</sup>her confidence and <sup>5</sup>her audience connection. <sup>6</sup>Wonderful alignment, posture, <sup>7</sup>physical tension balance. She stood upright with their shoulders back. <sup>8</sup> Her pitch is more comfortable than last time. Last time was very low. She just seemed really grounded and present, you know, in a really free and easy way. <sup>9</sup>And so, her breathing was quite nice this time. You know she still gasped a bit, but her breathing had improved dramatically. And <sup>1</sup>all her patterns and her expressiveness had improved as well. <sup>10</sup>Her articulation is good. So, I ended up giving her excellent. Because I could listen to her for quite some time. Well, I think her improvement has surprised me.

<sup>1</sup> Expressiveness: 4  
<sup>2</sup> Pace and pauses: 2  
<sup>3</sup> Voice projection: 3  
<sup>4</sup> Confidence: 4  
<sup>5</sup> Audience connection: 4  
<sup>6</sup> Posture & alignment: 4  
<sup>7</sup> Physical tension balance: 4  
<sup>8</sup> Modulation of pitch: 3  
<sup>9</sup> Breath: 3  
<sup>10</sup> Articulation: 4

## Lay assessor's comments – STG group (Participant A)

### Video 1

8 Her English is good. I can understand everything she says. 1But she was reading in a monotone. Sounds like she was doing a first reading. Sounds like she was not familiar with what she was reading. At times, it sounded like she was only just now starting to get used to the process of interpreting. 1There was not a lot of intonation. 2Not a lot of expression. 1More like a monotone. 2She needs more expression. 3She was not looking at audience. She was mainly looking at her notes. 5She is just standing and sort of looking down. She was trying to find her place (in her notes). 4The volume is a little bit low. 5No rhythm as she was just reading from the notes. She wasn't 2dynamic. Average impression. 7Not professional, but at the same time, she is still just only a student.

<sup>1</sup>Intonation: 1 (monotone)  
<sup>2</sup>Expressiveness: 1 (Reading rather than speaking)  
<sup>3</sup>Audience connection: 1 (no eye contact/looking at her notes too much)  
<sup>4</sup>Posture: 1 (Looking down)  
<sup>5</sup>Rhythm: 1 (Reading from the notes)  
<sup>6</sup>Confidence: 1 (not professional)  
<sup>7</sup>Articulation: 3

### Video 2

A lot better than the first video. 1she said 'um' and 'ah' a lot and also she also sort of said other things in between sentences. But not too much. he is pretty good. 2Good alignment. 7Much clearer. 3More expressive. 4She Looked at the audience. 3She had more meaning in her voice. 3More expression in her voice for sentences. 6She phrased her sentences better, so you could understand what she was saying and meaning. You got meaning even better. 6She was putting intonation into the phrases correctly. So, that was very good. Written comment: in second video, the speaker was much better in phrasing, intonation and expression and volume; video 2 is better.

<sup>1</sup>Fluent: 3 (Fillers)  
<sup>2</sup>Alignment: 3  
<sup>3</sup>Expressiveness: 3  
<sup>4</sup>Audience connection: 4 (Eye contact)  
<sup>5</sup>Phrasing: 3 (clearer meaning)  
<sup>6</sup>intonation: 3  
<sup>7</sup>Articulation:4

## Interpreting instructor assessor's comments – STG group (Translated transcription)

### Participant A

#### Video 1

Generally speaking, her interpreting delivery was pretty good. The only thing is, from the beginning to the end, 1she did not greet us and audience with any eye contact, which gave people a kind of feeling that she was just doing interpreting 3by reading mechanically from her notes (死背书) by rote method. This reduced the good impression of performance (打折扣) . 4Obviously, she lacked confidence. 2Her delivery was generally very smooth. But 5her voice was a bit soft and quite. Sometimes I can't hear her clearly. 1What a pity! No eye contact at all.

<sup>1</sup>Audience: 1  
<sup>2</sup>Fluency: 3  
<sup>3</sup>Expressive: 1  
<sup>4</sup>Professional confidence: 1  
<sup>5</sup>Voice projection:1  
<sup>6</sup>Confidence: 4  
<sup>7</sup>Fluency: 4  
<sup>8</sup>Audience: 4  
<sup>9</sup>Voice projection: 2  
<sup>10</sup>Expressiveness:3

#### Video 2:

1This time, she looks very professional. Good performance. 2Clear delivery, good eye contact. 2Excellent fluency. 5So very comfortable to her delivery. Very good. 5Her delivery is lively, much better than the previous one. One of the best. Apparent improvement than the first video (进步很明显). Written: Very significant improvement in her performance. She has established good eye-contact with the audience. 4Her voice is loud and clear. 2Her delivery is very fluent. Good job!

<sup>1</sup>Confidence: 4  
<sup>2</sup>Fluency: 4  
<sup>3</sup>Audience: 4  
<sup>4</sup>Voice projection: 3  
<sup>5</sup>Expressiveness:3

## Appendix N-1 Matrix of Numerical Ratings of Evaluative Content (STG)

Spoken voice categories	Spoken voice attributes	Voice Trainer Assessor										Lay Assessor										Interpreting Instructor Assessor										
		student A		Student B		Student C		student D		student E		student A		student B		student C		student D		student E		student A		student B		student C		student D		student E		
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	
production	Posture & alignment	1	4	1	3	1	3			3	3	1	3	1	3	1	3	1	3	1	3			1	3							
	Physical tension balance	2	4					3	3					1	1	1	1	3	3	2	4					3	3					
	Body movement			1	1										1	3	1	1														
	Breathing support	2	3	1	1	3	3			3	3							2	2													
	Release of speech muscles			1	1	3	3							1	1																	
Quality	Voice projection	1	3					1	3	1	3				3	3	4	4					1	3	1	3			3	3	1	3
	Resonance & timbre							1	3	1	1				3	3	2	2	3	3			1	3	3	3						
	Pace & pauses	1	2	1	1	3	3	1	3					1	3																	
	Speaking clarity	2	4	2	4	3	3			1	3	3	3	2	3	1	1	1	3	1	3			1	3			2	4	1	3	
	Modulation of pitch range	1	3	3	3											2	2															
Dyna-mics	Intonation and tone patterns					2	2					1	3	1	3			2	4	2	4											
	Rhythmic & fluent rendition			1	1			1	4	1	3	1	3	1	3	2	1	2	2	2	4	3	4			2	4	2	4	1	4	
	Expressive-ness	1	4	1	3	3	4	1	3	1	3	1	3	1	1	1	1	1	3	1	3	1	3	1	3					1	3	
	Vocal confidence	1	4	1	3	1	3	3	4	1	4	1	3	1	4	1	3	3	3	3	4	1	4	1	3	3	4	3	4	1	4	
	Audience connection	1	4	1	3	1	4	3	3	1	3	1	4	1	1	1	3	1	4	1	4	1	4	1	3	2	4	2	4	1	3	

\*4= Very positive 3= Positive 2= Neutral 1= Negative (blank) = No evaluative comment

## Appendix N-2 Matrix of Numerical Ratings of Evaluative Content (LTG)

Spoken voice Categories	Spoken voice Attributes	Voice Trainer Assessor										Lay Assessor										Interpreting Instructor Assessor									
		student A		Student B		Student C		student D		student E		student A		student B		student C		student D		student E		student A		student B		student C		student D		student E	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
production	Posture & Alignment			1	3	3	3	1	3	2	2			1	3	3	3	1	3	3	3	1	1	1	2			2	3		
	Physical tension balance	1	1	1	3	1	1	1	1	1	2	3	1					1	1									1	1	1	1
	Body movement	1	2	1	2	2	2			1	1			1	1									1	2						
	Breathing support	1	3	1	3	1	1	1	3	1	3					1	1	1	1	2	2										
	Release of speech muscles			1	1	1	1			1	1																	1	1	1	1
Quality	Voice projection			4	4	2	2	1	3	1	1					3	3	1	1	1	1					1	3	2	4	1	1
	Resonance & timbre	3	3	1	1	3	3	3	3							3	3	3	3	3	3	3	3	3	3	1	1				
	Pace & pauses	1	1	4	4	2	2	1	2	3	3	3	3					1	1	1	1									1	1
	Speaking clarity			4	4	2	3	1	3	1	3	4	4			1	3	1	3	1	3	1	1	3	3	2	2	1	1	1	1
	Modulation of pitch range	1	3	1	3	1	3	1	3							3	3	3	3	2	2										
Dyna-mics	Intonation and tone patterns					1	1			1	1	4	4			3	3	1	3			1	3								
	Rhythmic & fluent rendition	1	1	4	4			1	3	4	4	4	1	3	3	1	1	1	3	1	1	2	1	4	4	1	1	1	1	1	1
	Expressiveness	1	1	4	4	1	1			1	1	3	1			3	3	1	1	1	1									1	1
	Vocal Confidence	1	1	2	2			1	3	1	2			4	4	1	3	1	3	1	3	1	3	2	2						
	Audience connection	1	1	4	4	3	3	1	2	1	3	1	1	3	3	1	2	1	1	2	2			3	3	3	4	1	2	1	1

\*4= Very positive 3= Positive 2= Neutral 1= Negative (blank) = No evaluative comment

## Appendix N-3 Matrix of Numerical Ratings of Evaluative Content (ARG)

Spoken voice Categories	Spoken voice Attributes	Voice Trainer Assessor										Lay Assessor										Interpreting Instructor Assessor										
		student A		Student B		Student C		student D		student E		student A		student B		student C		student D		student E		student A		student B		student C		student D		student E		
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	
production	Posture & Alignment	1	2	2	3			1	1	1	2	1	1	2	2	1	3	1	1	3	3	1	3	2	4	3	3	1	1	1	3	
	Physical tension balance	1	1	1	3			1	1	3	3	1	1	1	3	1	1	1	1	2	2	1	1	1	1	2	3	1	1	2	3	
	Body movement	2	1	1	2	2	2															1	1			2	2					
	Breathing support	1	1	1	1	2	1	1	1	1	3	3	2	2	3	3	3	3	2	2	2	2					1	1	1	1		
	Release of speech muscles	1	1	1	3										1	3								1	1	3	3					
Quality	Voice projection	1	1	2	3	3	3	1	1	1	2	3	3	1	3	1	1	1	1	1	2	2	2	2	2	2	2	2	1	1	1	3
	Resonance & timbre			1	3									2	2	3	3	2	2	3	3					2	3	1	1	2	3	
	Pace & pauses					4	2					1	1	1	3			1	2			1	1	1	3	1	3	1	1	1	4	
	Speaking clarity	2	1	3	3	4	4			3	3	1	1	1	3	2	3	1	2	2	3	1	1	2	3	3	3	1	1			
	Modulation of pitch range	1	2	1	3	2	1			1	1	2	2								2	2	2	2	2			2	2	2	2	
Dyna-mics	Intonation and tone patterns	1	2	1	3					1	1	1	1	1	3	2	3	1	1	1	3	1	1	2	2	2	3					
	Rhythmic & fluent rendition	1	1	1	3					3	3	1	1			2	3	1	2	3	3	1	1	1	3	3	3	1	1	1	3	
	Expressiveness	2	2	1	3	2	1			1	1	1	1	1	3	1	3	1	3	1	3	1	3	1	3	2	2			1	3	
	Vocal Confidence	3	3			4	2	1	1	1	3	1	1	1	3	1	3	1	1	1	3			2	3	2	2	1	1	2	4	
	Audience connection	1	2	1	3	2	1	1	1	1	2	1	1	1	3	1	3	1	1	1	3	1	1	1	3			1	1	1	3	

\*4= Very positive 3= Positive 2= Neutral 1= Negative (blank) = No evaluative comment



## Appendix N-4 Matrix of Numerical Ratings of Evaluative Content (CoG)

Spoken voice Categories	Spoken voice Attributes	Voice Trainer Assessor										Lay Assessor										Interpreting Instructor Assessor									
		student A		Student B		Student C		student D		student E		student A		student B		student C		student D		student E		student A		student B		student C		student D		student E	
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
Production	Posture & Alignment	3	3	1	3	1	2			3	3			3	3	3	3	3	3	3	3			3	1						
	Physical tension balance	1	1	1	1	1	1	1	1	3	1	1	1	1	1	1	2	1	3	3	1	1	1	1	1	1	1	1	3	3	
	Body movement			1	3	1	1	1	1	3	1					1	1							1	1			1	1		
	Breathing support	1	1	1	2	1	2	1	1	3	3			1	1	3	3	3	3							2	1				
	Release of speech muscles	1	1	2	2			1	1														1	1							
Quality	Voice projection	1	3	3	3	3	3	1	1	3	3	3	3	3	3	3	3	3	3	3			2	2	3	3	3	3	3	3	
	Resonance & timbre			1	2	3	3	1	1	3	3			3	3	3	3	3	3	3				3	3			3	3		
	Pace & pauses	1	1	1	1			1	1	3	3			1	1	3	1	3	3	3	3	1	1	1	1	3	1				
	Speaking clarity	3	1			1	1	2	2	1	1	2	2	3	1	2	2	3	3	3	3	1	1	2	2	1	1	3	3	2	2
	Modulation of pitch range							1	3	3	3			3	3	3	3	2	2	2	2										
Dyna-mics	Intonation and tone patterns							1	1					2	2	1	1	3	3	4	1	1	1						3	3	
	Rhythmic & fluent rendition	1	3	2	2	1	1	1	1	3	3	1	1	1	1	3	1	2	1	1	1	1	1	1	1	1	3	1	3	3	
	Expressiveness					1	1	3	3			2	2	2	1	3	3	3	1	1	1	1									
	Vocal Confidence	1	2	1	1	1	1	3	1	3	3	3	3	1	1	3	1	3	1	3	3	1	1	1	1	3	3	1	1	3	3
	Audience connection	1	1	3	3	3	3	3	1	3	3	1	1	1	1	2	1	2	2	3	3	1	1	1	1	1	1	3	1	2	2

\*4= Very positive 3= Positive 2= Neutral 1= Negative (blank) = No evaluative comment