
Doctoral

Applied Arts

2010-05-01

Perception and Acquisition of Natural Authentic English Speech for Chinese Learners Using DIT's Speech Technologies

Yi Wang

Technological University Dublin

Follow this and additional works at: <https://arrow.tudublin.ie/appadoc>

Recommended Citation

Wang, Y. (2010) *Perception and Acquisition of Natural Authentic English Speech for Chinese Learners Using DIT's Speech Technologies*. Doctoral Thesis, Technological University Dublin. doi:10.21427/D7BC84

This Theses, Ph.D is brought to you for free and open access by the Applied Arts at ARROW@TU Dublin. It has been accepted for inclusion in Doctoral by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-NonCommercial-Share Alike 3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/)

Perception and Acquisition of Natural Authentic English Speech for Chinese Learners Using DIT's Speech Technologies

Yi Wang, BA

PhD Thesis

Dublin Institute of Technology

College of Arts & Tourism

Supervisors:

Mr. Dermot F. Campbell

Prof. Ciaran McDonnell

Prof. Michael McCarthy

September 2010

Volume 1/2

Abstract

Given that Chinese language learners are greatly influenced by their mother-tongue, which is a tone language, learning and coping with authentic English speech seems more difficult than for learners of other languages. The focus of the current research is, based on an analysis of spoken English and spoken Chinese, to help Chinese learners benefit from ICT technologies developed by the Dublin Institute of Technology. The thesis investigates the application of speech technologies in bridging the gap between students' internalised, idealised formulations and natural, authentic English speech.

Part of the testing carried out by the present author demonstrates the acceptability of a slow-down algorithm in facilitating Chinese learners of English in re-producing formulaic language. This algorithm is useful because it can slow down audio files to any desired speed between 100% and 40% without distortion, so as to allow language learners to pay attention to the real, rapid flow of 'messy' speech and follow the intonation patterns contained in them. The rationale for and the application of natural, dialogic native-to-native English speech to language learning is also explored. The Chinese language learners involved in this study are exposed to authentic, native speech patterns by providing them access to real, informal dialogue in various contexts.

In the course of this analysis, the influence of speed of delivery and pitch range on the categorisation of formulaic language is also investigated. The study investigates the potential of the speech tools available to the present author as an effective EFL learning facility, especially for speakers of tone languages, and their role in helping language learners achieve confluent interaction in an English L1 environment.

Contributions

The current study produces several original contributions to the field of formulaic language, and EFL learning and teaching by:

- Filling gaps in current literature on the relationship between formulaicity and prosody, which enriches the understanding and use of the communicative, pragmatic functions of formulaic language.
- Clearly demonstrating the effectiveness and acceptability of a slow-down facility, which can be applied not only to citation forms, but also to assist Chinese EFL learners imitating and re-producing the native acoustic ‘blur’ and intonation patterns of L1 speakers using formulaic sequences.
- Elaborating the implications of the study for a revised EFL pedagogy, specifically for Chinese learners and teachers, by exposing students to real speech samples occurring in everyday life in conjunction with language learning technologies.
- Developing an evaluation methodology for testing EFL speech production which incorporates an assessment of the messy ‘blur’ of rapid speech and communicative intonation patterns.
- Investigating an application of an innovative language learning resource, i.e., real, dynamic, L1-L1 native dialogues, which exposes EFL learners to authentic, natural English speech, and makes language learning advance in a real, contextualised environment, so as to enhance their EFL study and afford them access to L1 prosody.

These contributions form the basis of future investigation into effective perception and acquisition of natural English speech for EFL learners, and as such represent novel work in this field.

Declaration

I certify that this thesis which I now submit for examination for the award of Doctor of Philosophy, is entirely my own work and has not been taken from the work of others, save and to the extent that such work has been cited and acknowledged within the text of my work.

This thesis was prepared according to the regulations for postgraduate study by research of the Dublin Institute of Technology and has not been submitted in whole or in part for another award in any Institute.

The work reported on in this thesis conforms to the principles and requirements of the Institute's guidelines for ethics in research.

The Institute has permission to keep, lend or copy this thesis in whole or in part, on condition that any such use of the material of the thesis be duly acknowledged.

Signature _____ Date _____
Candidate

Acknowledgements

First and foremost, I wish to thank my parents, for their unselfish love and support. It was their enlightened support which influenced me to travel the academic road. I take strength from their unconditional love and encouragement which helped me to endure the difficulties and loneliness of my chosen path. I am very proud of my parents, and I owe the whole of my life and career to them.

I also wish to thank my supervisor, Dermot Campbell. Many thanks for his giving me the opportunity to start my academic exploration and helping me to mature professionally. Also many thanks for his endless patience. It is his consistent encouragement and support that gave me confidence and led me to this academic study. I enjoy my research journey with him and feel very lucky to be his student. Without his generous kindness and valuable guidance, in academic and personal friendship, my life would be very different.

I also wish to thank my supervisor Prof. Ciaran McDonnell. His insightful supervision in research methodology made me interested in and derive enjoyment from research. He influenced not only this present study, but also my further research career. I also very much appreciate his positive encouragement and kind understanding, which keeps me constant on this journey.

I also wish to give special, sincere thanks to Prof. Michael McCarthy. His keen mind provided me with an admirable model of professional behaviour and methodology; his outstanding, preeminent academic insight directed my research study; his gentle, unique sense of humour and his warm personality makes me enjoy every meeting and

discussion with him. I am also very much appreciative of his generous support and interest in my research, and his precious time in giving me feedback.

I also wish to thank my colleagues, Dr. Marty Meinardi, Dr. Bunny Richardson and Dr. Susan Norton. Many thanks for their generous sharing of their academic experience and giving me useful advice, and their warm friendly support, all of which eased my task and increased my confidence. I am very grateful for their kind support and personal friendship.

I also wish to thank the team of the SALERO and FLUENT Projects in the Digital Media Centre of DIT, for their lively brain-storming discussion sessions. And I also wish to thank the panel who contributed to the validation of the tests. Their support and the valuable data are much appreciated and contributed to the evaluation of my methodology.

I also wish to thank Charlie Pritchard, the EU SALERO Project and the Dublin Institute of Technology for their sponsorship which gave me the opportunity to commence my research work and supported me in the completion of the present study.

Finally, I also wish to thank the great co-operation from three Chinese schools, especially three teachers and 100 students, who gave me great support from June 2007 to May 2008 during the tests and training period for the present study. Their fulsome and generous co-operation are much appreciated.

Abbreviations

ASR	Automatic Speech Recognition
BBC	British Broadcasting Corporation
CANCODE	Cambridge and Nottingham Corpus of Discourse in English
CEFR	The Common European Framework of Reference
CSAL	Cognition, Speech and Audio Laboratory
DIT	Dublin Institute of Technology
DITCall	Digital Interactive Toolkit for Computer Assisted Language Learning
DMC	Digital Media Centre
DSC	Dynamic Speech Corpus
DSP	Digital Signal Processing
EFL	English as a Foreign Language
ELF	English as a Lingua Franca
FS	Formulaic Sequence
GpAc	Control Group A (third level students)
GpAt	Test Group A (third level students)
GpBc	Control Group B (senior cycle secondary students)
GpBt	Test Group B (senior cycle secondary students)
GpCc	Control Group C (junior cycle secondary students)
GpCt	Test Group C (junior cycle secondary students)
ICT	Information and Communications Technology
IELTS	International English Language Testing System
IPA	International Phonetic Alphabet
L1	First Language
L2	Second Language

NNS	Non-native Speech
NS	Native Speech
RP	Received Pronunciation
SC	Standard Chinese
STM	Short Term Memory
TELL	Technology-Enhanced Language Learning

Table of Contents

Abstract	ii
Contributions	iii
Declaration	v
Acknowledgements	vi
Abbreviations	viii
Table of Contents.....	x
Table of Figures	xvii
Table of Tables.....	xx
1. Introduction	1
1.1 Overview	1
1.2 Motivation of the Thesis: Discrepancies in Perception and Use of Authentic Natural English Speech between L1 Speakers and EFL Learners	4
1.2.1 Informal Spoken English Acquired and Used Naturally by L1 Speakers	4
1.2.2 Citation Standard Written English Taught to and Used by EFL Speakers.....	7
1.2.3 Mismatch in Production and Intelligibility of Natural English Speech between L1 Speakers and Non-L1 Language Learners in Real English Communication	9
1.3 Aims of the Thesis.....	11
1.4 Contents of the Thesis	15
2. Literature Review	17
2.1 Introduction	17
2.2 Aspects of Production and Decoding of Spoken English.....	17
2.2.1 Salient Phonological Characteristics Signal Potential Attitudes in Oral English...	18
2.2.2 Intonation Units and Communicative Features	24
2.2.3 Natural Flowing Connected Speech and Citation Form.....	30
2.2.4 Different Decoding Processes of English Speech by L1 and L2 Users	36

2.2.5 Summary.....	42
2.3 Perception of Authentic Natural Spoken English by L2/EFL Learners	42
2.3.1 Facilitating Intelligibility by Increasing Exposure to Authentic English Speech ..	43
2.3.2 Speaker and Listener Convergence Based on Shared Contextual Knowledge	47
2.3.3 Formulaic Language and Spoken English	52
2.3.4 Summary.....	59
2.4 Learning and Acquisition of Formulaic Language.....	60
2.4.1 Multi-criteria Categorisation of Formulaic Language.....	60
2.4.2 Phonological Perception of Formulaic Language	65
2.4.3 Acquisition and Use of Formulaic Language by L1 and L2 Learners	70
2.4.4 Enhancing the Learning and Acquisition of Formulaic Language for Non-L1 Speakers.....	76
2.4.5 Summary.....	80
2.5 Conclusion.....	81
3. Spoken Chinese and English Learning and Teaching in China.....	82
3.1 Introduction	82
3.2 L1 – A Barrier to Adopting English Language Speech Patterns.....	82
3.2.1 Introduction: Chinese Language and its Sounds	83
3.2.2 Basic Properties of the Chinese Language	88
3.2.3 ‘Staccato’ Connected Speech with Clear Word Boundaries	93
3.2.4 Summary.....	97
3.3 Constraints in Current Educational Situation – How is English Learned and Taught in China?.....	98
3.3.1 Static Learning and Teaching Process Influenced by Traditional Cultural and Education System	98
3.3.2 Large Classes and Limited Teaching/Learning Resources.....	101
3.3.3 Inefficient Pronunciation Teaching and Listening	103

3.3.4 Insufficient Exposure to Natural Authentic English	106
3.3.5 Summary.....	108
3.4 Problems Encountered by Chinese Language Learners in English Native-speaking Communities.....	108
3.4.1 Issues for L1 Listeners with Chinese Speakers’ Mispronunciation and Misprosody	109
3.4.2 Issues for Chinese Learners in Decoding English Speech	111
3.4.3 Summary.....	113
3.5 Conclusion.....	113
4. Discussion of Literature Review.....	115
4.1 Introduction	115
4.2 Issues in Current Research	115
5. Current Research into TELL Tools.....	119
5.1 Introduction	119
5.2 Traditional TELL Tools	119
5.3 Innovative Language Learning Technologies Devised in DIT	123
5.3.1 Slow-down Algorithm.....	124
5.3.2 Vowel Pronunciation.....	127
5.3.3 Access to and Use of Natural Dialogic Native English Speech Assets.....	129
5.4 Conclusion.....	135
6. Research Design.....	137
6.1 Introduction	137
6.2 Rationale and Procedures in Current Research	137
6.2.1 Analysis of Formulaic Language.....	137
6.2.2 Speech Technologies Employed.....	138
6.2.3 Research Undertaken in Application of Speech Technologies to Acquisition of Formulaic Language for Chinese EFL Learners	141

6.3 Conclusion.....	144
7. Analysis of Formulaic Language.....	146
7.1 Introduction	146
7.2 Analysis of Natural Authentic Interactive English Speech	146
7.2.1 Inclusive Transcription via Orthographic Text	146
7.2.2 Segmenting and Tagging the Units	149
7.2.3 Initial Findings – Formulaic Language	157
7.2.4 Limitations of Analysis	173
7.3 Conclusion.....	174
8. Speech Technologies Applied to Acquisition of Formulaic Language	176
8.1 Introduction	176
8.2 Application and Evaluation of Speech Learning Technologies	176
8.2.1 Test 1	177
8.2.1.1 DESIGN OF TEST 1.....	177
8.2.1.2 SUBJECTS FOR TEST 1.....	185
8.2.1.3 METHODOLOGY FOR TEST 1.....	188
8.2.1.4 ANALYSIS OF TEST 1.....	193
8.2.1.5 OVERVIEW OF TEST 1 AND SUGGESTIONS FOR TRAINING AND TEST 2.....	199
8.2.2 Training for Test Participants.....	202
8.2.2.1 PURPOSE OF TRAINING.....	202
8.2.2.2 EXPLANATION BEFORE TRAINING.....	202
8.2.2.3 TRAINING MATERIALS.....	203
8.2.2.4 MONITORING THE TRAINING PROCESS.....	204
8.2.2.5 STUDENT REACTION TO TRAINING SESSION.....	208
8.2.3 Test 2	217
8.2.3.1 DESIGN OF TEST 2.....	217
8.2.3.2 SUBJECTS AND METHODOLOGY FOR TEST 2.....	222
8.2.3.3 ANALYSIS OF TEST 2.....	223
8.2.3.4 PANEL FOR FURTHER VALIDATION OF METHODOLOGY.....	238
8.2.3.5 DISCUSSION OF RESULTS OF TEST 2.....	257
8.2.4 Discussion of Tests.....	258
8.3 Conclusion.....	262

9. Discussion of Present Research	264
10. Conclusion	269
10.1 Conclusion	269
10.1.1 Overall Conclusion	269
10.1.2 Summary of Work	270
10.1.3 Contributions to the Body of Knowledge	272
10.1.4 Limitations and Implications of the Current Study	273
10.1.5 Reflections on the Current Study	275
10.2 Recommendation for Further Research	277
10.2.1 Further Analysis of the Relationship between Formulaicity and Prosody	277
10.2.2 Analysis Related to the Realisation of Formulaic Language	278
10.2.3 Fuller Investigation of Flow Unit and Other Units	278
10.2.4 Further Development of DIT's Speech Technologies for the Acquisition and Use of L1 Speech by EFL Learners	279
10.2.5 Further Analysis of Construction of Dialogic Con-fluence	279
References	281
Appendix 1: Further Potential of Authentic English Speech for Teaching EFL Learners 1	
Appendix 2: Snippets for Testing and Training Samples	53
Appendix 3: Questionnaire 1 for Both Control Group and Test Group (English Version)	54
Appendix 4: First Version of Data Assessment for Test 1	55
Appendix 5: Refined Version of Data Assessment for Test 1	56
Appendix 6: Explanation of Training Purpose and Training Materials for Control Group	57
Appendix 7: Explanation of Training Purpose and Training Materials for Test Group ..	59
Appendix 8: Transcripts and Analysis of Training Materials 09/2007 for Control Group ..	62
Appendix 9: Transcripts and Analysis of Training Materials 09/2007 for Test Group ..	67

Appendix 10: Transcripts and Analysis of Training Materials 10/2007 for Control Group	73
Appendix 11: Transcripts and Analysis of Training Materials 10/2007 for Test Group	78
Appendix 12: Transcripts and Analysis of Training Materials 11/2007 for Control Group	83
Appendix 13: Transcripts and Analysis of Training Materials 11/2007 for Test Group	88
Appendix 14: Transcripts and Analysis of Training Materials 12/2007 for Control Group	93
Appendix 15: Transcripts and Analysis of Training Materials 12/2007 for Test Group	99
Appendix 16: Transcripts and Analysis of Training Materials 03/2008 for Control Group	105
Appendix 17: Transcripts and Analysis of Training Materials 03/2008 for Test Group	112
Appendix 18: Transcripts and Analysis of Training Materials 04/2008 for Control Group	119
Appendix 19: Transcripts and Analysis of Training Materials 04/2008 for Test Group	124
Appendix 20: Questionnaire 2 on Training Process for Control Group (Chinese Teachers of English).....	129
Appendix 21: Questionnaire 2 on Training Process for Test Group (Chinese Teachers of English).....	131
Appendix 22: Feedback 1 on Questionnaire 2 on Training Process for Test Group (12/10/2007)	133
Appendix 23: Feedback 2 on Questionnaire 2 on Training Process for Control Group (06/11/2007)	135
Appendix 24: Feedback 3 on Questionnaire 2 on Training Process for Test Group (20/11/2007)	137
Appendix 25: Feedback 4 on Questionnaire 2 on Training Process for Test Group (23/12/2007)	140
Appendix 26: Dialogues Provided by Chinese English Teacher for Reference.....	143
Appendix 27: Questionnaire 3 for Test Group (English Version)	152

Appendix 28: Questionnaire 4 for Control Group (English Version)	154
Appendix 29: Handout of Training Samples Practised by GpB.....	155
Appendix 30: Instructions for Panel Evaluation – Step 1	163
Appendix 31: Instructions for Panel Evaluation – Step 2	164
Appendix 32: Acceptability of Slow-down Facility – Questionnaire for Linguistic Researchers at BAAL Conference 2008.....	166
Appendix 33: Presentations and Publications	171

Table of Figures

Figure 1: Idealised phonological form vs. natural acoustic blur produced by L1 English speaker	34
Figure 2: Average normalised intensity for speakers A, C over 10 second frames with 50% overlap	52
Figure 3: Average normalised speech rate for speakers A, D over 20 second frames with 50% overlap	52
Figure 4: SC utterance produced at normal speed	97
Figure 5: SC utterance produced at a faster speed	97
Figure 6: Slow-down algorithm	127
Figure 7: Initial (proposed) interface of Dynamic Speech Corpus	131
Figure 8: Waveform representation of dialogue produced in the CSAL	134
Figure 9: Transcription of natural conversational speech	149
Figure 10: Segmentation of natural conversational speech into flow units	154
Figure 11: Segmenting and tagging the sequences based on flow unit	155
Figure 12: Pitch contours of FS 'you know' uttered at different speeds of delivery	159
Figure 13: Examples of FS 'I mean'	160
Figure 14: Different realisations of speed and pitch range of FS 'I mean'	161
Figure 15: Towards inclusion of prosody in FS typology	169
Figure 16: Sample analysis in Test 1	194
Figure 17: Analysis of Test 1 results	197
Figure 18: Q1 in Questionnaire 3 (Test Group)	209
Figure 19: Q2 in Questionnaire 3 (Test Group)	210
Figure 20: Q3 in Questionnaire 3 (Test Group)	210
Figure 21: Q4 in Questionnaire 3 (Test Group)	211
Figure 22: Q5 in Questionnaire 3 (Test Group)	212

Figure 23: Q6 in Questionnaire 3 (Test Group)	213
Figure 24: Q1 in Questionnaire 4 (Control Group)	215
Figure 25: Q2 in Questionnaire 4 (Control Group)	216
Figure 26: Q3 in Questionnaire 4 (Control Group)	217
Figure 27: Sample analysis in Test 2.....	223
Figure 28: Analysis of two samples used in both Tests 1 and 2.....	226
Figure 29: Analysis of Test 2 results.....	227
Figure 30: Improvement by participants (TG+CG) in Tests 1 and 2	230
Figure 31: Analysis of Tests 1 and 2.....	231
Figure 32: Improvement compared between Test Group and Control Group in Tests 1 and 2	231
Figure 33: Score '4' achieved by participants in Tests 1 and 2.....	233
Figure 34: Sample 2 scores achieved by Test Group in Tests 1 and 2 (across all score ranges)	234
Figure 35: Sample 2 scores achieved by Control group in Tests 1 and 2 (across all score ranges)	235
Figure 36: Sample 8/4 scores achieved by Test Group in Tests 1 and 2 (across all score ranges)	236
Figure 37: Sample 8/4 scores achieved by Control Group in Tests 1 and 2 (across all score ranges).....	236
Figure 38: Exemplar production of chunk 'you know' in Sample 6	237
Figure 39: Participant's performance of chunk 'you know' in Sample 6	238
Figure 40: Sample 3 evaluation comparison between naïve language users and the author (Step 2)	251
Figure 41: Sample 3 evaluation comparison between language teachers and the author (Step 2)	252
Figure 42: Sample 3 evaluation comparison between panel and the author (Step 2)...	252
Figure 43: Sample 4 evaluation comparison between naïve language users and the author (Step 2)	253

Figure 44: Sample 4 evaluation comparison between language teachers and the author (Step 2)	254
Figure 45: Sample 4 evaluation comparison between panel and the author (Step 2)... ..	254
Figure 46: Attribute Tree.....	5
Figure 47: Attribute values under the tag ‘speaker intention’	5
Figure 48: Intonation pattern and lexical filler used for the realisation of turn keeping ..	17
Figure 49: Interactive turn behaviour between speakers	17

Table of Tables

Table 1: Consonant phonemes in SC.....	85
Table 2: Vowel phonemes in SC	87
Table 3: Definition of speeds of delivery	156
Table 4: The fastest and slowest speeds of delivery of FS 'you know'	158
Table 5: Phonological realisation of collocations	163
Table 6: Phonological realisation of semi-fixed frames	164
Table 7: Phonological realisation of idioms	164
Table 8: Phonological realisation of unmarked chunks	166
Table 9: Phonological realisation of marked chunks	166
Table 10: Phonological realisation of marked chunks with a slow delivery speed.....	167
Table 11: Phonological realisation of marked chunks with a fast delivery speed.....	167
Table 12: Phonological realisation of grammatical paradigms	169
Table 13: Position of FS and its delivery speed (FS followed by a variable)	171
Table 14: Position of FS and its delivery speed (FS preceded by a variable)	171
Table 15: Position of FS and its delivery speed (FS preceded and followed by a variable).....	172
Table 16: Testing samples in Test 1	183
Table 17: Questionnaire 1 of participants in Test 1	187
Table 18: Analysis of correlation between score and amount of English exposure.....	198
Table 19: Linguistic features in training materials	203
Table 20: Q1 in Questionnaire 3 (Test Group).....	209
Table 21: Q3 in Questionnaire 3 (Test Group).....	211
Table 22: Q4 in Questionnaire 3 (Test Group).....	212
Table 23: Q5 in Questionnaire 3 (Test Group).....	212

Table 24: Q1 in Questionnaire 4 (Control Group)	215
Table 25: Testing samples in Test 2	220
Table 26: Choice of participant samples	240
Table 27: Sample 3 evaluation comparison between panel and the author (Step 1)....	243
Table 28: Swap patterns employed for ranking evaluation for both Samples 3 and 4 (Step 1)	244
Table 29: Sample 3 evaluation values based on swap patterns (Step 1)	245
Table 30: Sample 3 evaluation based on swap patterns and equivalent values (Step 1)	246
Table 31: Sample 4 evaluation comparison between panel and the author (Step 1)....	247
Table 32: Sample 4 evaluation values based on swap patterns (Step 1)	248
Table 33: Sample 4 evaluation based on swap patterns and equivalent values (Step 1)	249
Table 34: Sample 3 evaluation comparison between panel and the author (Step 2)....	251
Table 35: Sample 4 evaluation comparison between panel and the author (Step 2)....	253
Table 36: Segmentation of orthographic text into three hierarchical levels.....	3
Table 37: Segmentation of Harry-Sue dialogue based on three hierarchical levels.....	4
Table 38: Attribute values in the dialogue Harry-Sue.....	6
Table 39: Potential correlation between speaker intention and its realisations.....	8
Table 40: Turn constructions and their realisations	12
Table 41: Acceptability of the slow-down technology by linguistic researchers (1) ...	169
Table 42: Acceptability of the slow-down technology by linguistic researchers (2) ...	169

1. Introduction

1.1 Overview

This thesis is about spoken English. Its approach differs slightly from other research in the area, as it focuses on the production of speech and how prosody functions in the delivery of the speaker's intention in real, natural, dialogic conversation.

This chapter gives an overview of the thesis, its motivation, aims and research questions. It also takes an initial look at authentic language (as opposed to citation forms) from an L1 speaker and a learner perspective. The chapter ends with an outline of the main contents of the thesis.

Spoken language, as opposed to its written form, as indicated by Crystal (2002, pp.94-97, 2003a, p.291), is 'dynamic' and 'time-bound'. The information transmission is produced spontaneously; the communication is constantly modified, clarified and negotiated; interruptions and overlapping are very frequent. Most speech is unplanned, real-time production, with less complex and precise grammatical structures; and less elaborately balanced syntactic patterns. Fragmented utterances, repetition, rephrasing and replacing, and also false starts characteristically occur in natural informal conversation. Given that most spoken language occurs in face-to-face interaction, interlocutors can also make use of body language to aid their communication. The lexicon in conversational language often tends to be informal, vague and with more deictic expressions. Data obtained by O'Keeffe, et al. (2007) also show that more interpersonal lexical items are used in spoken language.

A unique feature of conversational language, as Crystal (2003a) points out, is prosody. ‘The many nuances of intonation, as well as contrasts of loudness, tempo, rhythm, and other tones of voice cannot be written down with much efficiency’. That is to say, the large range of nuances can mostly and best be expressed by prosody. This description, to some extent, echoes the investigation carried out by Mehrabian (1981) in terms of the communication of feelings and attitudes. These highlight the importance of non-verbal features in spoken language. In his studies, a ‘7%-38%-55% Rule’ is proposed, in which 38% of face-to-face communications is expressed by tone of voice (pp.75-80). Whenever there is a clash between the words uttered and the non-verbal cues used, the listener tends to believe the communicative value conveyed by tonality and facial expression (Mehrabian, *ibid.*).

The current research therefore starts with the discrepancy between spoken English and Chinese, especially the east-west prosodic divide (detailed in Chapters 2 and 3). It concentrates on how prosodic features, rather than form or phonological substance, contribute to the various realisations of formulaic language (outlined in Section 7.2.3) in authentic, interactive, native-to-native English speech.

A traditional, syntactic approach to language learning and teaching focuses on vocabulary acquisition, grammar explanation and translation. The advantage of this approach is to improve learners’ competence in respect of acquiring a large vocabulary and accurate grammatical forms. This linguistic competence is mainly acquired via the written word. This is the main learning model adopted by Chinese learners of English.

In contrast to this grammar-translation approach, the discoursal approach prefers to improve learners’ communicative competence. Rather than confining itself to the

conventional, ideational and stable written text, this approach appreciates that learning emerges from the ‘real’ speech samples of everyday language and acknowledges the pragmatic functions of language in a discourse environment. This holistic learning approach also facilitates lexical learning by making accessible multi-word expressions and chunks of formulaic language in the context in which they naturally occur, which potentially produces and improves conversational fluency; it also makes the acquisition of grammar more pragmatic and plausible. The corpus-informed approach proposes that language learners should be exposed to the ‘real’ speech production of authentic spoken language, in which natural features of spoken production, other than citation forms, can be demonstrated, i.e., co-articulatory production, natural intonation patterns, and the ‘messy’ and ‘sloppy’ flow of acoustic signals. Becoming immersed in natural transactional and interactional activities can provide language learners with the opportunity to hear and understand spoken language, and use it appropriately.

It is the latter descriptive, data-driven approach which establishes the theoretical framework of the present research. Given that most EFL students learn in a non English-speaking environment, they thus need an interface which allows them to access the ‘real’ English speaking community. Real, natural, L1-L1 dialogues (intended for inclusion in DIT’s Dynamic Speech Corpus) are recorded at a very high audio quality and have the ability to separate the speakers, even when they are engaged in cross-talking. In contrast to some existing spoken corpora (sometimes read out from scripts or performed during a television or radio programme), the early assets of the Dynamic Speech Corpus (DSC) contain many of the features of dialogic speech. Naturalness is one of the most salient properties. The incorporation of a slow-down facility also makes available to learner and researcher alike the acoustic ‘blur’ and intonation patterns of informal speech, especially with formulaic language. Other speech features included in

natural, interactive conversations are: false starts, self-repairing, topic-changing, turn-behaving and back-channelling. It is anticipated that the unique assets in the DSC will bridge the gap for language learners surrounded by non-native English speech and help them to access real, natural, dialogic English speech production, so as to communicate naturally and fluently with L1 English speakers without panic.

Since it is the features of native-to-native dialogic, informal speech which form the basis of the current study, it is natural that the interactive discourse present in DSC dialogues be chosen as the vehicle for training and testing the research questions included in the thesis.

1.2 Motivation of the Thesis: Discrepancies in Perception and Use of Authentic Natural English Speech between L1 Speakers and EFL Learners

1.2.1 Informal Spoken English Acquired and Used Naturally by L1 Speakers

Different genres and registers used, and different contexts in which language occurs produce different communicative values. Carter and McCarthy (1997) pinpoint the nature of spoken English naturally used by L1 speakers. There are two main activities in communication – transaction, which aims to fulfil the business or get things done; and interaction, which refers to the more interactive activities normally involved in transactional business and oriented by personal and social relations. As McCarthy (1991, p.137) points out, it is these unexpected interactive conversations which can cause non-L1 language learners to panic. Key features in unplanned conversations are investigated and displayed by Carter and McCarthy (1997, 1995, 2006), from the point of view of grammatical structure and vocabulary usage.

Spoken grammar, as McCarthy and Carter (2006a, p.27) argue, is characterised by 'factual utterances'. It frequently contains some 'mistakes', appearing as ungrammatical forms, incomplete sentences, situational ellipsis, etc., which are regarded as 'bad' performance in writing text. They occur, however, very often in informal, natural, unplanned interaction. L1 speakers do not pay much attention to these problematic 'flaws', and focus on maintaining the communication.

Some vocabulary usage is very typical of spoken data: for example, vague expressions, situational deixis and ellipsis in casual, moment-to-moment conversations. Vague expressions, such as 'or something', 'or whatever', often give the impression of careless or sloppy linguistic behaviour. From the L1 perspective, however, these expressions serve to soften the communication and sound less-authoritative, more relaxed and friendly. Some small words, such as 'well', 'actually', ' 'cos', 'just', as argued by Hasselgreen (2004), play an important role in interaction and contribute to fluency.

There are also large portions of interactional vocabulary expressing interpersonal functions, rather than transactional purposes in a communication. Two types are discourse markers and back-channelling. Discourse markers, investigated by Schifffrin (1987), such as 'I mean', 'you know', play many important roles. They can serve as a kind of navigation marker, and orientate the stretches of conversation and easily guide the listeners to the next topic or the following section of the discourse. They can also be used to change topic, return to an interrupted topic, act as boundary marker, etc., or be used for checking shared knowledge between the interlocutors. In addition, back-channelling, as investigated by Yngve (1970, cited in McCarthy, 1991), is also marked as one of the prominent features in informal, colloquial, native English speech, for

example, 'yeah', 'that's right', 'um'. Back-channelling can act as a sign to indicate a potential upcoming interlocutor utterance, or to show a positive engagement, without any intention of cutting in or taking possession of the floor. These coherent flow speech markers and natural back-channel responses are sometimes regarded as 'annoying' or 'redundant' by non-L1 language learners, but they are precisely the features which mark L1 natural speech.

Brown and Yule (1983a) also unfold a picture of how language is used, especially spoken language, based on the analysis of conversational English. The main function of spoken language, they assert, is to maintain a kind of social relationship, to be nice (or not) to other people, which is the interactional function. The transactional function is sometimes embedded within an interactional activity, and vice-versa. The difference between transactional language and interactional language lies in the fact that relatively clear and more specific language tends to be used in transactional contexts (p.14). However, the less precise use of language in interactional situations, as Brown and Yule (*ibid.*, p.15) point out, does not affect the communication.

Brown and Yule (1983a, pp.97-99) then list several features which normally occur in natural, informal spoken English: the use of reduced forms (as opposed to the full version in written English), the use of fillers (either verbal or non-verbal), the use of pauses, the use of repetition (Tannen, 1989), and the use of 'policy of ease of articulation' (Ladefoged, 1993, p.267). These natural, interactive features often appear in the spoken data of L1 speakers, and this kind of spoken communicative skill is generally acquired naturally by L1 speakers (Brown & Yule, *ibid.*, p.19).

The studies referred to above illustrate some high frequency regularities occurring in L1 speakers' everyday language behaviour. In general, L1 English speakers can naturally acquire and produce dialogic and interactive English speech.

1.2.2 Citation Standard Written English Taught to and Used by EFL Speakers

As mentioned earlier in Section 1.1, most language learners learn their L2/EFL mainly in a classroom environment, and thus written work dominates their learning activities, as opposed to spoken language and an orientation towards speech acts and discourse-level phenomena (Sinclair & Renouf, 1988).

In the literature of English learning and teaching, Brown and Yule (1983a) state that the situation of non-L1 language teaching is mainly governed by the study of written language. Text materials provided to EFL learners were mostly in written form, based on standard English, up until the end of World War II. The teaching of spoken English only started in the 1950s, beginning with the teaching of pronunciation, then expanded to the teaching of listening skills. These circumstances mirror the situation of English teaching and learning in China, where there was less emphasis on the spoken word.

Chinese students start learning English at the age of 11 when they are in their third year in primary school (based on guidelines given by the Ministry of Education in China). Only a few new words and simple drills are taught in class. Knowing how to read these words and sentences, and how to write and remember them are the main tasks for both English teachers and students. The course-books used are old fashioned – mainly written exercises dominate. Even though some new editions appeared in the last 20 years, there was little change over the previous decades. There was less realisation of

the importance of spoken English in the minds of pupils or even English teachers, nor were there any oral examinations. There was almost no practical oral communication or emphasis on intelligibility, or at least insufficient emphasis. Even today, the English class is like a static pool, with little or no interaction between English teacher and students, as pointed out by Gu (2003) among others. This mono-directional classroom dynamic is detailed later in Section 3.3.

Given that in China EFL learners are long and constantly exposed to and educated in a written-language controlled environment as mentioned above, when these language learners engage in real communication with L1 speakers, or even with other non-L1 speakers from different learning situations, they employ this scholarly defined, standard written English style which is odd and out of tune with the language used by their interlocutors.

This observation is also made by Brown and Yule (1983b, pp.14-19), when they point out the different features of written and spoken English, which they call ‘sentences’ and ‘utterances’. Sentences are usually heavy, dense with information and with a relatively complex and well-organised syntactic structure. Spoken language often used by EFL learners generally sounds like written texts read aloud with very clear pronunciation and regular, frequent pauses (Brown & Yule, 1983a, p.21). Their speech is characterised by a lack of natural prominences and rhythms. The signals are deliberately delivered one by one, rather than in a connected flow of blended phonemes. Speech of many non-L1 speakers is often more like a monologic announcement – without interactive expressions and responses. The speech is more precise with less repairs or modification. The language produced by many EFL speakers tends to contain complete, perfectly

formed sentences, with few short, or poorly styled (though understandable) phrases or sequences.

In summary, language learners learning their L2/EFL mainly based on written texts might produce correctly formed sentences; yet their speech might not be appropriate in the contexts in which it occurs, since, as Gimson (2001, p.4) points out, it is the spoken, rather than the written form, that represents the essence of the English language.

1.2.3 Mismatch in Production and Intelligibility of Natural English Speech between L1 Speakers and Non-L1 Language Learners in Real English Communication

When comparing the nature of spoken English and its idealised written form as used by most EFL learners, it is obvious that there is a difference in production and intelligibility of natural flow English speech between L1 speakers and non-L1 language learners. Field (2004, p.114) generally summarises some phonological characteristics of what he calls ‘foreigner talk’ (i.e., L1 speakers talking to foreigners), such as slower delivery speed, longer pauses, greater segmentation in the articulation of words, more stress marks, clearer and more deliberate articulation, and fewer assimilations. Zielinski (2008) supports this point of view by investigating the impact on L1 listeners, with respect to intelligibility, of different speech features used by non-L1 language learners. The findings illustrate that non-standard stress patterns and word segments cause L1 speakers uncertainty and finally lead to the misinterpretation of EFL learners’ intentions.

Brown (1990, p.2, pp.5-7, pp.144-48) also states that many foreign students of English who live in a non English-speaking country are potentially in danger of never having

the ability to acquire and use ‘an appropriate style of pronunciation’. One of the main reasons is that, since students learn English mostly in a classroom environment and learn from teachers, their English is learned in terms of ‘words’ and ‘sentences’ and teachers tend to present a slow, clear model to students. Brown (ibid.) argues that it is a teachable model, but not the only ‘correct’ or ‘acceptable’ style of natural speech, especially for advanced students. In the first place, a slow and clear pronunciation demonstrates nothing about the normal patterns of the stream of informal speech; secondly, the presentation of stresses and intonation patterns in isolation or in short sequences ignores the important communicative values of the prosody, and also in no way prepares language learners to listen to and cope with spontaneous conversations. There is still a considerable number of advanced language learners, especially Chinese EFL learners, who produce a ‘formal’ citation form of words in informal situations, where a natural flow of connected speech would be more appropriate.

If students are surrounded by and educated in this kind of citation and formal language environment, when they are involved in real communications with L1 speakers, they may find it very difficult to understand the conversation. In real-life communications, L1 speakers use the language simply to ‘get on with living’; they speak only clearly enough to make themselves understood in a particular context (Brown, 1990, p.2). When L1 interlocutors communicate, they only pay part attention to the incoming signals; mostly the listener uses the clues sent by the speaker to abstract the main ‘gist’, and at the same time prepare his reaction in his mind (in the present work, ‘his’ is to be interpreted as ‘his’ or ‘her’; similarly ‘he’ should be interpreted as ‘he’ or ‘she’). L1 speakers ask for clarification only when intelligibility breaks down. Most of the time they use their shared experience to compensate for unclear clues.

The discrepancy in learning and use of natural English speech between L1 speakers and non-L1 language learners is also reflected in the use or non-use of certain words, such as dialogic markers, in informal English conversations, as shown in the following investigations. A study carried out by Firth (1988, cited in McCarthy, 1991, pp.49-50) displays the different distribution of ‘reason’ markers in the speech of L1 English speakers and Danish learners of English. ‘Because’ is exclusively used by L2 learners, while ‘ ’cos’, ‘like’, ‘see’, etc., are variously employed by L1 speakers. These discourse markers unconsciously used by L1 speakers, as shown by Watts (1989), do seem to be one of the innate characteristics in L1 speakers’ speech. Lack of this natural routine makes the speech sound unnatural and incomprehensible (Tyler, et al., 1988).

It is the mismatch in perception and use of natural English speech between L1 speakers and EFL learners that motivates many scholars to explore a more efficient second language learning and teaching approach; and that is also the motivation of the current study.

1.3 Aims of the Thesis

Due to the considerable proportion of formulaic language in native English speech, the perception and acquisition of the natural use of formulaic language as used by L1 speakers is important for EFL learners. In the literature there is much analysis of frequently used formulaic language, as detailed in Sections 2.3.3 and 2.4. Most studies concentrate on its forms and functions, and some on its phonological makeup. However, there is no evidence based on real, dynamic dialogue to show a correlation between different categories of formulaic language and their relevant phonological characteristics. Therefore, one of the objectives of the present study is to demonstrate

the relevance of speed of delivery and pitch range to the realisation of formulaic language.

Given the gap between internalised, formal, written forms which language learners are exposed to and natural, authentic, dialogic English speech as acquired and used by L1 speakers, as outlined in Section 1.2, appropriate EFL pedagogy has become a fruitful area for investigation. Can innovative speech learning technologies, specifically the slow-down algorithm and the early assets of the DSC (Dynamic Speech Corpus) which is being produced in DIT, both available to the present author and both based on insights into the relevance of pronunciation and intonation at supra-segmental level, help bridge the gap and facilitate EFL learning and teaching activities, particularly in the learning and acquisition of formulaic language? This is the second objective of the current study.

The present research addresses the following research questions in depth:

RQ1: What influence do speed of delivery and pitch range have on the categorisation of formulaic language?

RQ2: Does the slow-down facility, coupled with suitable training materials, improve Chinese EFL learners' ability to perceive and produce formulaic language with NS-like quality?

The following, related, secondary research questions are also addressed:

RQ3: Can real natural English speech be incorporated into EFL pedagogy?

RQ4: How can EFL learning be evaluated in the context of natural speech?

RQ5: Can assets from the DIT Dynamic Speech Corpus improve EFL learners' facility with L1 speech?

The present study argues that, based on the analysis of speed of delivery and pitch range, formulaic sequences are not all equal. The same sequence of formulaic language, spoken at different speeds and displaying different prosodic features, produces different pragmatic outcomes. An analysis of various categories of formulaic language and their different phonological realisations is proposed in this thesis as a starting point for further research (see Chapter 7).

In order to demonstrate the validity of the slow-down algorithm in helping Chinese language learners produce native-like English speech, particularly formulaic language, tests and training sessions were designed and carried out from June 2007 to May 2008 (see Chapter 8). The considerable improvement in the Test Group performance demonstrates the high acceptability of slow-down technology in helping the participants to capture the messy acoustic ‘blur’ in the rapid flow of NS and to follow the intonation patterns, so as to facilitate intelligibility of L1 English speech.

The current study produces several original contributions to the field of formulaic language, and EFL learning and teaching by:

1. Filling gaps in current literature on the relationship between formulaicity and prosody. Different phonological realisations, e.g., speed of delivery and pitch range, and various categories of formulaic sequences are investigated and analysed in this study, which enriches the understanding and use of the communicative, pragmatic functions of formulaic language. (RQ1)
2. Clear demonstration of the effectiveness and acceptability of the slow-down facility, which can be applied not only to citation forms, but also to assist

learners imitating and re-producing the native acoustic ‘blur’ and intonation patterns of formulaic sequences. (RQ2)

3. Implications of the study for a revised EFL pedagogy, specifically for Chinese learners and teachers, by exposing students to real speech samples occurring in everyday life in conjunction with language learning technologies, rather than maintaining loyalty to the internalised, ideal citation forms which cause problems when learners are exposed to real target language L1-L1communication. (RQ3)
4. Development of an evaluation methodology for testing EFL speech production, which, rather than being based solely on the judgement of citation forms, also incorporates an assessment of the messy ‘blur’ of rapid speech and communicative intonation patterns. The test results obtained from 100 participants from different levels and various language learning environments provide a reliable body of data for qualitative and quantitative analysis. (RQ4)
5. Investigation of an application of real, interactive spoken English assets (currently being developed for DIT’s Dynamic Speech Corpus), which exposes language learners to authentic, natural English speech, and makes language learning advance in a real, contextualised environment, so as to enhance their EFL study, particularly of L1 prosody. (RQ5)

These contributions form the basis of future investigation into effective perception and acquisition of natural English speech for EFL learners, and as such represent novel work in this field.

1.4 Contents of the Thesis

The work of this thesis is contained within the following chapters:

Chapter 2 focuses on the review of spoken English and formulaic language. Firstly aspects of production and decoding of English speech are dealt with. The natural perceptive process of authentic English speech is also examined. Then the learning and acquisition of formulaic language is considered.

Chapter 3 concentrates on the review of spoken Chinese and the conditions under which English learning and teaching occur in China. A review of the unique phonetic and phonological features of Chinese, in contrast with spoken English, is given first. An overview of how English is learned and taught in China and what problems Chinese learners encounter when involved in a real English speaking community follows.

Chapter 4 discusses the literature review of Chapters 2 and 3. Some issues in current research are outlined: the east-west prosody divide, formulaic language and its phonological realisations, the need to expose Chinese EFL learners to authentic spoken English, and some pedagogical considerations in the use of technology in language learning.

Chapter 5 first describes an overview of current, conventional TELL tools. Then a description of the TELL tools developed by DIT researchers is given, which might enhance language learning and teaching so as to improve global English communication.

Chapter 6 presents research design, based on the above review and discussion. The scope of the current research is identified, the rationale for designing the study and the procedures and methodology for its implementation are described in detail.

Chapter 7 details the innovative contributions of the current research in the area of formulaicity and its phonological realisation of speed of delivery and pitch range, based on the transcribing, segmenting and tagging of real, natural, dynamic speech sequences.

Chapter 8 describes the application of the innovative speech technologies to the acquisition of NS-like production of formulaic sequences by Chinese learners of English. This chapter introduces the application of speech technologies to the tests and training sessions of 100 Chinese EFL students. An evaluation procedure to test the effectiveness of the slow-down facility is designed and carried out. In addition, a panel of L1 speakers was used to validate the methodology employed. Test evidence demonstrates the effectiveness of the training materials and the assessment procedures, as well as a high degree of student acceptance of the technology employed.

Chapter 9 discusses the present research and its contribution to the body of knowledge with reference to the literature and technology reviews in Chapters 2, 3 and 5.

Chapter 10 summarises all the work undertaken during the preparation of the thesis. A statement of original contributions is included. The chapter closes with reflections on the current study and suggestions for further work to develop these contributions.

2. Literature Review

2.1 Introduction

The previous chapter gave an overview of the present study. The descriptive, corpus-based approach mentioned informs the work of this thesis. When studying the literature on L1 speakers and EFL learners in natural, informal English communication, a gap became evident, which motivated the present author to explore a solution to bridging this gap using the language learning technologies being developed in DIT.

This chapter reviews spoken English and formulaic language: aspects of production and decoding of spoken English (Section 2.2), the natural perceptive processes of authentic English speech (Section 2.3), and the learning and acquisition of formulaic language (Section 2.4).

2.2 Aspects of Production and Decoding of Spoken English

This section on aspects of production and decoding of spoken English looks first, in 2.2.1, at the use of significant phonological features in signalling potential speaker attitudes. It considers stress, rhythm, pause, speech rate and intonation patterns. 2.2.2 then examines various models of intonation units and their communicative features, such as the tone unit, the information unit, the intonation group, the intonational phrase, and the paratone and the TM intonation model. The difference between natural, flowing English speech and its citation form is then reviewed in 2.2.3. Finally 2.2.4 deals with the different decoding processes of English by L1 and L2 users.

2.2.1 Salient Phonological Characteristics Signal Potential Attitudes in Oral English

As Field (2003a, p.126) points out, an additional resource which a listener has is prosody. Prosody refers to the variations in pitch, loudness, tempo and rhythm of speech, which covers the aspects of stress, pausing, speech rate and intonation. These supra-segmental phonological elements provide cues as to syntactic structure. They also provide an information focus, indicate contrast and emphasis and provide affective signals. In addition, Brazil (1997) also suggests they reflect the state of shared knowledge between speaker and listener, and he defines various tone patterns within a tone unit, as detailed in Section 2.2.2. Tatham and Morton (2004) also emphasise that, apart from the linguistic functions, prosody also conveys expressive content (p.295).

Ladefoged's (2001, p.15) view of the supra-segmental features is that they are 'characterized by the fact that they must be described in relation to other items in the same utterance'. This traditional model of prosody fits supra-segmentals to strings of syllables. Tatham and Morton (2006, p.123, p.130), on the other hand, argue that prosody has an independent existence and that prosodic effects 'span more than one individual speech segment'. Within this model 'syllable strings fit to an existing prosodic structure' which is hierarchical, rather than linear.

Some of the most important elements of prosody are looked at in the following paragraphs.

A. Stress

The term stress normally includes word stress and sentence stress. Word stress, also called lexical stress, is assumed to be part of a word's identity. Word stress itself is relatively stable and invariable, unless it is produced in a particular fashion within an utterance. Therefore, the main focus in the current study is on sentence stress, since it is of one of the supra-segmental features.

In an utterance, the speaker realises his communicative plan with stresses which emphasise his communicative intention. The listener's task is to perceive and correctly assign the planned stresses intended by the speaker. Brazil (1995) points out that, when the speaker chooses one of the words to carry the prominent stress, he is actually telling his listener that this word 'represents an existential sense selection' (p.241). Whether a word is selective or not depends on the particular context in which the actual speech is uttered (Brazil, 1994, p.86). In his analysis of 'tone unit' (as detailed in Section 2.2.2), Brazil (1997, p.9) also states that the last prominent syllable in a tone unit is defined as the 'tonic syllable', which carries the tonal contour and indicates the speaker's communicative intention.

Tatham and Morton (2006, pp.139-40) also point out that different listeners with their different internal and external perceptual 'environment[s]' assign different stress or prominence to the same speech signal.

B. Rhythm

The alternation of stressed and unstressed syllables forms a special rhythm in speech patterns of a category of languages (stress-timed languages), to which English is traditionally assigned. In this category there is a tendency to show approximately equal

intervals between stressed syllables with unstressed syllables fitting into the intervals. Knowles (1987, p.144) traces its long history back to Joshua Steele (1775), who insists that the rhythm of speech can be compared with rhythm in music. Mortimer (1985, p.138) claims that this perceptual rhythm ‘exists before the speaker speaks’.

However, some linguists offer strong counter-arguments on the irrhythmicality of spoken English. Roach (1982, p.78) points out that the impression which influences people’s subjective judgements might be on account of ‘vowel reduction in unstressed syllable[s]’. Indeed, the degree of rhythmicality can vary in the same speaker in real-time interaction. As Cauldwell points out, ‘spontaneous speech is functionally irrhythmic’ (2001, p.1); it is shaped by ‘the speaker’s choice’ (2005, p.2). Absolute, regularly-timed rhythmicality is an abstraction, without regard to individual speaker performance (Hawkins, 1984, p.178). Crystal (1996, p.8) goes so far as to state that it is unlikely for all the languages to ‘fall neatly into any two types’ (stress-timed and syllable-timed). Thus we can understand why Brazil (1996, p.9) expresses a negative opinion on the teaching of rhythm in language class, especially from a discourse point of view. A similar opinion is expressed by McCarthy (1991, p.92). Therefore, an indecisive conclusion is arrived at by Marks (1999, p.198) who asserts that, even though the evidence of stress-timing pattern is unproven, the potential tendency towards this rhythmicality is a ‘deep, inherent element of language users’ competence’.

C. Pause

With respect to rhythm, one aspect should also be mentioned, namely pause. Goldman-Eisler (1968, p.12) distinguishes three types of pause within the speech of individual speakers: articulatory pause, hesitation pause, and pause for breath. Chafe (1979, p.162)

also discusses ‘hesitation pause[s]’, where the speaker encounters momentary encoding process problems. According to Meyer, et al. (1980), pauses are divided into filled pauses and silent pauses. Filled pauses are either linguistic, i.e., ‘you know’, ‘well’, or non-linguistic, e.g., ‘er’, ‘mm’. Pauses normally coincide with syntactic units, i.e., clause boundaries or sentence boundaries.

Deliberate pauses have important discourse functions – either indicating that one has finished one’s turn and the interlocutor can take the floor, or be aware that more important communications are yet to come. Brown (1990, pp.48-51) outlines positive functions of pauses from the point of view of the rhythm of English speech. She gives an example to show how a brief pause and a nod performed by the speaker replaces the stressed syllable ‘thank’ in ‘thank you’, which ‘contributes to our perception of a fluent, rhythmical flow of language’. A study by Beattie (1983) also shows that a hesitant phase in speech is actually followed by a fluent phase. Pausing, at first glance, seems to cause disfluency; however, it can benefit both speakers and listeners.

D. Speech rate

‘Rate’ is the overall tempo of speaking. Based on Laver’s (1994, p.158) definition, there are two different terms indicating the delivery speed of speech. One is ‘articulation rate’ – the rate at which the actual utterance is produced by a speaker, excluding silent pauses. The other is ‘speaking rate’ – the rate at which the whole speaking-turn is produced, including all speech material and silent pauses. When a speaker is perceived to be a fast talker, it is not because he speaks at a fast articulation rate, but rather due to his fast speaking rate – fewer or shorter pauses are produced within the utterance (Field, 2004, p.273).

Tempo varies between speakers, different speaking environments and different types of speech, and it also varies within the speech of an individual. There is no clear ceiling to a speaker's speech rate; however, a rapid or slow tempo, compared with the speaker's average rate, is marked for attitude. Several functions are recognised in Deese's (1984) study. The first function correlates with turn-taking, in which there is an increase in speed near the end of a turn to prevent the interlocutor cutting in, since the speaker wants to 'keep the floor'. Another function relates to the manner of expression, in which a 'modest', 'non-assertive' attitude is expressed. Deese points out that this function always accompanies certain intonation patterns.

E. Intonation patterns

It is recognised by Roach (2000) that intonation, which, however, has not yet been satisfactorily defined (p.150), is strongly involved in pitch patterns. The English language is categorised as an intonation language (ibid., p.162), in which pitch value is primarily used for conveying syntactic and contextual meanings. Intonation is perhaps the most complex of the prosodic phenomena in speech production, in which some physical elements cause the change of intonation patterns. Everyone has an individual pitch span. Any extension or restriction in pitch range, together with changes in intensity (greater or lesser volume), can indicate the presence of a particular attitude or syntactic category. The same applies to pitch range on its own. Lowering or raising the pitch can indicate a non-neutral attitude and low termination can also be a turn-taking cue (Brazil, 1997, pp.88-92).

The most dominant property of connected speech in intonational languages, as Levelt (1989, p.307) points out, is its melody. The intonation of an individual word depends on which syllable carries the word accent (word stress in citation form), whether the word is selective or not (whether it carries the pitch accent – the main pitch movement), and the melody of the intonational phrase in which the word is embedded. Levelt also states that the intonation patterns are likely to be imposed upon an utterance, which could happen ‘at a late stage’, immediately ‘before articulation’, which later is called, by Field (2004, p.50), ‘mental buffer’. Field (2003a, p.81) also states that this intonation assignment stage may occur ‘at the same time as the building of a syntactic frame’. However, Tatham and Morton (2004, p.302) argue the opposite. They state that the expressive pattern is a ‘carrier’ of the message, which comes first and lasts longer than the message itself.

Traditionally, the normal ‘unmarked’ pitch in English is falling, and the raising of pitch demands a response by the interlocutor. A discourse-based approach to intonation choices is proposed by Brazil (1997), and also investigated by McCarthy (2001, p.65), in which the communication value is taken into consideration in the interaction between speakers and listeners. Intonation is not dependent on syntactic or grammatical structures, but is rather a speaker’s choice to decide how to package the information. Appropriate intonation patterns can, however, guide the listener as to the syntactic structure of the utterance. Intonation is also a potential meaning-carrier when it expresses a speaker’s affections and attitudes, and also when it directs the listener to interpret the information so as to achieve convergence between speakers and listeners. This point is supported by Selman’s (2009) personal experience of using correct intonation to improve his communication while learning Japanese. Appropriate encoding and rendering of intonation patterns depends on not only the conventions of

pitch direction, but also on shared knowledge between the interlocutors, and the dynamic contexts of the interaction. This discourse-based approach is adopted by the present study.

2.2.2 Intonation Units and Communicative Features

All spoken language is composed of different sized ‘chunks’. These speech units have been focused on by different researchers and are amenable to different types of linguistic analysis. The investigation in this section starts with some short intonational units which are characterised by having a tonic stress within each unit, i.e., the tone unit, the information unit, the intonation group and the intonational phrase. In addition, some other longer units which consist of more than one prosodic pattern, such as the paratone and the TM intonation model, are also outlined.

A. Units characterised by presence of tonic stress

a) Tone Unit

For analysing spoken language and its communicative value, the minimal intonational unit – the tone unit, is proposed and systematically analysed by Brazil (1997). The tone unit is ‘the stretch of language that carries the systemically-opposed features of intonation’ (p.3), which is similar to ‘sense groups’, ‘breath groups’ and ‘tone groups’ (p.5). According to Brazil, there is only one indispensable tonic syllable (which carries the operation of salient pitch movement) within each tone unit. This is located on the last prominent syllable, although the criteria for assigning the tonic syllable ultimately depend on speaker decisions.

Depending on the communicative event, either a P tone (proclaiming – providing new information) or an R tone (referring – indicating shared information) is assigned by the

interlocutors (Brazil, 1997, pp.68-73). Brazil defines the values of the P/R tone choices, from the social interactive point of view. In general, a proclaiming tone shows a distance between the speaker and the listener, and a referring tone normally expresses a co-operation between the interlocutors (p.82). Brazil emphasises that the P/R tone choices depend, among other things, on whether the speaker wants the speech to benefit himself or his listener (p.95).

Pause phenomena also play an important role in orientating a discourse and are defined as ‘dummy carriers’ of the tone (Brazil, 1997, p.139). Given the imperfect nature of tone units (i.e., the ‘incomplete tone unit’), Brazil acknowledges that ‘pauses are always treated as tone unit boundaries’ (pp.147-48), and the boundary of the tone unit often coincides with a grammatical unit, i.e., the sentence or the clause (p.150).

The tone unit is regarded as a ‘communicatively functioning whole’ (Brazil, 1997, pp.18-19), and the communicative value of intonation, as Brazil emphasises (1995, p.240), is associated with the purpose in which it occurs in a certain here-and-now context of interaction.

This stretch of sound continuum – the tone group – is also analysed by Brown (1990). She tends to break a stretch of speech by ‘immediate constituent[s]’ (p.93). The boundary most likely occurs between subject and predicate, and might be found within a long subject or a long predicate. The criteria for this division are not only pitch, but also lengthening of the final syllable and pause. In general, in spontaneous speech, the division into tone groups is less clear, but at least the tone group delimits major syntactic constituents (p.102). While Brown also locates the tonic syllable on the last

lexical word, she points out the phenomenon of tonic shift to the left. This would suggest a speaker-dictated prioritisation of discourse function over syntax.

Roach (2000) also discusses tone units from the perspective of the analysis of intonation patterns. He points out that the smallest tone unit can be of only one syllable. The tonic syllable carries the pitch movement of the tone unit, yet, when there is a tail following, the pitch movement will be extended (pp.167-73). In addition to silent pauses, ‘sudden change[s] of pitch level’ and rhythm are also defined for the identification of tone unit boundaries (p.178). He also agrees that usually tone units accord with grammatical units. A more generalised view on the position of the tonic syllable is given by Roach (ibid.) who claims that it ‘tends’ to occur on the last lexical item; on a few occasions it can be earlier due to the speaker’s decision on prominence (pp.194-95).

b) Information Unit

As mentioned above, an important function of intonation is to mark off new information from given information, thus an ‘information unit’ is proposed by Halliday (1967). The information unit is directly associated with the realisation of a ‘tone group’. Each tone group ‘represents what the speaker decides to make into one unit of information’ (Halliday, 1970, p.162). The prominent syllable in this group, i.e., that which carries the ‘main burden of the pitch movement’, is called the tonic syllable (ibid., p.4). There is one and only one tonic syllable within each tone group. According to Halliday, each tone group must begin with a stressed syllable, or a ‘silent ictus’ when the initial foot is unstressed. Although no criteria are given for the division of tone groups, Halliday (1967, p.203) emphasises that the tone group, due to its phonological characteristics, must occur within the limitations of rhythm. Since information is mostly ‘mapped on to

the clause' (1970, p.127), the clause is posited as the common unit of the information group.

The tone group is also analysed by Laver (1970, pp.68-69), from the speech production point of view, in which it is stated that it is 'handled in the central nervous system as a unitary behavioural act'. The average length for a tone group is about seven or eight syllables. The tonic syllable, according to Laver, is usually located 'at or near the end of the tone group'. The pause is seen as the primary tone group boundary. The syntactic clause often coincides with the tone group. In agreement with Halliday, Laver also states that the tone group corresponds with a rhythm unit.

c) Intonation Group

Another type of unit – the 'intonation group' – is proposed by Cruttenden (1997). In his analysis, four-levels of stress patterns are present within an intonation group: primary stress, secondary stress, tertiary stress and unstressed (p.18). Given the unclear delineation of an intonation group (apart from pauses), three other external criteria for the identification of intonation groups are proposed by Cruttenden, i.e., an anacrusis, a lengthening of the final syllable, and the change of pitch contour on an unaccented syllable (pp.32-34). Cruttenden points out the problems encountered when using pauses as the criterion, and when there is no clear pitch movement to indicate tonic stress. This indicates that the intonation group is in fact an abstraction, 'a theoretical construct' (p.37). While Cruttenden uses 'nucleus' to describe the tonic stress within an intonation group, he also argues that the contour following the nuclear accent always expresses the most salient intonational meaning of an intonation group (pp.44-49).

Intonation groups more commonly align with large syntactic constituents – more commonly clauses than any other grammatical unit (Cruttenden, 1997, pp.69-73). However, due to speaker-determined performance, the division of an intonation group is unpredictable. According to Cruttenden (*ibid.*), normally the average length for an intonation group is about five words, with very few groups over seven words (p.72). Given that the focus of the intonation group is the tonic syllable, Cruttenden argues that there are quite a lot of exceptions to its final lexical item placement (p.75). The reason for assigning and for the interpretation of different tone choices depends firstly on the syntactic type, and also it depends on the relationship between the speaker and the listener (p.104). Cultural factors are also emphasised, in that a high pitch register seems to be associated with certain cultural conventions (p.124). Like Brazil, Cruttenden labels the contrastive, confirmative information produced by falling intonation as ‘closed’, and the non-assertive, continuative associated with rising tone patterns as ‘open’ (p.163).

d) Intonational Phrase

Another term, defined by Levelt (1989) from the melodic point of view, is the intonational phrase. There is only one nucleus in each intonational phrase, and when there is only one pitch accent (there could be more), then the main pitch contour rests on the last lexical head. According to Levelt, an intonational phrase consists of ‘one or more phonological phrases or metrical groups’; ideally it spans about 2 seconds and ranges between 1 and 3 seconds (p.386). An intonational phrase is usually a sentence unit surrounded by grammatical pauses (usually more than 200 milliseconds). It can also be isolated by syntactic, semantic, or ‘operational’ definitions, as long as it is under the speaker’s control and carries one of a set of tones. Two different functions in one

intonational phrase are posited by Levelt – a qualifying and modifying function on the pre-nuclear part, and an intonational focus function (expressing meaning) on the nuclear pitch movement of an intonational phrase.

B. Units with more than one prosodic pattern

All the analysis given above is focused on the identification of the pitch movement and intonation functions within one intonational unit. Given the review earlier, many factors contribute to the isolation of intonational units, and the criteria are so complex and variable that in no case can the boundaries of intonational units easily be identified. Therefore, not all linguists agree how natural flowing speech can necessarily be divided into small intonational units, especially in the rapid continuum of informal speech. Some researchers abandon intonational units, and adopt different units for their different analytical purposes, as examined below.

a) Paratone

The ‘paratone’ is proposed by Brown and Yule (1983b). In order to analyse how speakers organise larger chunks of discourse and how the topic is smoothly changed between the interlocutors in natural conversations, Brown and Yule divide large pieces of discourse up into a series of small structural units, each on a separate topic, called paratones (p.100). Paratones, as Brown and Yule state, function as ‘spoken paragraph[s]’, and mark the beginning of new topics. The tone patterns in these units tend to display similar contours. A high pitch value usually occurs at the onset in order to draw the listener’s attention and mark the start of the turn. Brief pauses tend to be embedded in the middle of the paratone, and a gradual declination in pitch level or a long pause towards the end of the paratone indicate the closing remarks and a readiness

to hand over the turn. The paratone is more appropriately associated with topic structure, rather than with small individual pitch features.

b) TM Intonation Model

Another unit, called the TM intonation model, is employed by Tatham and Morton (2005) in the analysis of speech synthesis. In this unit four aspects are marked: lexical stress, a syntactic phrase within an intonational phrase, sentence and intonational phrase boundaries, and sentence focus on a single syllable. Only two levels of each prosodic characteristic are identified in the TM intonational model, which is different from Cruttenden's (1997) four levels of stress, as mentioned earlier. Based on Tatham and Morton's speech production theory, expression is a 'central' or 'enveloping' characteristic of utterances (p.112). In this model, they focus on identifying intonational functions by fitting the segments into a prosodic wrapper (p.287).

All the above listed different units identified by various researchers are based on the specific purposes of their linguistic analysis, and all of them are related to pitch movement and fit the prosodic intonation contours of English. However, there is a lack of agreement among linguists on the most useful unit of analysis in speech. In order to analyse the phonological realisation of natural speech, a minimal production unit – 'flow unit' (Campbell, et al., 2006) – is proposed and adopted in the current study, as detailed in Section 7.2.2.

2.2.3 *Natural Flowing Connected Speech and Citation Form*

This sub-section deals with the difference between the dynamic flow of English speech produced by L1 speakers and its citation form which is often produced by non-L1 language learners. It starts with the review of co-articulation and assimilatory processes

in the natural flow of connected speech, is then followed by the idealised citation form of spoken English produced by L2/EFL speakers. Finally it considers EFL speech and pedagogy.

‘I believe I need to learn what the word sounds like when it is used in the sentence. Because sometimes when a familiar word is used in a sentence, I couldn't catch it. Maybe it changes somewhere when it is used in a sentence’ (Goh, 1997, p.366). This is a problem often encountered by non-L1 language learners in stream-like English connected speech, which also uncovers the mismatch between the sounds produced in natural conversational speech by L1 speakers and those uttered in the language class by language teachers (most of them L2/EFL learners themselves). Since English speech is, as Steele (2005, p.1) states, ‘a continuous stream of sounds, without clear-cut borderlines between each word’, these ‘extremely messy products’ (Lass, 1984, p.298) of the stream of speech often cause failure on the part of EFL learners to process natural connected English speech.

A. The co-articulation process

In the dynamic flow of speech there are no clear-cut segments. As Levelt (1989) puts it, the main function of phonological encoding is to ‘prepare for fluent connected articulation’ (p.364). Another reason is that the speaker produces utterances based on a certain prosodic plan. All syllables within the utterance are chunked into smaller or larger rhythmic segments, realised by intonational phrases, as considered earlier in Section 2.2.2. The rhythm gives shape to the stressed and unstressed syllable structures and the duration of pauses. It transcends the phonetic level of individual sounds and covers the prosody of the whole utterance. The overall speech rate is regarded as the

main factor affecting the phonological encoding process. All segments are influenced differently (Tatham & Morton, 2006, p.145). If an utterance is delivered rapidly, not all the segments are reduced on the same scale. Vowels are normally more reduced than consonants. The sounds uttered in connected speech often do not conform to the citation forms produced in slow, careful speech.

This process of re-constructing the syllables and segments depends mainly on the context in which prosody occurs and has been described as the product of a ‘prosody generator’ – a processing component that consists of metrical and intonational properties (Levelt, 1989, p.406). The prosody generator interacts with the spellout procedure, in which words are slurred together, and the mechanism of production helps to create an acoustic flow of connected speech (p.410).

Kenworthy (1987), among others such as Brown (1990), also points out that, ‘no sound is an island’ (p.70). The acoustic signal is a continuum in which each realisation of each sequential phoneme flows into the next. Each phoneme is realised as various allophones in different phonetic environments, and the articulations of adjacent sounds are usually overlapped in the flow of speech so as to facilitate the production of the sound sequence. This kind of adaptation to surroundings is called co-articulation, and it spreads out from the syllable nucleus and provides the continuity of the speech production. Influenced by this process, it is normally not easy to separate individual sounds in the flow of natural speech, even within one single word.

B. The assimilatory process

There is another process, called the assimilatory process, which also occurs in connected English speech. The assimilatory process is often used to achieve a smooth speech production and encompasses assimilation, elision and catenation, etc. Dalton and Seidlhofer (1994, pp.24-31) give a detailed analysis of these phenomena. Basically, assimilation means the final consonant of a word is changed so as to ease the pronunciation of the initial part of the following word. Elision is the process of some sounds being omitted. Catenation refers to the phenomenon of sounds linking or merging across words. Confined by the stressed and unstressed rhythm in English (controlled by speaker's decisions), there are also some other modifications which occur in this production process, i.e., contractions and weak forms. Kenworthy (1987, pp.82-84) and Roach (2000, pp.112-13) respectively detail some rules for these phenomena.

C. Natural flow of L1 speech, and EFL speech and pedagogy

When L1 speakers become involved in fluent speech, both co-articulation and assimilatory processes are employed naturally. These processes are classified as features of rapid flow speech. A sample below (Campbell, et al., 2008, see Figure 1) clearly illustrates the different realisations of the phrase 'come here' between their idealised phonological forms uttered in isolation – the left part of the spectrogram, and the development of natural acoustic blur out of what is citation clarity realised in rapid flow connected speech on the right of the spectrogram.

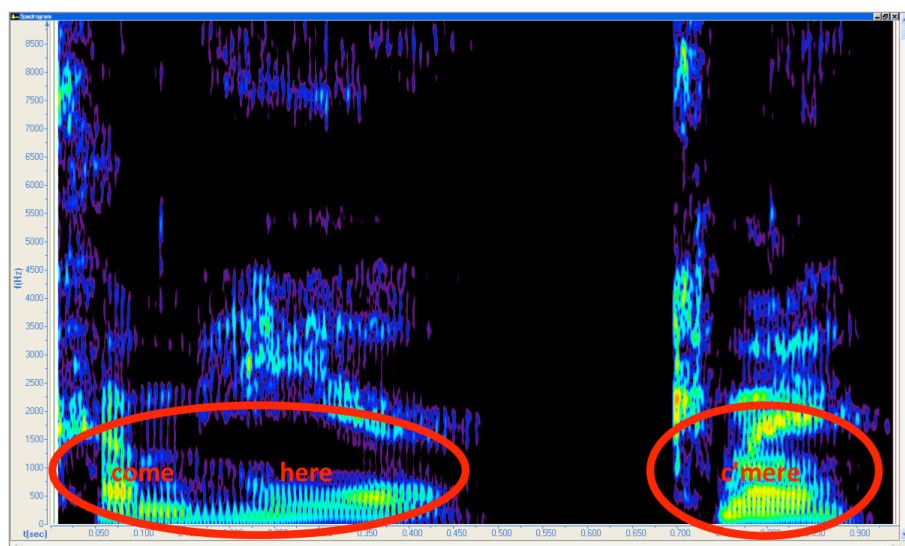


Figure 1: Idealised phonological form vs. natural acoustic blur produced by L1 English speaker

Source: Campbell, et al., 2008

Assigning word boundaries and recognising sounds which are simplified and altered in rapid speech is not an easy task for non-L1 language learners. Many Chinese EFL students complain that westerners speak too fast and consequently that they cannot recognise the words which could otherwise be easily retrieved in citation written form. Even advanced language learners suffer from this problem to some extent. Everyday English, as Cauldwell (2002, p.8) expresses it, is ‘messy’. ‘Connected speech is not just the sum of its individual words’ (Underhill, 1994, p.58). It involves a series of co-ordinated processes in which citation forms of sounds are connected and modified so as to form a smooth stream of speech. This is one of the most striking differences between slow, formal speech and casual, connected speech. It is the listener who segments the continuous stream of speech and assigns proper phonological labels, based on the segmental representation constructed in his mind (Tatham & Morton, 2006, p.14). This leads to a discrepancy between the learners’ idealised forms of individual words which he has internalised and the stream of non-segmental words which he is listening to.

Cauldwell (2003) later on supports his idea and pinpoints the reason why EFL learners often cannot retrieve the sounds in natural, informal English speech. In a classroom situation, when a teacher teaches students the pronunciation of isolated words, what the students learn is just the citation forms of the pronunciations, and these standard pronunciations are not the same as real English spoken in context. As Cauldwell (*ibid.*) puts it, they are the ‘misrepresentations of the essential stream-like nature of all speech’, and they are ‘obstacles to improvement of their [language learners’] listening skills’ (p.5). In classroom-based teaching activities, teachers typically present the ‘perfect’ citation forms of the English words – with pauses providing a clear word boundary, and indicating stresses and tone patterns. However, in the real world, the communication which language learners experience is a messy, non-segmented stream-like flow in which most of the sounds are unrecognisable and ‘pulled out of shape’ (p.1). Jenkins (1998a, p.43) also states that it is particularly those assimilatory items which actually ‘detract from intelligibility’ for EFL listeners and easily get them ‘lost’ in connected NS.

Given the features of English connected speech, many scholars point out that making non-L1 language learners become aware of and equipping them with the ability to cope with natural spoken English becomes pedagogically necessary. Language learners should at least be given a chance to study these features, and compare the different forms produced in a deliberately produced teaching environment and their phonetic realisation in normal informal speech. This will help teach them to compensate for gaps in word identification. As Field (2003b, p.331) puts it, ‘[a]wareness of this kind of feature can aid learners in producing these clusters, as well as recognizing what has been omitted’.

2.2.4 Different Decoding Processes of English Speech by L1 and L2 Users

This sub-section deals with different decoding processes of English NS employed by L1 users and L2 learners. It is structured as follows: L1 speakers' decoding process which involves bottom-up and top-down processing and making intelligent guesses; L2/EFL inefficient decoding process; nature of L1 decoding, EFL decoding and listening strategies.

A. L1 speakers' decoding process

Receiving incoming acoustic signals, assigning correct labels to the segments and arriving at reasonable interpretations: this is the process called decoding. L1 language users can generally cope with the decoding process automatically.

Brown (1990) gives a thorough description of the natural decoding process of L1 speakers. Recognising the main message given by the prominent units (either the stressed words, or the salient information chunks), L1 listeners reconstruct the information and then achieve an intelligible interpretation. This process involves getting the phonological information from the acoustic signals, narrowing the content down to a particular topic and correctly detecting the words used by the speaker, and then arriving at a correct interpretation.

a) bottom-up processing of speech

The first step – recognition and assigning of phonological code – is called 'bottom up' processing (Brown, 1990, pp.10ff, pp.150ff). This is an essential process for listening. The more phonological information the listener gets, the better he grasps the topic being

spoken of. This does not mean that the listener has to understand every single word so as to get the whole conversation. In fact, in listening, listeners normally pay more attention to ‘what’ is said, rather than ‘how’ it is said (Brazil, 1994, p.2), listeners do not perceive spoken language as a series of sounds, instead, they capture the gist of the communication. As Roach (2000, p.130) puts it: ‘when children are learning their first language, they acquire features rather than individual phonemes’. The salient phonological characteristics, as examined earlier, mark significant communicative information in everyday spoken English. Having internalised these native features of English speech, L1 listeners can easily retrieve sufficient phonological information and build them stepwise into a representation going from smaller units to larger ones, then to the whole utterance. Having acquired the basic information input can then lead to the step of ‘top down’ processing.

b) top-down processing of speech

‘Top down’ (Brown, 1990, pp.11-12, pp.151-52) processing means that, after getting sufficient contextual information on the topic in the ‘bottom up’ process, intelligent predictions are made. This process is dependent on listeners’ personal, formalised experience both of language and the world. In NS communication, people tend to use a large amount of ‘prefabricated’ language (reviewed later in Sections 2.3.3 and 2.4). In particular, discourse markers, as Chaudron and Richards (1986) discuss, function as a signpost for facilitating top-down processing. Apart from these clear signposts, due to the stream of English connected speech, it is impossible to pick up every single phoneme. The listener uses his personal language experience to compensate for any gaps in recognising the missed words. In addition, this process also involves completion of the interpretation by bringing in the listener’s world knowledge. As de Beaugrande

(1980, p.30) states, ‘the question of how people know what is going on in a text is a special case of the question of how people know what is going on in the world at all’. Armed with this set of stereotypical knowledge, L1 language users can generally predict what might be talked about in this situation, and what might be discussed by this specific speaker. This familiar, shared knowledge is acquired by L1 users since infancy, and is regarded as one of the automatic skills of L1 language users (Brown, *ibid.*, pp.153-55).

c) making inferences in speech

The process of ‘making inferences’ (Brown, 1990, pp.155-58) refers to getting extra information, which is not linguistically present in the specific expressions used by the speaker. This logical guessing operation depends on how relevant the extra information is in the context in which it occurs, and even depends on the interpretation of the ‘utterer’s beliefs and desires’ (Dennett, 1990, p.191).

In general, L1 language users are active listeners. They are primarily reliant on phonological cues which provide them with basic information to help them tune into the communication. Since in the rapid flow of informal speech, some phonetic clues are reduced, modified, or missing, L1 listeners do not depend on capturing all details of the speaker production; rather they only pay attention to the ‘shape’ of the word (McCarthy, 1990, p.35). L1 language listeners have the innate ability to make up for phonetic weakness or slips of the tongue (Boomer & Laver, 1968), and draw logical conclusions, and sometimes they can even finish the utterance for the speakers, which is termed ‘latch[ing]’ (Sacks, et al., 1974).

B. L2 listeners' decoding process

In contrast with L1 speakers, for non-L1 language learners, perceiving natural connected speech and achieving a correct rendering is not an easy task. Due to the classroom language learning environment (for most EFL learners in China), in which there is almost no exposure to authentic, natural, spoken English, when non-L1 language learners deal with real L1-L1 English communication, they are easily over-tasked in the decoding process, as investigated by the following researchers.

a) inefficient use of bottom-up processing

Research carried out by Tsui and Fullilove (1998) indicates that 'less-skilled listeners were more likely to process ... linguistic input without understanding the entire text' (p.447). This is a typical processing approach employed by most Chinese EFL listeners. When involved in informal native-to-native English conversations, typical L2/EFL listeners constantly struggle with scanning the incoming signals and looking for matches in their lexicon. However, there are some factors which impede L2 listeners in recognising every feature of linguistic input. Firstly, as considered in Section 2.2.3, given the nature of the natural flow of English speech, for L2 listeners, the problem is often that 'it is not clear to them how many words there are supposed to be in the utterance and where their boundaries might lie' (Brown, 1990, p.150). Another factor is, as pointed out by Field (2004), 'a limited vocabulary or grammar, or the inability to recognise known words in connected speech' (p.308).

b) insufficient use of top-down process and making inferences

The skills of narrowing down information and making intelligent guesses from the surrounding contexts are often inappropriately or insufficiently employed by L2/EFL

listeners. One of the problems is, as investigated by Aitchison (1994), the role of culture in the decoding of L2 language. Given that L2/EFL learners come from different social and cultural backgrounds, L2 listeners often lack sufficient 'familiar knowledge' (Brown, 1990, p.155) which can supply them with correct cultural information and facilitate their rendering of the L1 speakers' intention. Another factor which may also affect L2 listeners' decoding process is kinesic behaviour (i.e., body movement), as investigated by Kellerman (1992). Due to the lack of relevant L2 world knowledge, language learners often cannot narrow down the topic being talked about, although they can pick up some of the words in the utterance. They often do not feel confident enough to make logical inferences by employing their previous L1 experience, which will prevent them from achieving a degree of 'automaticity' in the way they decode L2 language (Field, 2004, p.308).

C. Nature of L1 decoding, EFL decoding and listening strategies

One point which needs to be clarified here is that a correct interpretation does not mean 100 percent identity between the speaker's intention and the listener's comprehension. As Brown (1990, p.10) emphasises, 'communication is a risky business'. Every listener has different personal experiences and world knowledge, which also triggers different interpretations of the speaker's intention. Even if there is agreement among the listeners as to what was said by the speaker, there might still be varying perceptions regarding the real intention behind the words the speaker used. L1 speakers, however, can normally minimize the potential misunderstanding, as they move towards maximal convergence in their communication (details reviewed in Section 2.3.2).

Given that listening is not a passive information-transmission process, how can one achieve Brown's (1990) goal for language learners – to 'listen as a native speaker listens' (p.148)? Cauldwell (2000, pp.2-3) emphasises that, 'it is a mistake' to abandon bottom up activities, since it demonstrates to language learners the essential characteristics of speech. In other words, from his point of view, language teachers have inherited a top down approach, and are of the opinion that learners do not need to understand every word. Cauldwell suggests that this seems illogical and unreasonable. The skill of the top down approach is 'a goal to be reached', rather than 'a means of getting there'. Language teachers should teach learners how to perceive and work towards an imitation of the L1 listeners' decoding style rather than teach them how to gain these abilities at an early stage. It seems that there needs to be some bottom up processing before the real top down skill can be achieved. Wilson (2003, p.335) has a similar idea and proposes a more plausible 'discovering listening' method, based on Marslen-Wilson's (1989) 'bottom-up primary' model, which improves language learners' performance by pinpointing their listening difficulties after re-constructing a text. In other words, language learners should 'spend more time with the signal' (Cauldwell, 2000) and study how the L1 speaker speaks, not just what the L1 speaker says.

Field (1999, pp.338-39) also discusses bottom up and top down approaches, and he points out that these two operations are actually processed interactively. Bottom-up information is the basis for narrowing down the range of possible predictions. Meanwhile, the contextual information gained by top down processing also influences or supports the basic phonological clues. A more efficient skills-based approach is advocated by Field, which shows the importance of conceptual and perceptual work in second language teaching activities. Teaching listening strategies is not 'a waste of

time’ (Ridgway, 2000, p.184), and Field (2000, p.194) argues, ‘[l]et us improve the lifebelts rather than relegate our swimmers to the paddling pool’. The present author agrees with Cauldwell’s and Field’s approaches of L2/EFL listening teaching.

2.2.5 Summary

Knowledge of aspects of production and decoding of spoken English, as reviewed in this section, is essential for non-L1 language learners to better understand English speech and to learn to naturally and automatically process spoken English in much the same way as L1 language users do. These important features are highlighted at the beginning of the analysis, whose aim is to pinpoint the essence of spoken English which is the basic barrier impeding language learners in overcoming their L1 influence and adopting native English language patterns.

Given that spoken language is a dynamic communication-transmission process, efficient processing of natural, authentic English speech for non-L1 learners involves some necessary aspects and procedures, i.e., exposure to natural speech, negotiated convergence of speaker and listener, and use of formulaic language, all of which are further outlined in the next section.

2.3 Perception of Authentic Natural Spoken English by L2/EFL Learners

This section deals with the perception of natural English speech by non-L1 learners. Firstly, in 2.3.1 it considers facilitating intelligibility by increasing exposure to authentic spoken English. 2.3.2 then examines issues of speaker and listener convergence, based on shared contextual knowledge. Finally, formulaic language in spoken English is briefly reviewed in 2.3.3.

2.3.1 Facilitating Intelligibility by Increasing Exposure to Authentic English Speech

This sub-section deals with issues of how to facilitate L2 learners' intelligibility by increasing their exposure to real, natural spoken English. It is structured as follows: various interpretations of 'intelligibility'; issues of intelligibility which include segmental, supra-segmental and accent elements; importance of exposure to natural English speech in L2 acquisition.

A. Various interpretations of 'intelligibility'

Intelligibility is a widely researched area in L2/EFL learning and teaching. Different scholars, however, have differing interpretations of intelligibility. A catch-all term is given by Bamgbose (1998, p.11), in which intelligibility means 'a complex of factors comprising recognising an expression, knowing its meaning, and knowing what that meaning signifies in the sociocultural context'. This interpretation corresponds to Smith and Nelson's (1985) terms 'intelligibility', 'comprehensibility' and 'interpretability'. Intelligibility here means recognition of a word or utterance; comprehensibility means getting the meaning of the word or utterance, while interpretability refers to perceiving and rendering the speaker's intention in word and utterance. Brown (1995, pp.10-11), in her analysis of speaker and listener communication, gives a simpler version – 'identification' and 'understanding' – in which 'identification' is similar to 'intelligibility', and 'understanding' seems to cover both 'comprehensibility' and 'interpretability' based on Smith and Nelson's definitions. It seems, therefore, that there is no general consensus on the definition of these terms. In this study, the present author adopts the position, also advocated by Smith (1992), that intelligibility, emphasising

word and utterance recognition, is the basic level to achieve comprehensibility and interpretability (Nelson, 1995, p.274).

B. Issues of intelligibility

a) segmental elements

Word and utterance recognition firstly involves decoding the sounds uttered by the speaker. As reviewed in Section 2.2.3, natural English speech is unbroken and connected, and sounds are not produced in isolation, all with equally clear pronunciation, but co-articulated. Most Chinese language learners are exposed to a classroom learning environment in which they get used to a model of artificial speech where every segment is clearly articulated. When language learners listen to the ‘untidy’ flow of natural English speech, as Brown (1990, p.60) emphasises, they are going to ‘experience a devastating diminution of phonetic information at the segmental level’. All familiar words seem to disappear, even in the case of advanced learners. The reason is that the learners are not ‘given’ any opportunity to ‘learn to understand an informal style of speech’ (ibid., p.5).

b) supra-segmental elements

Features on the supra-segmental level contribute even more to word recognition; for example, the ‘selection slot’ assigned by the speaker on the prominent stress (Brazil, 1994, p.86), stressed and unstressed syllable rhythm, facilitating meaning rendering, and intonation pitch variation patterns expressing the speaker’s intentions. Prosody is important for intelligibility, even though it seldom really results in the breakdown of recognition at word and utterance level, yet, as Kenworthy (1987, p.19) puts it, its effect can be ‘cumulative’, which can lead to misunderstanding just as a mispronunciation

does. Most Chinese language learners are exposed only to their language teachers or their classmates, with a high possibility that they share the same L1 background. Being accustomed to this non-native input, they will find it difficult to easily tune in to natural English speech patterns and understand the speech of L1 speakers in real-life encounters.

c) accent elements

Another important aspect of intelligibility cannot be neglected, namely accent. Received Pronunciation (RP), which is associated with similar labels such as ‘Queen’s English’, ‘BBC English’ and the like (Crystal, 2003b, p.3), is traditionally regarded as a model for language learners of English. With the increasing growth of English used as an international language, acquisition of native-like accent may not be any longer the ultimate goal for many language learners (Jenkins, 1998b, p.119). Communication with L1 speakers is widely changing to communication between non-L1 language learners. To communicate effectively in ‘interlanguage talk’, Jenkins (2000, p.69) emphasises the importance of perception and intelligibility of various accent characteristics.

According to Purcell and Suter’s (1980) study, accent is affected by four significant factors: mother tongue, aptitude for oral mimicry, length of time living in the target language environment, and concern for pronunciation accuracy. It has been noted by Kenworthy (1987, p.15) that a conversation is more easily understood between L2/EFL learners coming from the same L1 background. Kenworthy also claims that language teachers are not ideal judges in assessing the speaker’s intelligibility, since the teacher can tune in to the learners’ accents. This demonstrates the salient influence of exposure and familiarity in achieving efficient intelligibility. As Kenworthy emphasises, ‘the only

thing which will lead to a ‘permanent’ re-tuning is long-term experience and exposure to the new language’s sounds’ (p.49).

Research carried out by Gass and Varonis (1984) on the effects of familiarity on L1 speakers comprehension of accented speech, indicates that familiarity with NNS, a particular type of accent, and a particular speaker all had an effect on intelligibility. As Varonis and Gass (1985) put it, ‘[t]he less interlocutors know about each other, the more likely they are to misunderstand each other on a linguistic, social or cultural level’ (p.327). Wingstedt and Schulman (1984) also conclude that familiarity with a particular kind of accent facilitates intelligibility. Another study undertaken by Derwing and Munro (1997) demonstrates that familiarity with a particular language and recognition of the speaker’s L1 background is associated with greater success in language identification.

C. Importance of exposure to natural English speech in L2 acquisition

The importance of and necessity for exposure for language learners so as to cope with real, authentic English speech are stated not only by the above researchers; the following researchers also highlight these essentials from different points of views.

Brown (1990, pp.46-47) states that the ideal isolated forms of words are never spoken by L1 to L1 speakers. If language learners are constantly exposed to this kind of spoken English, inevitably they will find it quite impossible to understand normal English speech. From the point of view of methodology, Brown emphasises that language learners who are much exposed to natural English speech will learn English much faster than those who are not, because they have the chance to begin to ‘build up stereotypes

of familiar expressions' and learn to 'understand through context clues' rather than segmental signals (p.168).

Field (2000, p.190) in his reply to Ridgway (2000) argues that one of the reasons to use authentic material in teaching listening is that the material is 'unscripted', in which important characteristics of natural speech, e.g., natural rhythm, pause, hesitations, repetitions, false starts, tongue slips, are featured. Field (ibid.) emphasises that language learners 'need to be exposed' to this authentic language, since this is what learners will experience in the outside world. Another reason is these authentic materials are 'ungraded'. Listening to natural materials occurring in real life can provide language learners with the experience of natural interpretation processing by employing the same everyday listening skills that an L1 speaker uses, and which they themselves use in their own native language. Walter (2008) shares a similar idea and also emphasises that 'time is well spent in more exposure to spoken language'.

2.3.2 Speaker and Listener Convergence Based on Shared Contextual Knowledge

This sub-section reviews issues of interlocutors' achieving of communicative convergence on the basis of shared contextual knowledge, and it is structured as follows:

- A. what communication is
- B. context knowledge, role and effects
- C. communicative convergence between interlocutors
- D. acoustic confirmation of communicative convergence

- A. What communication is

Communication involves, on the one hand, the transfer of the speaker's ideas into the listener's mind, and on the other hand, the interpretation of the listener's intention by his utterance. This is not a simple, straightforward process. Instead, it is a dynamic negotiating process, in which the stances of both the speaker and the listener are always shifting; a communicative 'loop' – a constant feedback, checking and monitoring process (Moore, 2005) is involved, and a 'collaborative' contribution (Schober & Clark, 1989) is needed to ensure a maximum convergence between the speaker and the listener within the context in which the conversation occurs.

B. Context knowledge, role and effects

According to Field (2004, pp.76-77), context is widely used to cover any of the following: 'immediate situation', 'meaning representation', 'topic' and 'co-text'. There are two effects according to the ways in which 'general features' (i.e., a blanket term to describe context, used by Nation and Coady, 1988, p.102) are interpreted. One is related to the interpretation of a word or an utterance, and the other is linked to lexical access. Rendering the speaker's intention involves firstly interpretation of the words used. The main argument made by Field (2003a, p.10) is that a word is a 'movable unit of meaning' which cannot be broken down into small individual elements. Understanding a word must be based on its links to its surrounding words, and also depends on the context in which the word is embedded. Three types of schema pointed out by Field (2003a) are involved in conversational interpretation. They are a world knowledge schema, a contextual schema and previous experience schema (p.40). These external factors are generally referred to as context knowledge, which plays an essential role in understanding a communication.

C. Communicative convergence between interlocutors

Convergence, as defined by McCarthy (1998, p.177), ‘is an ideal state where speakers’ minds mesh, where they are on the same wavelength, pursuing the same goals, and each participant sees the same need to co-operate and get to the desired outcome’. In communication, neither the speaker nor the listener has a privileged controlling right to the topic, they must negotiate it. Their participatory role in the dialogue of the interlocutors is constantly shifting. There is a constant flow of interruptions, floor-grabbing, arguments, contention for possession of the floor and securing the floor. This co-operative cohesive achievement of discourse is heavily dependent on real-time adjustments, demonstrated to some extent by speakers’ needs to ‘negotiate meanings’ of the vocabulary used (Carter & McCarthy, 1988, p.xi). Speaker and listener actively accommodate each other and jointly contribute to the convergence by ‘playing the same game’ (Cicourel, 1973, p.87), in which similar lexical patterning, for example prefabricated formulaic expressions (which is outlined later in detail in Sections 2.3.3 and 2.4), is co-presented by both speaker and listener. This approach to the use of formulaicity, as argued by McCarthy (1998), can facilitate fluency, project the learner’s personality and establish appropriate socio-pragmatic, interactional relationships in communication (pp.109-15).

Communicative value is also discussed by Widdowson (2007), in which he points out that communication is not simply about transferring knowledge into agreement, but ‘a degree of convergence’, in which a ‘quite complex negotiation’ process is involved (p.26). Widdowson argues that language is only produced when there is an occasion to use it, and the occasion for language use takes place in the ‘continuous and changing contexts’ of our everyday life (p.19). Communication is heavily shared-knowledge-

based, speaker and listener can understand each other only within a common situation, called context (p.22). According to Widdowson, both speaker and listener have their own schematic structures of knowledge. In communication, new information constantly emerges to fill out the existing schema and enrich the personal frame of reference of the speakers. The less knowledge is shared by speaker and listener, the more pronounced the divergence between the interlocutors, and the greater the need for accommodation and negotiation between the speakers.

Achieving communicative convergence is taken as a kind of ‘communicative competence’ (Campbell & Wales, 1970, p.249), in which a negotiation process for common agreement between the interlocutors is involved. This convergent process, as argued by Widdowson (2007, pp.63-67), is pragmatically oriented by communicative intentions. For an efficient communication, there is a composite process towards convergence between speaker and listener via ‘give and take on both sides’. On the one hand, both speaker and listener have to insist on and protect their own stance – ‘individual reality’, ‘a sense of self’, and ‘a personal territory of identity’. On the other hand, a collaborative relationship has to be established and retained between the interlocutors, which runs a risk of compromising individual identity. These two aspects are termed by Widdowson as ‘territorial imperative’ and ‘co-operative imperative’. It is therefore not only meaning that is negotiated in conversation, but also social and human relationships, for example friendliness, politeness, and individual attitude. Therefore, to some extent, communication is ‘an exercise in control’ – to ‘assert one’s own position’ and to ‘persuade the other to accept it’. Another point made by Widdowson is that convergence can only be achieved indeterminately and partially, which is a very common feature of communication. No matter how well interlocutors know each other,

perfect understanding never occurs. This subjective perceptual opinion conforms to Brown's (1990) 'adequate' understanding as examined in Section 2.2.4.

D. Acoustic confirmation of communicative convergence

Speaker and listener convergence in a single dialogue is also investigated by Kousidis, et al. (2008) within the context of speech recognition. In this study, four acoustic features are investigated between two speakers in a natural, unscripted dialogue – mean pitch, mean intensity, pitch range and speech rate. A direct comparison between time-aligned frames shows a persistent convergence in intensity (see Figure 2 below) and speech rate (see Figure 3 below). This evidence indicates that speakers readily adjust their acoustic behaviour to accommodate each other, which echoes Tatham and Morton's (2002) collaboration notion between speaker and listener. Another study undertaken by Derwing (1990) shows that, in order to accommodate L2 learners, L1 speakers (10 out of 16 test persons) increase the pause time in their narrations. Even though, as Chaudron (1982) argues, some adjustments may have 'adverse effects' on communication, nevertheless, the majority of researchers' investigations show a general tendency of co-operation and accommodation between interlocutors.

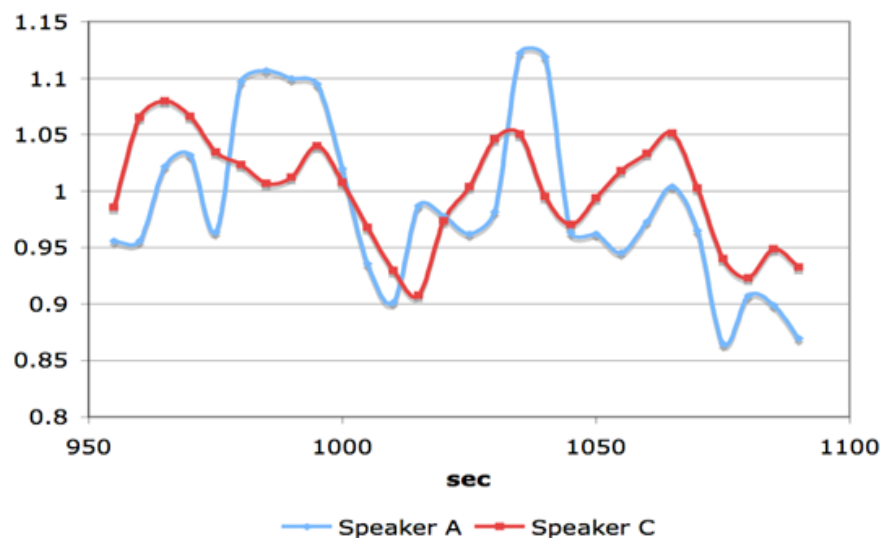
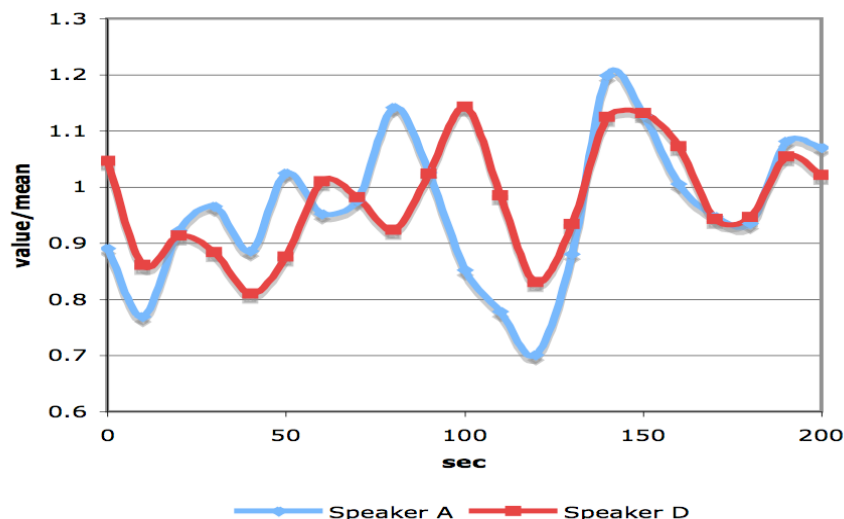


Figure 2: Average normalised intensity for speakers A, C over 10 second frames with 50% overlap

Source: Kousidis, et al., 2008

**Figure 3: Average normalised speech rate for speakers A, D over 20 second frames with 50% overlap**

Source: Kousidis, et al., 2008

2.3.3 Formulaic Language and Spoken English

This sub-section deals with a general review of formulaic language and English speech. It first looks at the concept of formulaic expression, i.e., its definition, its processing and storage. Then it examines the role of formulaicity from the perspectives of social interaction, processing economy and language learning. These then lead on to further consideration of formulaic language in Section 2.4.

A. Concept of formulaic expressions

Research in recent years has convincingly shown the importance in informal speech of formulaic language, i.e., pre-fabricated linguistic segments. This long-recognised linguistic phenomenon has been investigated by many scholars. As Hymes (1968,

pp.126-27) points out, a vast proportion of verbal activity is composed of ‘recurrent patterns’, and ‘linguistic routines’, which are of conventional significance within a particular society. ‘[T]he unit of actual speech is the holophrase’ (Firth, 1964, p.83), the language we use is not built from scratch, but contains ‘an incredibly large number of prefabs’ (Bolinger, 1976, p.1). Irrespective of the size of these prefabricated elements, they are always considered as a unit which ‘can not be further analyzed or decomposed in the way a free combination can’ (Jespersen, 1976, pp.85-88). The fixedness of prefabricated units is also noticed by Saussure (1966, p.177) who claims that, ‘the mind gives up analysis’. Later, a working definition of ‘formulaic sequence’, based on the manner in which they are stored in the brain, is given by Wray (2002, p.9) who sees formulaic language as being ‘a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar’.

a) defining ‘big words’

More than forty different terms are quoted by Wray (2002), to describe the phenomenon of ‘big words’ (Ellis, 1996, p.111), such as chunks, collocations, idioms, fixed expressions, formulaic speech, and so on, although some of them overlap. There is also a range of terms defined by other researchers depending on their analytical focus. Irrespective of terminology, formulaic sequences share a common advantage of having a wide acceptance and use in natural L1-L1 informal speech. Erman and Warren (2000) calculate that all the various types of formulaic language make up 58.6% of the spoken discourse they analysed. Altenberg (1990, p.134) gives an estimation that about 70% of the running words in the corpus he analysed constitutes some type of formulaic string,

and later this figure is even increased to 80% (Altenberg, 1998, p.102). Certainly, there are also some scholars who provide a relatively low figure for formulaicity due to the lack of agreement on criteria for their identification and measurement. Nevertheless, as Perkins (1999, pp.55-56) puts it, in our everyday speech, ‘the patterning of words and phrases ... manifests far less variability than could be predicted on the basis of grammar and lexicon alone’. This striking similarity of everyday English conforms to the ‘idiom principle’ proposed by Sinclair (1987, pp.319-20), and it does seem to contribute to the success of real English native-to-native communication.

Research into language acquisition, as examined below, demonstrates clearly that language is learned, stored, retrieved, and produced in whole phrases and other multi-word constructs, alongside individual words. Williams (1988) points out that native-like ability of lexical selection includes the selection of preferred sequences from a number of grammatically acceptable varieties, and ‘it is our ability to use lexical phrases ... that helps us speak with fluency’ (Nattinger & DeCarico, 1992, p.32). The advantage of fluency can allow speakers to focus their attention on the larger structure of the discourse, rather than struggling narrowly with individual elements. Native-like control of the language, for normal adult L1 speakers, does not actually make use of all the productive possibilities of lexical and grammatical rules to reconstruct the language bit by bit each time it is needed. In many circumstances, particular communicative functions are realised by particular linguistic forms; the speaker only needs to employ pre-stored complete units to fill in the ready-made slots, depending on the context.

- b) evidence supporting holistic processing and storage of formulaic language

A dual-system solution, proposed by Wray (1992), contains analytic processing and holistic processing. It is the holistic processing, which emphasises the retrieval of pre-composed strings of words from memory which is favoured by adult speakers, rather than analytic processing based on decoding individual linguistic elements from scratch while conforming to grammatical rules (p.18). This explanation backs up the theory that language is stored redundantly (Bolinger, 1975). The same information is stored not only as partially assembled pieces in the lexicon, but also as holistic, memorised chunks which free speakers from the enormous burden of building up every expression ‘from scratch on the spot’ (Nagy, 1978, p.289).

There is also some psycholinguistic and neurological evidence supporting the holistic storage and production of formulaic language, as outlined below. Evidence given by Van Lancker (1987) shows that (for a right-handed person) the right hemisphere processes complex patterns and can store and process non-propositional language (p.101). In addition, given the very limited capacity of working memory (i.e., short term memory) for handling the current information and limited processing speed in the human brain (Crick, 1979, p.219), listeners need to rapidly transform the language they hear and read into pieces of abstract information before they decay. One linguistic solution pointed out by Miller (1956, p.93) is that ‘we can increase the number of bits of information that [the communication] contains simply by building larger and larger chunks, each chunk containing more information than before’. In other words, chunking fragment items into larger units can actually increase the capacity of short term memory (Anderson, 1983, p.39).

In terms of neurological reality, a study carried out by Oppenheim (2000) suggests that ‘when speakers need ... to express the same message, the same neural networks [are]

excited, causing speakers to repeat many same-word sequences' (p.229). For processing fluent speech production, a significant amount of automaticity is required, which is supported under 'open-loop control' where whole chunks of unanalysed language are automatically retrieved and produced (Code, 1994, pp.139-40).

B. Roles of formulaic expressions

Central to the 'pragmalinguistic competence' (Leech, 1983, p.11) – the ability to 'select and retrieve ready-made form/function composites' (Nattinger & DeCarrico, 1992, p.13), the roles of formulaic language are treated by many scholars, as detailed below.

a) facilitating social interaction

Adult L1 language speakers master a great many various frozen or half-frozen formulaic expressions, which makes them easier to understand. As Pawley (1991, p.339) puts it, 'the conventions shared by a speech community ... specify ... what things may be said about a particular subject or topic, how these things are said, idiomatically, and when they are said, appropriately'. These formulae are socially 'licensed' (Smith, 1991) to perform a particular function within a certain community. It is 'a quick means to be communicative' (Schmitt & Carter, 2004, pp.11-12) by 'bypassing the difficulties of processing' (Wray & Perkins, 2000, p.17).

Another feature of these prefabricated expressions is that they are also culturally-coined interaction formulae. This is the way we say it; it 'sounds right', and is 'regularly considered by a language community as being a unit' (Moon, 1997, p.44). Any departure from this is culturally inappropriate, and may 'render the magic wholly ineffective' (Sebeok, 1964, p.356). The reason why these fixed expressions are

remembered and repeated, as Watkins (1992, p.393) states, is not ‘because they delight the ear, rather they are signals, ... , of the relations of things’. Culturally sanctioned formulaic phrases can be used either as an inclusion (Garvey, 1977, p.43) or exclusion (Schmidt, 1983, p.156), by which some individuals are included and grouped, some are not. Mastered and used appropriately, these social-cultural conventionalised codes for language learners are key for maintaining and establishing social relationships (Yorio, 1980, p.438).

b) reduced processing effort

Firstly, according to Perkins (1999, p.56), the main reason why formulaic expressions are so popular in adult speech is their ‘simple processing principle of economy of effort’. As mentioned previously, using ‘ready-made frameworks’ needs little encoding work, which is more economic than going through the labour of reassembling an utterance at the time of production. By reducing the processing load, speakers can accommodate other aspects of social activities. As for listeners, taking shortcuts in decoding the packaged formulaic information allows them to devote their attention to novel segments.

Secondly, central to lessening processing effort by using pre-assembled language is saving working memory capacity, which allows for fluent production and faster processing (Raupach, 1984). A study carried out by Kuiper (1996) shows how traditional oral poets and singers largely use memorised bits of speech retrieved ‘from the dictionary’ (p.3) and produce their output ‘in stereotypical form’ (Lord, 1960, p.24) in order to lower the on-line performance pressure. Later, this formulaic property is further investigated in terms of facilitating the pressured situations of auctioneers or

sports commentators by their extensive use of formulaic speech to convey fluently large amounts of transactional information and commentary under severe time constraints. Kuiper (2004, p.38) concludes that it enables them to ‘do a particular job’.

Apart from relieving time pressure, using multiword expressions can also save processing effort by buying time for planning subsequent discourse and promoting fluent output, as argued by Wray (2002). According to Wray, saving processing time and effort is not simply a matter of ‘taking short cuts’, but ‘about regulating production’, even sometimes taking the long way round, to fulfil the ultimate goal of ‘maintaining a particular preferred rhythm and flow’ (p.75). This view conforms with and is extended by McCarthy and Carter. McCarthy and Carter (2006b), by investigating automatic retrieval of multiword strings of language in the data they analysed, confirm the pragmatic and interactive functions of these pre-formulated units, as they put it, and ‘make fluency a reality’ (p.23). This notion corresponds with Towell, et al.’s (1996) study on fluent production of an increase in length and complexity of proceduralised units by British advanced learners of French (pp.112-13).

Comparing the idealised phonological notion of fluency, which is linked with speech rate, pausing time and percentage of repair fluency, as seen in a study carried out by Xu and Ferguson (2008), a dialogically fluent performance, termed ‘confluence’, is proposed by McCarthy (2006, in press), i.e., interlocutors’ shared responsibility for creating and maintaining conversational flow by using formulaic discoursal expressions. ‘[T]he ability to retrieve, quickly and automatically, items from a repertoire of ready-made chunks, especially the core, most frequent ones in everyday talk’ is one of the central determinants in constituting flow and creating confluence (McCarthy, 2008, p.33).

c) improving language learning

As Weinert (1995, p.184) points out, language learners, by analysing formulaic sequences, derive linguistic rules and can create their own productive output. A similar opinion is shared by Shin and Nation (2008). A double role of formulaic sequences in language acquisition is described by Wood (2002, p.5) in that, ‘they are acquired and retained in and of themselves, linked to pragmatic competence and expanded as this aspect of communicative ability and awareness develops. At the same time, they are segmented and analysed, broken down, and combined as cognitive skills of analysis and synthesis grow. Both the original formulas and the pieces and rules that come from analysis are retained’. In this, prefabricated chunks of unanalysed language seem to be an ideal unit for language learning. ‘It is impossible to perform at a level acceptable to native users, in writing or in speech, without controlling an appropriate range of multiword units’ (Cowie, 1992, p.10).

2.3.4 Summary

In this section, natural authentic English speech is reviewed from the aspects of facilitating mutual intelligibility, speaker convergence via shared knowledge between interlocutors, and using formulaic language. Pragmatic competence, compared to linguistic competence, seems to be more plausible and significant in natural English conversations. Internalising these communicative strategies for L2/EFL learners can enlarge negotiation convergence, so as to facilitate native-like communication.

In view of the considerable proportion of formulaic language employed by L1 language speakers in their everyday speech, the nature of this kind of formulae and their overall

functions in spoken language are introduced in this section. Further analysis is reviewed in greater detail in the following section.

2.4 Learning and Acquisition of Formulaic Language

Based on the introduction to formulaic language in Section 2.3.3, some detailed aspects of this phenomenon are now considered in this section. It first looks at the categorisation of formulaic language (2.4.1), followed by its phonological perception (2.4.2). The different acquisition and use of formulaic language by L1 users and L2 learners is then reviewed (2.4.3). Finally, the issues of enhancing the learning and acquisition of formulaic language for L2/EFL speakers are dealt with (2.4.4).

2.4.1 Multi-criteria Categorisation of Formulaic Language

Setting criteria for the identification of formulaic sequences is the basic prerequisite for its classification.

A. Wray's (2000) criteria

Four kinds of structural features are defined by Wray (2002, p.47), they are: form, function, meaning and provenance.

a) form

The form criterion is normally regarded as an easy way to categorise formulaic strings, in which collocation is one of the dimensions being widely addressed. According to Sinclair (1991, p.170), collocation is defined as 'the occurrence of two or more words within a short speech of each other in a text'. Compared with other types of multi-word strings, collocation is 'much more fluid', and is only about 'tendencies and preferences' (Wray, 2002, p.51). Three types of collocations are analysed by Moon (1998, pp.26-28),

in terms of surface lexical co-occurrence, categorical association and syntactic relationship. Teubert and Cermakova (2007, p.113), from a language understanding point of view, define three criteria for assigning collocation status based on whether a lexical pattern can be repeatedly paraphrased, whether it can be used in a metaphorical way, or whether it can be translated as a whole. Thus two kinds of collocations are classified. One is fixed expressions with a certain grammatical pattern, and the other is lexical co-occurrences restricted within a certain context (p.116). From the categories given above, it can be seen that, using a form-based criterion, it is not easy to classify a formulaic unit, because '[f]orm and meaning are inseparable' (Stubbs, 1993, p.17).

b) function

According to Wray (2002), the second criterion she considers as an important factor is function, which, as Coulmas (1981, pp.2-3) puts it, ties the expressions to 'more or less standardized communication situations'. Since the conventionalised forms are coined within their social-cultural contexts, which facilitates communication by reducing the complexity of choices, many scholars set this dimension as their starting point, as detailed below.

Nattinger (1988, pp.76-77) describes six types of 'lexical phrase' (DeCarrico & Nattinger, 1988) based on the characteristics of function-structure composites, they are: polywords, phrasal constraints, deictic locutions, sentence builders, situational utterances and verbatim texts. Later Nattinger and DeCarrico (1992, pp.65ff) revisit this preassembled unit, and re-categorise its types, mainly focusing on its pragmatic function. Three functional categories are defined, that is, social interactions, necessary topics and discourse devices. Collocations, idioms and syntactic strings are excluded

from this category, due to their lack of clear pragmatic function. Thus, the six-fold categorisation of form is modified to four: polywords, institutionalized expressions, phrasal constraints and sentence builders (pp.37-47). Classifying the categories in this manner, as Nattinger and DeCarrico (ibid.) point out, is based on frequency of occurrence rather than structure.

Moon (1998) also uses a function-based criterion in her analysis, which is termed FEI (fixed expressions and idioms) (p.3). Three factors are principally employed to identify these sorts of holistic units: institutionalization, lexicogrammatical fixedness, non-compositionality, along with three other factors based on length, syntactic integrity and phonological criteria. Classified by these dimensions, five discoursal roles are defined which perform informational, evaluative, situational, modalising and organisational functions, to cover her three macrocategories: anomalous collocations, formulae and metaphors. Moreover, an alternative classification is also given, based on the restriction of paradigmatical, syntactic, meaning and transparency. Given the overlap between the groups, nearly 47% of FEIs in Moon's data have at least two functions. Therefore, in conclusion, Moon (ibid., p.23) asserts that, 'it is often impossible to assign an FEI to a single category'.

c) meaning

According to Wray's (2002) description, the third criterion for identification of formulaic expressions is meaning, in which transparent and non-transparent are two distinctive features. Idiom is deemed to be the most common category.

Wood's (1986, p.2) definition of the 'true' idiom is 'a complex expression which is wholly non-compositional in meaning and wholly non-productive in form'. Flavell and Flavell (1992, p.6) state that idioms 'break the normal rules'. Thus, the central definition of an idiom is that its meaning derives from the holistic interpretation, rather than simply the sum of its individual constituents. A study carried out by Gorokhova (2008) shows evidence of incorrect retrieval of lexical items due to decomposition of an idiom. Two kinds of idiom are generally classified by Moon (1998, pp.4-5). Narrow idioms, or 'pure idioms' (Fernando & Flavell, 1981), which are fixed, semantic opaque or metaphorical, and rendered as a whole; while broad idioms loosely refer to many types of multi-word strings, no matter whether they are semantically opaque or not.

McCarthy (1998, pp.129-31) also discusses idioms and points out that both syntactic, lexical and phonological form, as well as semantics and pragmatic functions of the unit are, to some extent, fixed. Seven categories are detailed, that is, prepositional expressions, binominals and trinomials, frozen similes, possessive 's phrases, opaque nominal compounds, idiomatic speech routines, gambits and discourse markers, and cultural allusions. One point is emphasised – that idioms are different from the other categories of formulaic language in that they are more culturally rooted, and this requires a necessary shared cultural background in order to interpret them appropriately. McCarthy (ibid.) also asserts that drawing an absolute distinction among these categories is not only 'problematic', but also 'impossible'.

d) provenance

The last dimension defined by Wray (2002) is provenance, which means 'the way that formulaic sequences come about' (p.59). Classifying formulaic language based on their

provenance, as Wray argues, can accommodate both smaller strings being stitched together and those longer units which have never been broken down due to their complexity of meaning (p.61). Two possibilities are defined by Peters (1983, pp.2-3), in her description of ‘speech formula’, which are ‘either through social negotiation or through individual evolution’. Based on Peters’ analysis, a seven-way division is proposed by Weinert (1995, pp.182-83) in relation to acquisition, that is, phonological coherence, greater length and complexity, non-productive use, community-wide use, idiosyncratic/inappropriate use, situational dependence, and frequency and invariance. Meanwhile, Weinert also states that there is quite a large degree of overlap in the identifying criteria.

B. Continuum model

In view of the overlap between the four criteria reviewed above, a continuum model is envisaged, which views formulaic language as a dynamic continuum. Some scholars are more interested in variability, and, as Howarth (1998, p.35) points out, ‘[i]t is essential to see the categories as forming a continuum from the most free combinations to the most fixed idioms, rather than discrete classes’. Another continuum is defined by Cowie (1988, pp.133-35), which combines the variety of lexical form and literal meaning. Two main groups of word combinations are proposed, i.e., ‘*pragmatically* specialized’ and ‘*semantically* specialized’. Givon (1989, p.258) also proposes a continuum approach, called ‘automaticity continuum’, in which different scales of categorises are conceived as occurring along a continuum from the most conscious patterns to the most automatic patterns. The continuum model seems to be able to better describe the category of formulaic strings; however, there is still more than one version and no consensus among scholars.

C. Categorisation using multiple criteria

Wray (2002) comes to the conclusion that all four themes in the definition of formulaic sequences are ‘closely interrelated’ (p.65). The four criteria are not mutually exclusive, but rather they are overlapping. This, however, causes some problems in assigning a category to a particular formulaic string. However, it is not necessary to insist on a separation of these criteria, since formulaic language is profitably explored based on a usage-based language theory, and various analyses vary according to the area of focus. Using clear-cut criteria is running a risk of ‘misrepresenting the nature of the native speaker’s knowledge’ (Pawley & Syder, 1983, p.212). Thus these scholars agree that a multi-criteria classification is an effective and plausible solution, so as to better demonstrate both the linguistic and the pragmatic features of formulaic language.

2.4.2 *Phonological Perception of Formulaic Language*

A. Phonological analysis in current literature

Compared with the other aspects of formulaic language, the amount of research into the phonological realisations of unanalysable chunks of speech is relatively modest, and there are only a few studies undertaken by some scholars. The consideration is mainly on the grounds of articulation, speed of delivery, stress, pause and intonation patterns, as examined below.

a) articulation

Features in the articulation of formulaic strings are mainly represented by less precise articulation and reduced phonemic production. Van Lancker, et al. (1981) provide a

thorough investigation of phonological cues in the production of formulaic language. In terms of articulation, they point out that less clear pronunciation is apparent in the production of idioms compared with the equivalent novel lexical units (pp.333-34). Plunkett (1993, pp.46-47) also gives evidence of the segmentation of an utterance into pieces or chunks, in which reduced clarity of pronunciation is taken as one of the criteria to identify formulaic units. The process of formulaic sequencing from conceptualisation to realisation is explained by Hudson (1998, p.2), in which phonetic reduction is regarded as making a contribution. The liaison phenomenon of French speakers is investigated by Bybee (1998), in which she proposed that the distribution of liaison indicates the structure of formulaic storage, and is 'evidence for the size and nature of processing units' (p.432). Thus the phonetic reduction of preassembled strings seems to contribute to the description of the single semantic and phonetic identity of formulaic units.

b) speed of delivery

As mentioned above, since there are some reductions in pronunciation, a faster processing seems to be a legitimate consequence. Van Lancker and Canter (1981) report that an idiomatic expression is normally articulated faster than a literal one when rendered within context. Underwood, et al. (2004) also provide evidence, by eye movement, that 'terminal words in formulaic sequences are processed more quickly than the same words when in nonformulaic contexts' (p.167). Another observation by Lin (2006) also provides support for the changes in articulation rate which can be a potential indicator of formulaic processing.

c) stress

In addition to articulation and speed of delivery, stress is also seen as a key feature of identification of formulaic language. Jespersen (1976, p.83) argues that, when talking about the variability of formulaic sequences, the reason why chunks of speech are ‘felt and handled as a unit’ is that ‘no one can change anything in them’. It is impossible to pause within such units or to stress them in a different way (ibid.). This notion is consistent with Peters’ (1983) finding that regular patterns of word-initial stress are likely to be, amongst others, a candidate to segment an unanalysed string and identify salient formulaic units (p.36).

d) pause

An experience described by Brown (1973), while learning Japanese, shows that a novel lexical string is regarded as a single word until similar utterances are encountered with the same structure. The main reason for this is ‘it is spoken without pause’ (p.5). This description corresponds with the study by Goldman-Eisler (1968) on pauses in speech production. According to her investigation, pauses more frequently occur in novel speech strings than in formulaic units. An experiment carried out by Van Lancker, et al. (1981, p.331) suggests that for a string carrying literal meaning, interword pauses and word durations are longer than a string with idiomatic meaning, because they contain a greater number of pauses and also because the key lexical items in literal readings are produced more slowly. All these analyses are in line with Raupach’s (1984, pp.114-16) assertion that pauses or hesitation phenomena may function as a cut-off point, which leads to a preliminary identification of formulaic sequences as those not being interrupted by unfilled pauses. Pause as an indicator of multi-word expressions is also investigated by Dahlmann and Adolphs (2007), based on an approach of combining frequency and psycholinguistic description of multi-word expressions. By analysing an

interview corpus of two Chinese learners of English, they suggest that pause is a valuable factor to indicate possible boundaries; consequently, the placement of pauses might be an additional criterion for the identification of the holistic storage of prefabricated language (p.55).

A potential limitation to using pauses as an objective empirical criterion is the fact that ‘pauses in natural speech only occur every 12 syllables or so’ (Field, 2003b, p.327). Pawley and Syder (2000) also express the regularity of ‘the average number of words per fluent unit is about six’ for fluent L1 speakers (p.195). This means pauses potentially occur regularly and might not always correspond with formulaic string boundaries. These findings challenge to some extent the role of placement of pause in identifying the boundaries of formulaic units.

e) intonation patterns

The features of ‘phonological coherence’ (Hickey, 1993, p.32) of formulaic language are also exhibited by overall fluency and intonation shapes. Evidence given by Van Lancker, et al. (1981) shows that pitch contour is one of the main acoustic cues distinguishing the idiomatic strings from literal readings. Formulaic strings are typically produced in a fluent manner with a coherent intact intonation contour, whereas literal utterances tend to contain more pitch changes. This phonological criterion is proposed earlier in Makkai’s (1972, p.29) study, in which compositional and noncompositional strings are distinguished by intonation patterns. Formulaic units are deemed to be encoded without a break in intonation contour. An example given by McCarthy (1998, p.129) emphasises that an idiom like ‘rough and ready’ is produced within a single tone unit. Cowie (1988, p.134) also states that one requirement for the successful usage of a

formulaic expression is an intact intonation pattern. Consistent intonational and rhythmical shapes are also investigated by Pawley (1991) and Kuiper (1996), in which sports commentaries are ‘delivered extremely fluently in a droned intonation’ (Pawley, *ibid.*, p.340) and with a ‘regular syllable-per-second delivery rate’ (Kuiper, *ibid.*, p.19). So are the cases of auctioneers and race callers (Kuiper, *ibid.*, pp.36-37). On the analysis of internal cues for segmentation of units, Peters (1983, p.37) asserts that ‘rhythm and intonation should play a part in determining segmentation points’.

Lin and Bahlmann (2008) also analyse the role of prosodic features in the identification of formulaic language, in which many factors are found associated with broken intonation contour, such as, final syllable lengthening, pause, a global declination in pitch, pitch reset and falling tone.

B. Link between different categories of formulaicity and their phonological realisations

As reviewed above, many phonological features are identified, linked to and defined, to some extent, as phonological cues for the identification of formulaic language. However, the present author has found no evidence in the literature referring to real dynamic dialogue and showing the link between the different categories of formulaic language and their relevant phonological characteristics. This is the gap the present author aims to bridge for the body of formulaic language, which is outlined in Section 7.2.3.

2.4.3 Acquisition and Use of Formulaic Language by L1 and L2 Learners

This sub-section deals with the difference in acquisition and use of formulaic language by L1 language users and L2 learners. It first gives an overview of the acquisition and use of formulaic language by L1 and L2 speakers. It then reviews natural acquisition and holistic processing of formulaic language by L1 users. The then follows a contrast between various types of acquisition and use by L2 learners. Finally, the importance of appropriate acquisition and use of formulaic language for L2/EFL learners is considered.

A. Acquisition and use of formulaic language by L1 and L2 users

Wray (2002) gives a comprehensive overview of formulaic language acquired and used by L1 speakers and L2 learners. The first point she makes is that L1 users have a greater tendency to use formulaic language as a routine shortcut in their everyday processing and communicative interaction. The second assertion she makes is that there is some similarity between L1 users and L2 learners in their early stages of acquisition, in which formulaic language are heavily used as a useful starter to facilitate their initial social and linguistic interactions.

B. Natural acquisition and holistic production by L1 speakers

a) phases of acquisition

Formulaic language is seen as a feature of L1 language learning. First of all, based on Locke's (1995, 1997) work, a model of four different phases of L1 speakers' acquisition and use of formulaic expressions is described by Wray (2002, pp.132-35). Phase 1 starts at the time of birth till about 20 months when the child begins to use simple grammar to

build up utterances by combining small units. Much of the linguistic knowledge acquired at this very early stage is unanalysed chunks of language, which is oriented by ‘specialization in social cognition’ (Locke, 1995, pp.295ff) and ‘supports an affectively oriented developmental growth path that channels infants in the direction of spoken language’ (Locke, 1997, p.269). Phase 2 then begins when the child is 20 to 30 months old. At this period, the child becomes more aware of vocabulary and grammar, which leads to a change in balance from formulaic language to novel utterances. In this stage, analytic strategy, supported by formal education and particularly literacy, reaches its maximum potential. Then, in Phase 3, aged from 8 to 18 years, a greater formulaic output returns and occupies a larger part of the individual’s lexicon. The reason for this is that it is inefficient and represents an extensive effort to reconstruct utterances, which are regularly called for, from scratch, every time they are used. Following the holistic storage and use of formulaic strings, in Phase 4, which starts in the late teens, the balance between holistic and analytic processing is fixed. Evidence from adult L1 speakers’ significant proportion of formulaic language in their everyday speech supports the description of formulaic language as an ultimate solution for both transactional and socio-interactional activities.

Wray’s (2002) analysis is in line with some investigations done by other scholars. Peters (1983) describes the units encountered by a child as ‘an intermittent stream of speech sounds containing chunks, often longer than a single word, that recur with varying frequency’ (p.5). From the child’s viewpoint, however, each string of words (morphemes) may be seen as only one unit, since these basic strings naturally correspond with a comprehension strategy from which the child captures general meaning and saves for future use, without having to go down to the lowest level of linguistic knowledge (Golinkoff & Hirsh-Pasek, 1995, p.430). This process of

extracting units from speech input is assumed to be much beyond the child's current linguistic capabilities. Therefore a holistic or gestalt approach (Peters, 1977) is adopted, which defines that the unit has to be 'produced without pauses between words, with reduced phonemic articulation, and with the effect of slurred or mumbled speech but with a clear intonation pattern enabling the listener to construct the target utterance *in context*' (Nelson, 1981, p.174). The central role of formulaic acquisition at this early stage is emphasised by Cruttenden (1981).

b) acquisition through 'fusion'

Another way to acquire prefabricated chunks, based on Peters (1983), is through the processing of 'fusion' (p.80). Children tend to utilise any available means to overcome the disadvantage of linguistic limitations, so as to express themselves and get the things done. Some 'stereotyped expressions' are created by children, which are 'neither copied directly from nor even directly reduced from adult usage' (p.82). By this process, some often-used speech sequences are stored as preformulated units for quick and easy retrieval on subsequent occasions (Lennon, 2000, p.39). This fusion process can go beyond childhood and continues into the adult period. Two relatively independent continua, as defined by Peters, are involved in this process. One is the extent of grammatical transparency/opacity, which mainly refers to 'idiosyncratic formulas' unilaterally stored for an individual use. The other is the degree to which a particular expression is accepted and becomes fused in a community, which, especially in the case of very opaque expressions, can be used as a kind of verbal fence to signal the identification of the learners within a linguistic community.

C. Various types of acquisition and use by L2 users

Given that formulaic language is a dynamic and specific response to processing and interactional needs (Wray, 2002, p.123), as examined above, a corresponding dynamic acquisition process consequently occurs for L1 speakers. By way of contrast, L2 data show various learning types and strategies adopted by L2 learners at different stages.

a) young learners

Very young children, investigated by Wray (2002, pp.153-56), use both memorised chunks and fused constructions to facilitate interactional activities. Segmentation processing occurs at this stage for later analysis into syntactic rules. Evidence from Itoh and Hatch (1978, p.83) also shows that repetition is particularly effective, and leads to further expanding use of formulaically-based chunks. This imitation processing is seen as natural as native-like acquisition without interference from learners' L1 language (Huang & Hatch, 1978, pp.123-24). Therefore, at this very early stage, full language acquisition by L2 learners occurs naturally.

b) primary school learners

When children are in primary school, the acquisition and use of formulaic strings is mainly oriented towards establishing social relationships and addressing the communication shortfall. A study by Wong Fillmore (1979, p.280) indicates that the spectacular success of one of the candidates, Nora, in actively integrating into an L2 social community, is attributable mostly to her principal goal of learning the language, that is, 'to enact a socially significant event in order to construct identities as competent students ... and construct collaborative relations with one another' (Willett, 1995, p.490). With increasing age, children's attitudes to L2 learning are more and more aligned to that of an L2 adult learner (Wray, 2002, p.205).

c) adults outside the classroom

Research into L2 adult learners in naturalistic settings shows that, in the initial stage, a preformulated store of strings ‘figures frequently in the speech of all learners’ (Ellis, 1994, p.85), and L2 learners do rely on formulaic expressions (although without accuracy) to bypass the processing difficulties (p.388) and achieve initial communicative success. However, while being aware that the word is probably the basic unit of language, in social interactions, adult learners may feel uncomfortable not knowing how to break down a memorised chunk into small units. This results in their tendency to focus on individual word acquisition, and they ignore formulaic sequences. The lexical composition of words and formulaic strings is very distinguishable from the lexicon storage of an L1 speaker (Wray, 2002, p.206).

d) classroom-based adult learners and teenagers

Compared with adult L2 learners who learn outside the classroom, classroom-taught adult learners and teenagers are more ready to apply analytic strategies to holistically learned word strings, as reviewed below. Explicit linguistic knowledge results in classroom learners’ absence of awareness of formulaic language (Howarth, 1996, p.186). Evidence supplied by Bishop (2004, p.239) supports this notion, in which unknown, non-salient formulaic strings are glossed less frequently than unknown words, due to their not being readily recognised as holistic units by L2 learners. Poor knowledge of routine expressions mastered by L2 classroom learners is investigated by Irujo (1993) and Schmitt, et al. (2004), among others, which shows that very few formulaic clusters are holistically produced and stored by L2 learners. As Pawley and Syder (1983) point out, one of the major difficulties for L2 learners, even those of

advanced proficiency, is to select the most idiomatic expressions customarily used by L1 speakers from a large stock of candidates exhibiting perfect vocabulary and grammatical structures. Thus native-like formulaic output is the most challenging for non-L1 learners. Since the main means to acquire formulaic expressions for L2 learners is by repeating and memorising the sequences learned whole from classroom input, however, not all memorised strings are idiomatic for L1 speakers, some of these sound native-like, but some do not. L2 learners have no idea of which is which (Cowie & Howarth, 1996, p.91), which leads to their inability in encoding and decoding them.

D. Importance of appropriate acquisition and use of formulaic language by L2/EFL learners

L2 classroom learners often have difficulty in mastering formulaic utterances, which consequently results in their language being unidiomatic. An observation by DeCock, et al. (1998) gives a thorough description, in which proficient L2 learners use, in some cases, more prefabricated routines than L1 speakers. However the prefabricated sequences they produce are either not the required target language formulas, or are not used with the same frequency, or have different syntactic or grammatical uses, or are used for different pragmatic functions (p.78). The importance of knowing ‘code responses’ is emphasised by Olsen (1972, p.145). Non-native-like use of formulaic language is also reflected in cross-culture interactions, which easily results in communication breakdown, or even conflict, as investigated by Nelson, et al. (1996). Therefore, more attention should be paid to contrastive pragmatics due to cultural variation and non-L1 language pedagogy (Aston, 1995, p.57).

2.4.4 Enhancing the Learning and Acquisition of Formulaic Language for Non-L1 Speakers

In this sub-section, the review is first based on the investigation of how to gain holistic input of formulaic language, then proceeds to the consideration of how to internalise natural use of formulaic expressions for non-L1 learners in a target language community.

A. Holistic learning in L2/EFL classroom setting

Given the different acquisition and use of formulaic language by L1 users and L2 learners, as examined in the previous sub-section, it is impossible for non-L1 language learners to adopt the same natural process to access and acquire formulaic language as L1 speakers, especially for classroom-learning L2/EFL speakers. The first step is, as Sinclair and Renouf (1988) propose, to undertake holistic intake, rather than to amass vocabulary bit by bit.

According to Sinclair and Renouf (1988), ‘the lexical syllabus does not encourage ... piecemeal acquisition’, especially at an early stage (p.155). Instead, it proposes that learners should make the best use of all the words they have learned. Building up a rich stock of expressions from chunks of language is more worthwhile than only mastering less-frequently used individual items. The correlation of success in an L2 language and learners’ ability to learn conventional routines is emphasised by Ellis (1996, p.91), in which ‘individual differences in learners’ ability to remember simple verbal strings in order’ is the most critical factor to link with the successful acquisition of L2 language. ‘[I]n order to survive in society we’ve got to know what to say, and we usually know it in advance by memorizing it’ (Becker, 1975, p.27). This pedagogy using whole-phrase

inputs is still widely used by some British secondary schools. As Mitchell and Martin (1997) put it, 'it is clear that prefabricated phrases have maintained a place in contemporary classroom practice' (p.6). The advantage for teaching communication using memorised formulaic word-strings, as stated by Hakuta (1976), is that it allows for expressions that learners are as yet unable to construct creatively, so learners do not need to wait until they acquire enough grammatical and lexical knowledge, otherwise, they will 'run into serious motivational difficulties' (p.333).

Classroom-taught success in the use of formulaic language is investigated by Schmitt, et al. (2004, p.68), in which, under semi-controlled input, participants' progressing 'from a partial receptive mastery to a more complete productive mastery' does indicate the successful input of formulaic language in an intensive language programme. Another means, proposed by Nattinger and DeCarrico (1992, p.113), is teaching conversation. The main advantage for this is that classroom learners can learn how to use routine strings to create the flow of a spontaneous unfolding conversation, rather than to learn isolated, individual words. A case studied by Wray (2004) demonstrates the considerable success of holistically memorising prefabricated, multiword sequences by an adult classroom L2 learner of Welsh. Two interacting factors – 'successful automatization and the absence of analysis' – are recognised as contributing to the general success (p.262).

B. Holistic learning by exposing learners to a L1-speaker environment

As argued by Cowie (1988, p.137), mature L1 speakers mainly achieve their linguistic competence by exposing themselves to everyday interactions, through which conversational expressions are customarily and eventually internalised. Therefore,

increasing exposure to natural ‘day-to-day lexical performance’ can effectively compensate for the lack of real routine formulaic input of L2 learners in their language learning classroom. ‘[F]orc[ing] [learners] to engage with the L2 for meeting basic needs’, as Wray (2002, p.148) proposes, is the best way for children to learn an L2.

The influence of exposure to a natural target-language environment on the acquisition and use of formulaic language is investigated by Adolphs and Durow (2004). Comparing the interview data over a period of seven months, it appears that more lexical or phrasal sequences replace the hesitation sequences which occurred in the initial interview (p.116). The different results stemming from two participants also suggest a relationship between engaging with L1 speakers in a genuinely interactive environment and the acquisition and use of conversational sequences (p.124). An investigation into the knowledge of formulaic language of multilingual advanced learners across their various languages is carried out by Spöttl and McCarthy (2004). Five participants, who achieve a ‘holistic transferring and automatic processing’ level, share a common feature – long-term exposure to an L2-speaking country or intensive contact with L2 speakers either as a family member or as a partner (p.204). The evidence indicates that, in order to effectively acquire and contextually use formulaic language, apart from the language learners’ general linguistic competence, natural exposure to authentic, non-classroom-based, social interactions is also significant (p.217).

C. Individual factors in the learning and acquisition of formulaic language

In addition to enhancing the holistic learning of unanalysed formulaic patterns and increasing exposure to full, natural interaction in the L2-speaking community, some

personal factors relating to the acquisition and use of formulaic language are also discussed by some scholars.

a) learners' perceptive abilities

The first aspect is the learner's perceptive abilities, which includes the necessary attention, awareness and sensitivity to the use of formulaicity in everyday interactions. As Nation (2001) claims, there are three psychological processes which necessarily function in the full command of a new language; they are noticing, retrieving and generating. Only when a word is noticed and understood, can it be identified and retrieved from lexical storage, and can be contextual-appropriately output when needed. One suggestion given by House (1996) on the promotion of idiomatic production of proficient language learners is to raise awareness. Willis (1990, pp.63-64) also points out that the best way to perceive the subtle difference between formulaic and non-formulaic word strings is by observation and imitation.

b) learners' motivation

Another aspect related to individual learners is whether the learner has the need and desire to use formulaicity. As Stevick (1976, p.36) puts it, language learning is better achieved when it is concerned 'with our plans, with our most important memories and with our needs'. Krashen and Scarcella (1978) also point out that, if the use of routines and formulae does not function importantly in L2 acquisition, at least it is useful 'for establishing social relations and encouraging intake' (p.298).

Given that formulaic language is deeply socioculturally rooted as mentioned earlier, effective acquisition, from the personal psychological point of view, also includes the

learner's sociocultural integration into the target language society, that is, sociocultural acculturation.

Acculturation, as defined by Schumann (1986), refers to 'the social and psychological integration of the learner with the target language group' (p.379). Learners' attitudes towards the L2 speakers and their culture, that is, whether the learners appreciate and adapt to the L2 culture and are willing to engage with the L2 community, is emphasised by Aston (1988) as a prerequisite leading to the successful acquisition of L2 language. The degree of learners' social solidarity with the host population, as Furnham and Bochner (1989, p.128) state, links importantly with the attainment of the target language. Individual differences in sociocultural acculturation and learners' acquisition of a formulaic competence is investigated by Dörnyei, et al. (2004), who suggest that successful acquisition of formulaic language is heavily dependent on the learners' breaking out of the 'international ghetto' and actively engaging with the target sociocultural community (pp.104-05).

2.4.5 Summary

Given that formulaic language is one of the subject matters of the current study, a detailed review has been undertaken in this section based on a multiple criteria categorisation, various phonological realisations, different patterns in acquisition and use between L1 and L2, and how to enhance the learning and acquisition of formulaicity by non-L1 language learners. The above analysis in the area of formulaic language leads to the innovative study undertaken by the present author, as further detailed in Chapter 7.

2.5 Conclusion

In this chapter, spoken English and formulaic language are considered in the first part. The aspects of production and decoding of English speech are reviewed, e.g., significant phonological features, various intonation units, continuous flow of connected speech, and different decoding approaches processed by L1 speakers and non-L1 language learners. Then follows an investigation of authentic, natural English speech in terms of facilitating intelligibility, and improving convergence between interlocutors. The last section in this part focuses on the review of formulaic language from the aspect of its multi-criteria categorisation, phonological perception, different acquisition and use between L1 users and L2 learners, and its enhancing of learning and acquisition for L2/EFL learners.

In view of the fact that Chinese learners of English are from a tonal language background, and given the nature of the Chinese language, how English is learned and taught in China, and the question of whether Chinese learners can easily cope with authentic English speech are the main topics reviewed in the following chapter.

3. Spoken Chinese and English Learning and Teaching in China

3.1 Introduction

Following the review of spoken English and formulaic language in the last chapter, a description of Chinese speech and the situation of English learning and teaching in China is presented in this chapter. There are unique phonetic and phonological features in spoken Chinese (reviewed in Section 3.2) which act as an obstacle to Chinese learners of English acquiring English language speech patterns. The current situation of English teaching in China (as examined in Section 3.3) also results in problems for Chinese learners when involved in a real English speaking community (as considered in Section 3.4).

Differences between spoken English and Chinese speech, and different perception and acquisition processes between L1 English speakers and Chinese EFL learners are of great importance to this study, and this provides the basic background for discussion in Chapters 4, 5 and 6, and also forms the basis of the novel work presented in Chapters 7 and 8.

3.2 L1 – A Barrier to Adopting English Language Speech Patterns

This section on aspects of L1 influence which prevents Chinese EFL learners from attuning to spoken English, gives first in 3.2.1 an overview of the Chinese language and its sound system. In 3.2.2 it then examines basic properties of the Chinese language, such as tone, syllable and stress, and intonation. Finally, issues of ‘staccato’ connected Chinese speech and the east-west prosody divide are considered in 3.2.3.

3.2.1 Introduction: Chinese Language and its Sounds

This sub-section deals with the Chinese language and the Standard Chinese sound system. An overview of the Chinese language is first considered, then follows disparities between the Standard Chinese and English sound inventories, with respect to consonants, vowels and rhotacisation. Standard Chinese phonemes and the Pinyin system are also reviewed in this sub-section.

A. Overview of the Chinese Language

The Chinese language, according to Sun's (2006, p.6) analysis, consists of seven mutually unintelligible dialect families. Among them, the Mandarin dialect family is the largest, with more than 70% of speakers from northern and southwest regions of China. Broadly speaking, Chinese refers to all kinds of dialects spoken by the *Han* people (the largest ethnic group in China). But, normally, 'Chinese' or 'Mandarin' narrowly refers to the official language of mainland China and Taiwan, which is called 'Standard Chinese', 'Mandarin Chinese' or 'Standard Mandarin'.

Standard Chinese (SC) is also called *putonghua* 'common speech' in China, which is defined as 'the standard form of Modern Chinese with the Beijing phonological system as its norm of pronunciation, and Northern dialects as its base dialects, and looking to exemplary modern works in *baihua* 'vernacular literary language' for its grammatical norms' (Chen, 1999, p.24). In 1958, in order to annotate standard Chinese sounds and facilitate the promulgation of *putonghua*, a new phonetic scheme was designed and adopted by the Chinese government. The new romanised spelling system is called *hanyu pinyin fangan* 'Chinese spelling system', or simply *hanyu pinyin* or *pinyin*, which

is the standard transcription system of Chinese characters used in China and widely accepted by Chinese language learners outside China (Sun, 2006, p.21).

B. Disparities between SC and English sound inventories

Compared with the English language, which belongs to ‘the Germanic language within the Indo-European language family’ (Yule, 1985, p.168), the Chinese language family is genetically classified as a major branch of ‘the Sino-Tibetan language family’ (Li & Thompson, 1981, p.2). Given the difference in nature of the various language families, some disparities in sound inventory between SC and English language are reviewed as follows.

a) SC phonemes and the Pinyin system

SC, like all other languages, makes use of a set of constituent consonants and vowels to complete its sound system. There are two different approaches among Chinese philologists to defining the number of phonemes in the Chinese Pinyin system – over-analysis and under-analysis (Chao, 1934, p.42), both of which lead to fewer phonemes in Chinese Pinyin system.

Therefore, according to The Hanyu Pinyin Syllabus, a combination of twenty-one consonants, six monophthongs and four diphthongs is defined, which is the standard Chinese Pinyin system learned and used in China. However, some scholars also provide other versions depending on a different analysis. Lin (2007), for example, from the point of view of the discrepancies of SC and English sounds, defines the Chinese Pinyin system as constituting nineteen consonants, five monophthongs and four diphthongs.

b) differences between SC and English consonants

In view of the fact that there are some overlaps between the different analyses, the main focus of this section is to demonstrate the differences between SC and English sound system, and Lin's (2007) version is adopted as the basis of the present analysis. There are nineteen consonants in SC (see Table 1), which seems very close, in number, to English consonantal phonemes; however, the phonetic realisations of these consonants appear differently to their English counterparts. Instead of going into the details on every individual phoneme, only the main differences are reviewed.

	bilabial		labio-dental	dental		post-alveolar		velar	
stop	p	p ^h		t	t ^h			k	k ^h
fricative			f	s		ʃ		x	
affricate				ts	ts ^h	tʃ	tʃ ^h		
nasal	m			n				ŋ	
(central) approximant						ɹ			
lateral (approximant)				l					

Table 1: Consonant phonemes in SC

Source: Lin, 2007, p.50

Firstly, in SC, aspirated and unaspirated stops are separate phonemes, while in English they are different phonetic variants of the same phonemes. For example, in English, /b/ and /p/ are two phonemes, and the phoneme /p/ has two allophones realised as [p] and [p^h]. In contrast, there is no /b/ phoneme in SC, and [p] and [p^h] ([b] and [p] respectively in SC Pinyin system) are distinctive, and function as two separate phonemes. All stops in SC are found only in syllable initial position, and are different to English in that unaspirated stops occur after [s]. Secondly, contrasting pairs of voiced

and voiceless consonants do not exist in SC. SC consonants, except [w], [j], and [ɥ], which are more likely to be realised as vowels, are all voiceless. SC voiceless consonants are either aspirated or unaspirated. Thirdly, the sounds [t], [tʰ], [n], and [l] in SC can be categorised either as dentals, or as alveolars, or both, depending on the speaker. Another difference lies in the fact that SC has three alveolo-palatals, i.e., [tɕ], [tɕʰ] and [ɕ], which involve both the post-alveolar and palatal regions, realised with the front part of the tongue raised higher and close to the hard palate, which are different to those in English. Lastly, in SC only the nasal consonants [ŋ] and [ɲ] are allowed in non-rhotacised syllable final position, realised with the vowel before them.

c) differences between SC and English vowels

SC vowels are also investigated by Lin (2007) (see Table 2). In the monophthong category, firstly, there are two additional vowels which do not exist in English – the high front rounded vowel [y] and a mid back unrounded vowel [ɤ]. Secondly, in SC, there is no phonemic contrast between tense and lax vowels, as [i] and [ɪ] in English, thus, the SC tense and lax vowels do not differentiate meaning in words as in English. Thirdly, in the SC Pinyin system, the letter *e* can be used for both [ə] and [ɤ] vowels, depending on its position in the syllable. And last, three vowels [j], [w], and [ɥ], also called glides, are more properly treated as allophones of their corresponding high vowel phonemes, rather than phonemes in SC. They do not occur in nuclear position, but only in syllable onset position.

	Front		Central	Back	
	Unrounded	Rounded		Unrounded	Rounded
High	i	y			u

Mid			ə		
Low	a				

Table 2: Vowel phonemes in SC

Source: Lin, 2007, p.82

Regarding diphthongs, the SC and English sound systems have quite similar diphthongs, except that [ɔɪ] is absent in SC. There are small disparities mainly laid out as follows. The first point is that the English diphthongs end in a lax high vowel [ɪ] or [ʊ], while in SC, the diphthongs end in a tense vowel [i] and [u], due to the lax vowels' absence in the SC Pinyin system. The second point is that the English diphthong [aʊ] has a central low vowel, but the diphthong [au] in SC has a back low vowel. The last point is that, in SC, there are only falling diphthongs (i.e., the sonority level falls from higher to lower), no rising diphthongs or triphthongs.

d) SC rhotacisation

In addition, there is a phonological change typical of SC called rhotacisation, also known as *erhua*, which is produced by suffixing a diminutive marker *-er* to the finals of words in the spoken language. This phenomenon is nevertheless absent in the English sound system.

As reviewed above, SC has a smaller sound inventory, and there is only a restricted overlapping of phonemes in SC and English sound systems. This basic disparity in the phoneme inventories is one of the main reasons for the problems in pronunciations for Chinese learners, as considered in Section 3.4.1.

3.2.2 Basic Properties of the Chinese Language

This sub-section deals with basic properties in Standard Chinese, such as tone, syllable and stress, and intonation. It also considers whether there are consonant clusters in Standard Chinese, and whether Standard Chinese has stress and how stress is shown, and how intonation is exploited and how tone and intonation interact.

A. Tone

The Chinese language is a tone language (Chan & Li, 2000, p.76), in which changes in the pitch value are used for differentiating the lexical meanings of a word. According to Lin (2007, p.4), tone in SC is taken as ‘the third type of speech element’. In addition to consonants and vowels to form a word, as in English, SC also uses tone to distinguish word meaning. For example, one consonant and vowel combination in SC *ma* can be translated as either ‘mother’, ‘hemp’, ‘horse’ or ‘scold’, depending on different tones (*mā*, *má*, *mǎ*, *mà*), and within each of these tones, there is also more than one homophonous word. The tone feature, therefore, is defined as a unique property of a tone language.

Tone in SC, as analysed by Howie (1976, p.218), is carried by the rhyme (a combination of consonant and vowel, or consonant, vowel and terminal consonant or nasal). Lin (1995) proposes that tone is carried by the nuclear vowel only. Although there is some disagreement between these two analyses, the common notion is that tone is a feature of the lexicon, and is a property of the whole syllable (Cheng, 1973, p.11). Duanmu (2002, p.211) also emphasises that, tone is ‘an integral part of the syllable’, not ‘something extra that can be stripped away’.

Since tone is illustrated by the pitch of the voice, each tone is defined according to both pitch level (how high or low the pitch is) and pitch contour (the pattern of pitch change). The pitch value is used for the transcription of tone. According to Chao's (1968, p.26) analysis, in SC there are four phonemic tones, whose pitch value is based on a scale of 1 to 5, with 5 indicating the highest pitch and 1 the lowest. Tone 1, a high-level tone with a pitch value of 5–5, means that the tone is on a relatively higher pitch level, and no pitch variation occurs within the syllable. Tone 2 is a high-rising tone, with a pitch value of 3–5 i.e., with the pitch movement starting at Level 3 and ending at Level 5. Tone 3 is known as a low falling-rising tone with a pitch contour of 2–1–4, which means the pitch starts to go down at Level 2, then rises from Level 1, and ends at Level 4. Tone 4 is illustrated as a 5–1 high-falling tone, with a high start at Level 5 and a low end at Level 1. In addition, in SC, there is also one neutral tone which typically occurs with some grammatical words and in highly restricted contexts. The phonetic pitch value of a neutral tone is mainly shaped by the extension and influence of its preceding phonemic tone.

B. Syllable and stress

In SC, each syllable generally bears a tone. A syllable is a prosodic unit for carrying tone and stress. SC has a rather smaller syllable inventory and a simpler syllable structure than English. According to (DeFrancis, 1986, p.42), there are 1,277 syllables including tone, or about 398 to 418 syllables ignoring tone. A description given by Duanmu (2002, p.51) is that, SC has only two kinds of syllables – full syllables and weak syllables. Full syllables are mostly lexical words, with the underlying structure of maximally four sounds CGVX, i.e., an initial segmental consonant, a medial (also

known as on-glide), a vowel, and a syllabic terminal (or off-glide). Weak syllables are normally grammatical words, and have a light structure of CV.

One important phenomenon worth focusing on is that, in SC, CG is primarily a consonant, because CG is ‘phonetically realised as a single sound’ (Duanmu, 2002, p.58). Chao (1934, p.42) also gives the reason that ‘there is only one slot in the onset, which C and G must share’. Put another way, the articulation of C and G happens at the same time ‘without leaving any appreciable duration’ for C or G to stand alone. Therefore, SC does not allow consonant clusters in a word, which explains why consonant clusters always result in problems of pronunciation for Chinese learners.

According to Jespersen (1922, p.369), ‘each [Chinese] word consists of one syllable, neither more or less’. That is to say, there is no stressed and unstressed pattern within a Chinese word. Then the question arises how SC exhibits stress, or even whether SC has stress. Stress, like tone, is also a supra-segmental property of the syllable. Shen (1989, pp.59-60) explains that, in SC stress is phonetically manifested by the extension of pitch range and time duration, and sometimes by an increase of loudness. Duanmu (2002) seems to support this notion, in that full syllables are louder, and have greater duration and amplitude than weak syllables (p.135), and a vowel in a weak CV syllable is about half as long as one in a full CV syllable (p.42). All full syllables are stressed (Luo & Wang, 1981, p.135) and have tone, while weak syllables have no stress and generally no tone. However, when each syllable carries a full tone, it is not easy to detect which syllable is stressed. As Selkirk and Shen (1990) note, in Shanghai ‘Chinese native speakers do not feel stress’ (p.315). Chen (2000, p.288) also states that, there is difficulty in obtaining agreement from SC speakers on the relative stress among full syllables.

Given the difficulty in identifying phonological stress in SC, in the literature some scholars claim that SC has no stress. However, Duanmu (2002, p.10) argues that ‘Chinese has stress’, which not only plays a key role in Chinese phonology, but also influences the word length and word order. In contrast with English, Chinese stress is not intuitively clear, especially in the case of full syllables. The main reason is that, SC is a tone language in which the most important phonetic cue for stress – the F0 (fundamental frequency) contour – is ‘taken up for lexical contrast’, so ‘cannot be freely altered to indicate stress’ (p.144).

C. Intonation

Regarding the features of tone and stress, in SC there is another supra-segmental property which spans more than one element – that is, intonation. Intonation appears to be a language universal, even for tone languages. However, in a non-tone language like English, pitch variation is used only for intonation which conveys syntactic and contextual information. In SC, however, the main pitch effort is used for tone in order to keep the semantic meaning of a word, which results in the change of pitch contour for intonation being restricted to phrase or sentence level (Cruttenden, 1997, p.9).

So, how is intonation exploited in SC, and how do tone and intonation interact? Firstly, as Chao (1933) points out, many functions of intonation in other languages are fulfilled in Chinese by the use of particles. SC makes use of particles to indicate syntactic and contextual meanings. These neutral-tone particles are function words, with grammatical meaning only, and are located in sentence-final position. For example, *ma* is a question marker, and *ba* is used for making suggestions, or soliciting agreement (Lin, 2007,

pp.228-29). Apart from sentence-final particles, three basic types of intonation are defined by Shen (1989, pp.26-27). The first is statement mode, which starts mid, moves higher to mid-high, then falls to finish at low pitch. The second pattern is question mode with high-final pitch contour, in which the pitch movement starts at mid-high level, moves to high, then goes into a slight drop before finishing at high or mid-high level. The third pattern of intonation pitch movement is also for a question, but with low-final pitch movement, in which the pitch starts at mid-high, moves to high, then drops to finish at low.

Although in SC pitch variation is used for both tone and intonation, the basic contour of a tone remains intact and recognisable, and easily perceived. The main reason is tone and intonation function at two different levels. Based on Chao's (1933) analysis, intonation expression is superimposed onto word tone, by which the overall pitch range of an utterance may be raised, lowered, expanded, or compressed, but the tone pitch contour for each word is retained. Simultaneous interaction between tone and intonation is supported by the study of He and Jing (1992), in that the question intonation raises the pitch level of the whole utterance without changing the distinctiveness of word tones.

Influenced by the crucial role which tone plays in SC phonology and semantics, when engaging in English communication, Chinese language learners frequently concentrate on pronouncing individual words, ignoring English stress and intonation patterns. This causes the problems confronted by Chinese learners in their encoding and decoding processes, which is further examined in Section 3.4.1.

3.2.3 'Staccato' Connected Speech with Clear Word Boundaries

This sub-section considers aspects of 'staccato' connected speech in Standard Chinese. On the one hand, some co-articulatory processes which contribute to connected speech are reviewed. On the other hand, some special properties obstructing the natural flow of Chinese speech, e.g., sonority sequencing principle, zero onset and dissimilation, are then examined. The east-west prosody divide is also dealt with in this sub-section.

A. Co-articulatory processes contributing to Chinese connected speech

Similarly to English, connected speech also occurs in SC spoken communication. Various kinds of segmental and tonal processes realised in different contexts are investigated by some scholars. A thorough analysis is given by Lin (2007), in terms of assimilation, vowel insertion, syllable contraction, *erhua* and tonal changes.

a) assimilation

As examined earlier, in SC there are four phonemic tones (high-level, high-rising, low falling-rising and high-falling), which are also called four 'citation tones' (Lin, 2007, p.95). When involved in the flow of natural speech, a citation sound is generally induced by some contextual factors and causes some changes. Assimilation is seen as one of the most well-known phenomena. When a sound becomes more similar to adjacent sounds or some sounds within the same syllable or word, it is called assimilation. Four cases of assimilation are discussed by Lin (*ibid.*, pp.150-65). This includes the phenomenon of consonant weakening, vowel reduction, rime reduction and vowel devoicing. The assimilation phenomenon normally occurs in a syllable with a neutral tone. The main reason is to facilitate faster and smoother transition between sounds so as to achieve ease of articulation.

b) vowel insertion, syllable contraction and *erhua*

Apart from the most frequent assimilatory changes, vowel insertion is also identified in SC connected speech (Lin, 2007, pp.175-80). Another phenomenon is syllable contraction, which refers to the process of two syllables in a sequence being merged into one syllable by combining some segments and tones from both syllables (ibid., p.180). An r-suffixed syllable, also called *erhua* (as mentioned in Section 3.2.1), is another case undergoing morphological changes, as detailed by Lin (ibid., pp.182-89).

c) tonal changes

The above are some general segmental processes under the principle of co-articulation, and given that SC belongs to the tonal language family, articulatory processes consequently exhibit some tonal changes. ‘The change of tone due to the influence of adjacent tones’ is called tone sandhi (Lin, 2007, p.100). The word ‘sandhi’ originally comes from Sanskrit, and means junction, connection, combination, or liaison (Chen, 2000). Tone sandhi in SC refers to the phenomenon of tonal alternations when syllables are connected in natural speech. The most productive tone sandhi is Tone 3 sandhi which involves consecutive Tone 3 syllables. The general rule described by Sun (2006, p.41) is that, in a sequence of two Tone 3s (low falling-rising), change the first Tone 3 to Tone 2 (high-rising). Another tonal change also applied in SC is Tone 2 sandhi, as described by Chao (1968, pp.27-28) and Yip (1980, p.291).

B. Special properties obstructing natural flow of Chinese speech

The main co-articulatory processes in SC are briefly examined above. These segmental and tonal changes ease the articulation and contribute to natural continuous connected

speech. However, having in mind that Chinese language is a tone language, there are also some special properties which impede SC speech in achieving real natural flow.

a) sonority sequencing principle

The first block can be explained by the ‘sonority sequencing principle’, which requires that the syllable have increasing sonority before the nucleus and decreasing sonority after the nucleus (Selkirk, 1982). Given that each Chinese character represents only one syllable, and with the basic syllable structure of CV (consonant + vowel), this leads to a general tendency for every character to follow the pattern: lower sonority plus high sonority. The rhythm of a sequence of CV syllables with the alternation of low sonority and high sonority results in clear syllable boundaries in a natural flow of connected speech.

b) zero onset

Another prosodic factor is due to the ‘zero onset’ in SC, which means when a syllable does not begin with a C, a G, or a CG combination, there is still an articulatory effort in the onset (Li, 1966). According to this analysis, in SC a zero-initial syllable does not attract the coda consonant of the preceding syllable to become its own onset, as English does, since there is already an onset in the zero-initial syllable. The ‘obligatory’ zero onset (Duanmu, 2002, pp.82-83) results in the relatively fixed C + V syllable structure and also causes the clear-cut boundary between words produced by Chinese learners when they speak English.

c) dissimilation

The third fact which prevents the continuous flow of SC speech is the phenomenon of dissimilation (Lin, 2007, p.145). Dissimilation occurs when a sound becomes less similar to a neighbouring sound or another sound within the same syllable. The tone sandhi process as examined earlier, especially Tone 3 sandhi and Tone 4 changes for the word *yi* ‘one’ and *bu* ‘not’, is accounted for as a kind of dissimilation process which better distinguishes two originally identical tones from each other. The dissimilation process therefore also contributes to unevenness in SC connected speech.

C. East-west prosody divide

The two graphs below (Campbell, et al., 2008) show the contrast of an identical SC utterance produced by an L1 Chinese speaker at normal speed (see Figure 4) and nearly twice the speed (see Figure 5). It can be seen that, by contrast with a natural blurred outcome in English (as considered in Section 2.2.3), in SC despite the different speeds of production, the spectrogram representations of both utterances show a remarkable similarity, since the intrinsic tones for each word must be kept in order to maintain lexical integrity, which is the basic *sine qua non* for all the co-articulatory variations which occur in connected speech. This kind of discrepancy easily results in Chinese learners sounding ‘staccato’ in their delivery of natural English speech, which is therefore an east-west intonational gap which Chinese learners of English need to learn to cross.

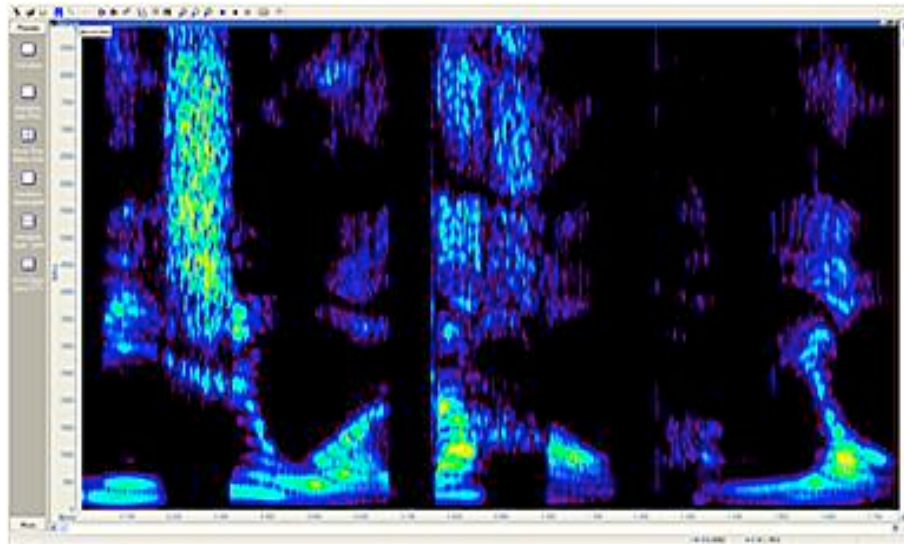


Figure 4: SC utterance produced at normal speed

Source: Campbell, et al., 2008

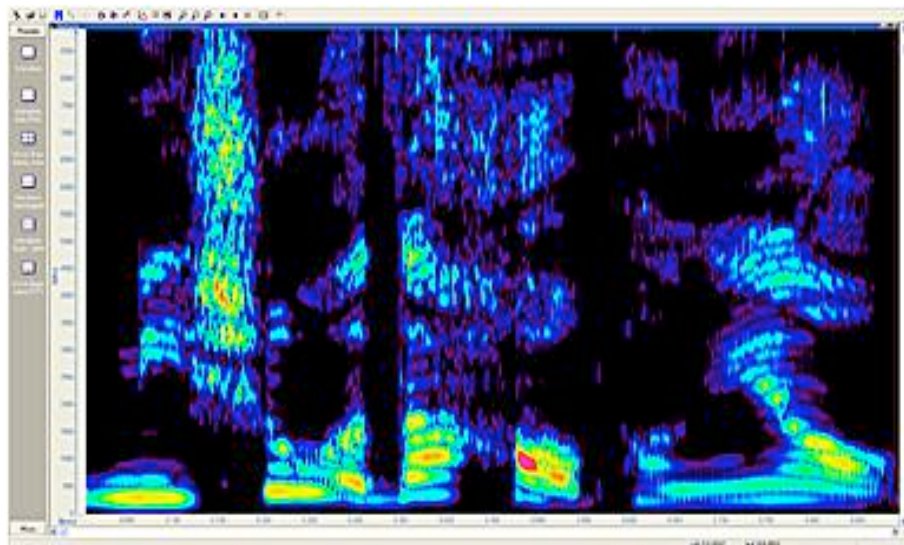


Figure 5: SC utterance produced at a faster speed

Source: Campbell, et al., 2008

3.2.4 Summary

In this section, the basic characteristics of SC are reviewed in terms of its phonemic sound system, simpler and fewer syllable structures, tone patterns and the interaction with intonation, and the segmental and tonal realisations of connected speech. The disparities between Chinese and English, especially the unique property of SC, in which

tones are embedded in each lexical word and pitch variants are used for lexical distinction, undoubtedly create difficulties for Chinese learners of English when they become linguistically involved in an English-speaking community. Only when there is an awareness of the differences between these two languages, as Huang (2010) points out, can appropriate pedagogies be adopted to improve Chinese learners' abilities in coping with the natural rapid stream of English speech.

In the following section, the current situation of English learning and teaching in China, e.g., a static learning and teaching process, limited learning-teaching resources, inefficient pronunciation and listening teaching, and insufficient natural authentic English exposure, is further considered.

3.3 Constraints in Current Educational Situation – How is English Learned and Taught in China?

This section reviews the situation of English learned and taught in China under the current educational constraints. It first looks, in 3.3.1, at the static learning and teaching process. Then, 3.3.2 examines the phenomenon of large classes and limited teaching/learning resources. The aspects of inefficient pronunciation teaching and listening are considered in 3.3.3. Finally, aspects of insufficient exposure to natural English speech are dealt with in 3.3.4.

3.3.1 Static Learning and Teaching Process Influenced by Traditional Cultural and Education System

Before 1978, China appeared completely closed to the outside world. International contact or exchange was not encouraged. The State controlled the education orientation and education structure, and nowadays these conformist traditions still have an impact on education. Details are reviewed under the following two aspects.

A. Teacher-centred learning

As discussed on the English Education website, the examination-orientated educational pattern is un-educational, and great pressure is put on teachers to teach to a specific syllabus and to prepare for structurally based, traditional exams. A teacher-centred learning and teaching approach dominates teaching activities, as argued below.

As pointed out on the Ninth Software website (p.1), the whole class dynamic shows a static and boring learning-teaching model characterised by the teacher's large input and the students' passive intake. About three quarters of class time is taken up by the teacher explaining and presenting language knowledge (Feng, 2003, p.2), and students get little or no chance to practise and improve their communicative skills (Xiao, 2004, p.4).

As pointed out on the Secondary School English Teaching Resources network (p.1), the traditional educational approach focuses on a normed, prescriptive education mode. The same teaching methodology is used with all students, which results in students' individual needs being neglected. According to the present author's personal learning experience, memorisation and re-production of the standard answers is what is required. Unique opinions are often regarded as out of the ordinary and will run the risk of being met with negative comments on the part of the teacher. As Zou (2008, p.1) points out, the teacher is typically ready to interrupt and correct his students, even in the rare spoken communications.

Under the traditional ‘Grammar-Translation method’, as Gu (2003, p.4) points out, ‘... in some places [in China] English was being taught like a dead language’, and there was often ‘no concentration on communication’. Real interactive features in authentic classroom settings, as investigated by Walsh (2009), are uncommon in the traditional teacher-dominated classes. This explains the fact that Chinese language learners ranked lowest in the oral test of IELTS (International English Language Testing System) in 2004 and 2005.

B. Conservative social and cultural environment

In addition, the conservative social and cultural system also greatly affects the concept of learning and the process of cognition of students. As Cortazzi and Jin (1996) state, ‘Western and Chinese cultures of learning sometimes weave past each other without linking’ (p.10).

Firstly, Chinese culture is regarded as ‘with a long tradition of unconditional obedience to authority’ (Liu, 1998, p.5). In China, teachers are seen as the holders of knowledge and authority. Out of respect, as Littlewood (2000, p.31) notes that, ‘[in Hong Kong] no one wants to voice their opinion and challenge what lecturers say’.

Secondly, ‘[t]raditional Chinese culture places a very high value on learning’ (Chang, 2001, p.322). Chinese students generally prefer a solitary learning process, rather than joining with their peers in groups (ibid.). The approach adopted by Western teachers of teaching English through games and communications is thought as ‘wasting too much time’ (Doyle, 2006).

Thirdly, in China, the teacher-centred learning and teaching model puts more responsibilities on schools and their teachers. Teachers are often present in the classroom. ‘Students will be accustomed to finding staff ready, willing and able to talk to them privately – and at length, without an appointment’ (Turner, 2001, p.46). As Stone (2008) points out, ‘Chinese students tend to rely on the teacher, which is unhelpful for students’ autonomous learning.

3.3.2 Large Classes and Limited Teaching/Learning Resources

This sub-section first deals with the phenomenon of large classes. It considers its negative effect on pronunciation teaching and the development of individual communicative competence. This sub-section then examines the phenomenon of limited teaching/learning resources, on the aspects of teachers’ professional status, teaching materials, resources and equipment, and scarcity of qualified teachers.

A. Large classes

A survey published by the Ministry of Education of China in 2003 shows that there are 3,221,000 large classes (56-65 students per class), which make up 27.65% of all junior secondary schools. There are also 272,400 super-large classes with more than 65 students per class.

With regard to the teaching of large classes, Lu (2007, p.1) emphasises that it is impossible to give every student the chance to perform an efficient interaction in a large class. Moreover, in a large class the teacher cannot possibly listen to every student’s pronunciation and instruct individually. The large class is therefore not suitable for English learning and teaching in primary or secondary schools, and especially not for the learning and teaching of pronunciation.

In addition, as pointed out by Xu (2004, p.2), the traditional large class model promotes a ‘dull’ and ‘stressed’ teaching style, which undoubtedly neglects students’ different learning needs, and is detrimental to develop students’ individual cognitive abilities in learning and using language, particularly listening and speaking skills.

B. Limited teaching/learning resources

The large class phenomenon is the inevitable result of limited teaching/learning resources. Four aspects are reviewed as below.

The first aspect considered by Burnaby and Sun (1989, p.228) is the teachers’ professional status. As non-L1 language speakers, Chinese teachers often lack sufficient ability in choosing teaching materials. Moreover, they also feel less confident in controlling the dynamic class and dealing with spontaneous questions given by students. Another reason is linked with social acceptability. Traditional teaching methods and examination-oriented learning motivation results in those teachers who deal with spoken English undervaluing their work and decreases their enthusiasm and innovation.

The second aspect is teaching materials. As reviewed by Burnaby and Sun (1989), most of the materials are produced in China and focus on vocabulary learning and language analysis. Therefore students have very few chances to access natural foreign language materials created by L1 English speakers. Moreover, nearly all the teaching materials are recommended by the central government, teachers have less control over the adaptation of teaching content which results in that teaching cannot meet individual needs. Another disadvantage, as also pointed out by some Chinese scholars, is that the

development and renewal of a course-book is relatively slow. Various versions of traditional teaching course-books had been used for decades. Until 1993, relatively new types of teaching materials came into use under pilot exercises in some provinces.

The third factor which prevents learning and teaching activities is limited resources and equipment. As mentioned by Burnaby and Sun (1989), even in third level colleges, there is a lack of audiovisual equipment, photocopiers, etc. A personal communication (Chen, 12 June 2008) with a lecturer from the Teachers Training School shows that the phenomenon of shortage of funding and equipment is severe in primary and secondary schools in her city. Where facilities do exist, exploitation is inadequate. No funds are available other than via the local education authority, and the rural situation is even worse.

Another point is number of qualified teachers. In 2005, there are 176.7 million English learners in China (as estimated by Graddol, 2006, p.95), which requires a highly qualified teacher pool. However, of 550 thousand English teachers in secondary school, only 80.4% of them are qualified to teach in junior secondary and 55% in senior secondary schools (Bao, 2004, p.4). The situation in primary schools is even worse. The need for extra teachers is met by drafting in staff unqualified in linguistically and in methodology.

3.3.3 Inefficient Pronunciation Teaching and Listening

Due to the prominent position of the traditional grammar-translation model in English language learning and teaching as reviewed earlier, spoken language teaching in China has had little attention paid to it and has been neglected in the teaching syllabus. This sub-section looks at this in greater detail.

A. Insufficient teaching of pronunciation

As pointed out by Han (2007, p.2), 'pronunciation teaching, like single phoneme instruction, only happens in short teaching blocks'. Chinese students' pronunciation proficiency as a whole does not seem to match the requirements of the syllabus (Wang, 2005). There are four main problems, as considered below.

The first problem, discussed by Hu (2005, pp.2-3), is reflected in the pronunciation syllabus in junior secondary school. The whole process is split up into several phases and lasts too long, which results in a very passive learning situation and means that students cannot read out new words independently until they have completely finished learning the phonemes. The delay of IPA teaching is also discussed by Qin and Lian (2008, p.1).

The second factor is that the teaching phases of the alphabet and phonetics are isolated and separated, and the teaching of phonemes does not receive sufficient attention. Students have little knowledge about phonemes. The link between spelling and phonetics is not established, which leads to the mechanical memorisation of words letter by letter. The survey undertaken by Hu (2005) shows that 53.5% of the spelling mistakes results from incorrect pronunciation.

The third aspect, discussed by Yang (2008), is that the supra-segmental level of pronunciation teaching is not given enough emphasis. Insufficient instructions and practice are given to students which results in 80% of senior secondary school students having difficulties with stress, rhythm, and intonation (p.2).

The last factor is that much emphasis is put on the teaching of knowledge, while the training of practical ability is neglected. According to Xiao's (2004, p.2) survey, only 28.7% of students can pronounce new words (using IPA symbols) correctly in the first academic year in junior secondary school.

B. Inefficient teaching of listening skills

A similar situation, or even worse, applies to the teaching of listening skills. There is little mention in the literature of the teaching of listening and learning, since there is almost no effective listening teaching. The main reasons are as follows.

Firstly, since listening test is only a small part in students' written examination (20 points out of 150 in junior secondary school), teachers prefer to spend their time explaining grammar rules, vocabulary usage and helping students with written exercises. The listening class is seen as the easiest teaching activity. Generally, listening teaching tends to exist at the presentation level, there is no real linguistic engagement involved.

Secondly, as Sun and Zhan (2001) comment, 'explaining new words, playing the recording and checking the answers' are the traditional steps in the listening class. This is also the experience of the present author when teaching listening classes in China. The teacher does not normally spend time in the post-listening phase on giving adequate instructions and helping students find out where and why the comprehension failed. As Brown (1990, p.8) points out, the listening process is more like 'testing' than 'teaching'.

The next factor, discussed by Fu (2008), is inadequate procedures in listening training. Teachers normally do not have a systematic teaching plan. The listening activities may be arranged too close to each other, or may be too isolated. Listening training is not allocated sufficient time, and in addition, the time for each session is either too long or too short. As regards the listening materials, there is also a problem in choosing realistic content. Besides, the totality of the training materials is neither continuous nor progressive.

The last aspect, as emphasised by Xu (2007), which is also significant is that, the listening training process is not scientific, or efficient. Firstly, the general length of each tape extract should not be too long, which overloads students' memory capacity. Secondly, the style of training is too static. The third prerequisite often absent in Chinese listening classes is adequate instructions before the listening phase. There is no relevant information given to students and on which students can adopt a top-down approach or make intelligent inferences. Another step is that there is no systematic listening training. In addition, no listening assignment is given after class to help students to continue their progress. Listening training mostly occurs in a limited classroom setting and is built upon answering comprehension questions (Hong, 2008, p.2).

3.3.4 Insufficient Exposure to Natural Authentic English

The main environment in which Chinese students experience and learn English is within a classroom setting. This inevitably results in insufficient exposure to natural English, which is mainly examined under the following four aspects.

The first factor is insufficient classroom teaching time. According to the regulations issued by the Ministry of Education of China in 2002, the time for classroom English teaching for primary and junior secondary school occupies only 6-8% of all teaching time (Bao, 2004, p.2). The implementation of the official policy carried out by different schools produces an even worse situation. A report shows that, only 11.38% of twenty-three primary schools in Yibin city offer three classes per week (p.2).

For classroom-based language learners, classroom teaching activities, especially the language used by teachers, are the main way in which students are exposed to spoken English. A survey carried out by Hu (2008) shows that 93% of twenty-eight teachers in the primary schools investigated teach English through Chinese. Another report shows similar results on the English teaching in senior secondary school.

The next aspect which restricts students' exposure to a natural English-speaking environment is that there is little or infrequent use made of multi-media equipment. Given that the current English class revolves around the teacher, as reviewed earlier, students cannot easily access multi-media teaching presentations. According to Hu's (2008, p.2) survey, only 5% of teachers use multi-media equipment due to lack of availability. Xiao (2004, p.6) reports that 27.7% of the students questioned think that multi-media teaching materials are only presented to them in observation classes. Another report (2006) also claims that electronic devices are used in only 32% of the 118 observation classes investigated (p.2).

The last factor is students' reading training after class. As presented by Jiang (2000, p.8), only 20% of the students complete the auxiliary reading materials. A questionnaire (2007) given to students shows that 43% of their teachers seldom give them reading

tasks (p.2). Burnaby and Sun (1989) also state that third-level college students participating in their survey have very few opportunities to access authentic materials in English (pp.227-28).

3.3.5 Summary

The current situation of English as it is learned and taught in China is reported in this section. The main barrier for Chinese learners to perceive and acquire real, authentic English speech lies in the traditional, static teaching approach and the teacher-fronted teaching model. Due to the limitations in learning and teaching conditions and pedagogies, English learned and taught in China cannot meet the real needs of ideal English learning and teaching. These disadvantages in current English learning and teaching inevitably result in the failure of Chinese language learners to cope with real, authentic English speech.

The discrepancies between Chinese and English and the detrimental situation Chinese students are facing in learning English, result in problems when students become involved in a target language community, which is further outlined in the following section.

3.4 Problems Encountered by Chinese Language Learners in English Native-speaking Communities

Based on the review above in Chapters 2 and 3, some problems encountered specifically by Chinese learners when involved in a target language speaking environment are considered in this section. Issues of Chinese learners' mispronunciation and misprosody, which confuse L1 listeners and cause problems of intelligibility, are examined in 3.4.1. Issues of decoding natural English speech are then dealt with in 3.4.2.

*3.4.1 Issues for L1 Listeners with Chinese Speakers'
Mispronunciation and Misprosody*

‘The more differences there are, the more difficulties the learner will have in pronouncing English’ (Kenworthy, 1987, p.4). Given the considerations examined earlier, there are certain discrepancies between spoken English and Chinese. These disparities can cause problems of intelligibility for English speakers when listening to Chinese learners.

The main problems, based on the analysis of speakers of Cantonese (spoken in Hong Kong) and Hokkien (spoken in Singapore), are detailed by Kenworthy (1987, pp.128-31). Some of the descriptions are specially emphasised in this sub-section from the point of view of SC, at both segmental and supra-segmental levels.

A. Issues at segmental level

a) consonants

Firstly, as also investigated by Ma (2001), SC has neither of the ‘th’ sounds (voiced or unvoiced). Chinese speakers tend to substitute /z/ for both of them. Secondly, there is no /v/ sound in SC either. It is easily substituted by /w/ by SC speakers in word-initial position, and also be substituted by /f/ or completely deleted when it occurs at the end of a word. The next problem is in dealing with word-final consonant, some of SC speakers tend to add an extra vowel (Hewings, 2004, p.234) or simply delete it (Ma, *ibid.*). Furthermore, since there are no consonant clusters in SC as examined in Section 3.2.2, consonant clusters or sequences are clearly a difficult area for many SC speakers, especially when they occur in contractions and grammatical endings (Li, 2007, p.37). Their dominant coping strategy, as Tajima, et al. (1997) observe, is either to

delete one or more consonants from the group, or to insert a short epenthetic vowel between consonants in a cluster at the beginning and the end of words. Another phenomenon also pointed out by Hewings (*ibid.*) is that Chinese language learners tend to pronounce the strong form of words instead of the weak form when the word should be non-prominent.

b) vowels

Due to the slight differences between English and Chinese vowels, as reviewed in Section 3.2.1, there is no big difficulty for SC speakers in producing these sounds. One problem however is the vowel length. As addressed by Kenworthy (1987, pp.129-30), Chinese speakers tend to shorten a diphthong, and use a similar monophthong instead.

B. Issues at supra-segmental level

a) rhythm, stress and intonation

Influenced by the crucial role which lexical tone plays in Chinese phonology and semantics, as mentioned earlier, when engaging in English communication, Chinese language learners usually concentrate on pronouncing individual words, ignoring English intonational patterns. In addition, as Juffs (1990) points out, the syllable structure of Chinese, i.e., the relatively fixed C + V distribution (see details in Section 3.2), also affects stress assignment, and therefore impacts the rhythmic pattern of the utterance. The English stress pattern of a word is part of the stored code of the word's profile and it is crucial for the word's identity. For example 'record' and 're'cord' have different word classes depending on where the stress falls. But there are 'no SC words that are distinguished by stress alone' (Duanmu, 2002, p.134). Therefore, sometimes Chinese speakers do not stress one syllable more than the others, or stress

the wrong syllable, which easily causes difficulties for English speakers in identifying the mis-stressed word under the wrong stress pattern (Brown, 1990, p.51), or sometimes Chinese learners stress all the words equally, which makes English speakers struggle to recognise the prominent parts of the information.

b) connected speech

Given that the relatively fixed C + V syllable structure and clear-cut word boundaries in Chinese language, as Kenworthy (1987, p.18) points out, ‘Chinese learners often do not use smooth transitions’. Linkage is one of the greatest problems for Chinese speakers. A lack of skills in this area results in Chinese speakers’ speech sounding ‘staccato and jerky’ (ibid.).

3.4.2 Issues for Chinese Learners in Decoding English Speech

Three perspectives are mainly dealt with in this sub-section, i.e., inappropriate pronunciation exposure, insufficient listening and improper listening strategies.

Firstly, Chinese learners are exposed mostly to the English of their teachers during their class time. The English teachers always produce every segment clearly, even unstressed syllables. This kind of pronunciation exposure results in Chinese learners being accustomed to listen to clear articulation of English speech. However, this model, as argued by Brown (1990, pp.46-47), is exclusive to language learners. In reality it never happens between English L1 speakers. Constant exposure to this sort of artificial English speech presents Chinese students with some difficulties in understanding natural spoken English.

A second reason is because of the improper listening goals. As examined in Section 3.3.3, both teachers and students focus on the outcome of listening, rather than the process. 100% correct comprehension is the goal which Chinese students expect in their listening activities. In reality, however, this is not achieved, even between L1 English users. This causes Chinese students to ‘experience panic’ in listening to natural English speech (Brown & Yule, 1983a, p.59).

Apart from an inappropriate exposure to pronunciation and an insufficient listening model, listening strategies are also problematical. As considered in Section 2.2.4, Chinese learners rely much more on the ‘bottom up’ process when listening, and have difficulty in progressing to the L1 ability of ‘top down’ processing of speech. However, due to the nature of connected English speech, it is impossible for Chinese learners to capture every sequence of the signal; thus, the decoding process easily breaks down once they encounter a new word or an unfamiliar expression. This kind of ‘bottom up’ activity is necessary, yet as Brown (1990, p.10) points out, ‘it is insufficient on its own’.

L1 speakers are active both as listeners and processors. They mostly use the top down process to ‘figure out a particular word’ (Kenworthy, 1987, p.15) or ‘repair the defective signals’ (Tatham & Morton, 2006, p.197). This reconstruction process is much more complex than a simple process of straightforward decoding of the signal. In fact, information is not encoded in the signal. The listener merely uses the acoustic signal as a ‘trigger’ to assign meaning to the spoken sequence (Tatham & Morton, 2005, p.105). As for Chinese learners, inadequate exposure results in a poor level of background knowledge and a consequent lack of contextual clues. Consequently, unrealistic hypotheses are made based on their L1 experience, and the understanding process is interrupted. The absence of an ability to assign correct words to appropriate signal

segments based on their personal knowledge is a crucial block for Chinese learners in rendering the speaker's intentions.

Making inferences is another kind of intelligent guessing ability which L1 speakers use to gain a good deal more information than is actually contained in the original message (Brown, 1990, p.155). The main problem for Chinese learners is that they tend to adhere to the original utterances and not go beyond that, since they are unsure of the shared experience between the speaker and themselves.

3.4.3 Summary

In this section, the main problems encountered by Chinese learners when involved in natural English-speaking community are reported. On the one hand, due to the differences between English and Chinese languages and influenced by L1 transfer, Chinese learners inevitably experience some pronunciation difficulties, which result in failure of intelligibility in their oral English communication.

On the other hand, the relatively poor level of comprehension ability of Chinese language learners is also considered in this section. Given the inefficient pronunciation teaching and listening process in Chinese classroom setting, and insufficient exposure to natural English speech, Chinese learners mostly employ a bottom up approach and fail in arriving at a reasonable understanding by narrowing down the possible expectations and making inferences.

3.5 Conclusion

This chapter mainly focuses on the description of Chinese speech, and English learning and teaching in China. Given that the Chinese language is characterised as being a tone language, its sound system, unique tone features, and non-connected speech, are

outlined initially from the perspective of why they impede Chinese learners in adopting spoken English patterns. Then an overview of English as it is learned and taught in China is given from the perspective of static teaching pedagogy, limited learning-teaching conditions, inefficient pronunciation teaching and listening process, and insufficient natural English exposure. Finally, the problems encountered by Chinese learners in their oral English communication are considered.

Knowledge and understanding of the above review in Chapters 2 and 3 is essential to this study, which provides an overview of key features of spoken English and the basis of the most common linguistic difficulties facing Chinese language learners when confronted with a native English-speaking community. Building on these two chapters, a further discussion follows in Chapter 4, of what Chinese learners really need in order to become proficient in natural spoken English communication.

4. Discussion of Literature Review

4.1 Introduction

In the previous chapters, a review of spoken English and Standard Chinese, and the current situation of English as learned and taught in China are considered.

This chapter offers a discussion of the literature review, specifically some issues in current research, such as the east-west prosody divide, formulaic language and its phonological realisations, the need to expose Chinese EFL learners to authentic spoken English, and some pedagogical suggestions on access to efficient ICT (Information and Communications Technology) technologies.

4.2 Issues in Current Research

Referring to current research in EFL in general, some issues relating to learning difficulties were identified in Chapters 2 and 3, as follows:

A. East-west prosody

The discrepancies between spoken English and spoken Chinese, especially the unique tone property and the fixed C + V syllable structure, present Chinese learners with identified difficulties in adopting spoken English patterns and achieving natural English communication. Therefore how to help Chinese learners to overcome the east-west prosodic divide and acquire natural, informal English intonation patterns emerges as one of the areas of the research to be carried out in this work. This relates to RQ1.

B. Formulaic language and its phonological realisations

Although a very considerable proportion of formulaic expressions is used by L1 speakers, very little formulaic language is learned and acquired by EFL learners, due to current classroom pedagogical practice. Thus, it is important to concentrate on formulaic language and help language learners to learn and use formulaic expressions appropriately in an English-speaking society.

Research in the area of formulaic language is relatively new. As argued by the present author, there is no evidence in the literature so far to show the link between the different categories of formulaic language and their relevant phonological characteristics. Therefore, this is a gap the present author aims to bridge in the study of formulaic language.

Most existing phonological analysis on formulaic language is based on recorded spoken language with its own phonological characteristics (but not necessarily dialogue). The analysis of the current research, on the other hand, is based on data drawn from natural, spontaneous, dynamic dialogues.

Of the many phonological characteristics, the present author mainly focuses on the speed of delivery and pitch range since these are two of the parameters which most influence the realisation of formulaic language and exemplify different communicative values. In addition, Chinese EFL learners tend to deliver flat intonation patterns when engaged in English communication. Therefore, concentration on the variety of pitch ranges mostly associated with the speed of delivery can benefit Chinese learners in learning western intonation patterns. The above relates to RQ1.

Based on the discussion in Section 3.4, the present author proposes that Chinese EFL learners need to be exposed to authentic English speech.

In addition, by exposure to a ‘virtual language community’ (Johnstone, 2007), language learners can not only perceive authentic language as it is used by L1 English speakers, they can also build up background knowledge. The present author agrees with Kuiper (2004, p.39), who addresses the role of exposure from the point of view of acquisition of formulaic language, and states that formulaic expressions cannot be taught, but only occur ‘when there is the chance for exposure’.

Considering the analysis above, and given that the long-standing ‘grammar-translation’ teaching model is still in operation in English classroom learning and teaching in China, it is the goal of the present researcher to use highly natural L1-L1 informal recordings and concentrate on the prosody of the formulaic sequences found in such speech in abundance. This relates to RQ3.

D. Access to effective ICT language learning technologies

Restricted by the very limited classroom teaching time and the traditional teaching pedagogy, increased exposure in class seems unlikely and impractical. Access to ICT technologies, therefore, becomes an ideal option.

As early as 1997, Graddol (1997, pp.30-31) emphasises that those technologies associated with computers and communications would bring major changes in culture and language. Similarly, as Huang (2007, pp.1-2) points out, a computer-assisted interactive teaching environment is the best way to implement an individual education,

in which learners can choose appropriate learning methods to improve their autonomous study ability.

The above discussion shows that linguists are becoming increasingly convinced of the advantages of ICT technologies in assisting language learning, especially for Chinese EFL learners, which is one of the gaps the present research aims to bridge. This leads to the investigation of RQ2 and RQ5.

The remain research question, RQ4, relates to Tests 1 and 2, and is described in Chapter 8.

In the following chapter, Chapter 5, firstly, some current TELL (Technology-Enhanced Language Learning) tools are briefly reviewed. Then an overview is given of DIT's innovative language learning technologies, e.g., the slow-down algorithm, the approach adopted in the project *Articulate!* with respect to segmentals, the development of the Dynamic Speech Corpus and its natural, spontaneous, authentic L1-L1 speech assets.

5. Current Research into TELL Tools

5.1 Introduction

The review and discussions in Chapters 2, 3 and 4 show that what Chinese language learners need in order to avoid the problem encountered in natural, informal English communication and achieve good interlocutory skills, is to increase their exposure to real L1 English speech. However, the current limited possibilities for English learning and teaching in China obstruct this kind of exposure to some extent and make efficient exposure unattainable. Thus, ICT (Information and Communications Technology) English learning and teaching programmes seem to be an ideal option for Chinese EFL learners.

In this chapter, firstly, some conventional TELL (Technology-Enhanced Language Learning) tools are briefly outlined in 5.2; in terms of their main functions, advantages and disadvantages. Then, in 5.3 an overview of some speech technologies developed by DIT is given. These innovative language learning technologies are devised specifically to help with the identified linguistic and pedagogical deficiencies.

5.2 Traditional TELL Tools

As Bush and Terry (1997) point out, from ‘curricular objectives to lesson planning ... from teacher training to software applicability, there will be no aspect of foreign language learning that will not be influenced by the technological revolution’ (p. xiv). Given the overall advantage of technology-enhanced language learning, many software/tools are developed to facilitate English learning and teaching. Some of them are briefly reviewed in the following, from the perspective of general English language

learning, pronunciation learning and teaching, listening comprehension, and speech analysis. Audio tools developed in China are also discussed below.

A. General English language learning tools

Tell Me More is one of the global online multi-lingual language learning programmes, which covers not only the training of essential skills in listening, speaking, reading and writing, but also vocabulary, grammar and culture. Its apparent advantage is that it provides a systematic English learning syllabus and makes English learning engaging and interactive. However, its drawback is also widely acknowledged: that it does not really provide any meaningful feedback on the pronunciation accuracy of language users, even though graphical wave forms are presented.

B. Pronunciation learning and teaching tools

Apart from *Tell Me More*, there are also some other language learning tools, especially those on pronunciation learning and teaching, which integrate automatic speech recognition (ASR) technology in the learning system. The reliability of ASR technology is investigated by Kim (2006), via *FluSpeak*, software for learning and teaching pronunciation. The results show that the overall accuracy of recognition is still not as accurate as human analysis, in which the correlation coefficient at word level is not high and near zero for intonation level (p.330). This implies that automatic speech recognition software may be only useful for learners to practise some aspects of pronunciation. This, as argued by Bacalu and Delmonte (1999), in general seems to be a shortcoming in most pronunciation-recognition tools, for example, *PLASER* and *Fluency Pronunciation Trainer* (Heffernan & Wang, 2007).

A pronunciation scoring algorithm, presented by Witt and Young (1998), is shown to calculate accurately a language user's pronunciation at phonemic level, in which both individual mispronunciations are detected and a general assessment of which sounds tend to be mispronounced are indicated as well. However, it cannot be used as an independent learning tool, but has to be embedded within an interactive language teaching system containing modules for, *inter alia*, error analysis, pronunciation feedback and assessment.

There are also some pronunciation learning tools specifically focusing on English vowel training. *Electronic visual feedback*, investigated by Lambacher (2001), is a computer-assisted training tool for accent reduction. The deviation between the acoustic features of the master's production and the learner's are displayed visually, in which the real-time evaluation and assessments of learner's mistakes and progress can be provided. As it is not devised originally for language learning, the interface, therefore, is not really user-friendly. It is often too complicated for a user to understand and interpret, and a basic knowledge of acoustic phonetics is needed. This is also the drawback for the development undertaken by Brett (2004) on an application of the *PRAAT* programme to calculate and plot the formants on a graph to provide learners with real time feedback on their vowel production.

C. Listening comprehension tools

Apart from the general English learning tool and those specifically concentrating on pronunciation learning and teaching, there are also some tools more interested in helping learners to improve their listening skills. *Electronic dictionary*, produced by Speechinaction, is an efficient pronunciation and listening training system, which makes

fast streaming speech easily understandable by exposing language learners to various possible phonetic environments of an utterance. The advantage of this programme is that, by being exposed to different delivery speeds, language learners are given the chance to perceive various phonetic changes of a sound in which some linguistic characteristics can be compared, e.g., articulation, sound reduction, so as to facilitate the understanding of natural L1 English speech. However, the productions at different speeds are mainly spoken by means of artificially adjusting them to fit into different speed bands. These productions are not natural spontaneous speech, which inevitably weakens its advantage of exposure to different phonetic environments.

D. Speech analysis tools

In addition to the technologies mentioned above used for language learning, there are other speech analysis toolsets which can also facilitate language learners in advancing their linguistic abilities. For example, *The Speech Analyzer* produced by SIL can provide detailed analysis on waveform, pitch, intensity and spectrogram. However, this speech analysis tool might not easily be used by language learners due to the fact that they are aimed at researchers and are technically sophisticated.

E. Audio tools developed in China

In China, there are also some products developed by Chinese researchers for English learning and teaching. For example, *StepbyStep* is one of those tools widely aimed at primary and secondary school students for pronunciation and listening training, in which language learners can adjust five different bands of delivery speeds from very slow through to very fast. However, this kind of speed change is not really helpful, since the slow-down is not linear.

Given this brief review of some representative English learning and teaching tools, it is not difficult to see that these TELL systems are useful to some extent in helping language learners enhance their English learning. Their potential drawbacks, however, impede language learners in further progressing their communicative competence.

5.3 Innovative Language Learning Technologies Devised in DIT

The advantages of TELL resources in English learning and teaching are well accepted. However, looking closer at those programmes, it is not difficult to see that there are still some disadvantages which prevent a broad application.

On the one hand, many online listening materials are authentic native dialogues, but the ‘messy’ rapid flow of speech is only delivered at normal (i.e., 100%) speed, which is not a comfortable speed for Chinese learners to capture the real ‘blur’ of English speech flow. Even if the same speech sequence can be re-played again and again, the excessively fast speed (from the students’ perspective) still prevents them from arriving at intelligibility and communicative ability. On the other hand, as noted above, there is in spontaneous conversations – inevitably – cross-talk. Mixed acoustic signals result in difficulty for Chinese learners in segmenting and interpreting the affected sequences, which can then not be individually retrieved by language learners.

In contrast, in addition to simple exposure to authentic English speech, some novel language learning technologies being developed in the Dublin Institute of Technology (DIT) may offer more advantages to Chinese EFL learners in enhancing their language learning and improving their linguistic abilities. These are detailed as follows.

5.3.1 *Slow-down Algorithm*

The development and application of a time-scaling slow-down algorithm was sponsored by Enterprise Ireland in a project which started in 2002. It is one of the technologies within the Digital Interactive Toolkit for Computer Assisted Language Learning (DITCall) project whose objective was the development of a digital interactive package to assist non-L1 language learners of English to enhance their listening and speaking skills in self-study. The main advantage of the slow-down technology is a real-time, variable slow-down facility for speech recordings allowing students to capture details in natural, authentic L1 spoken English without the pitch distortions induced by a mechanical slow-down facility. This allows language learners more time to study real speech action.

A. Need for slowed-down speech

Research carried out by Derwing and Munro (1997) shows that, among eight factors affecting comprehensibility, speaking rate ranks the second with 38%, only lower than enunciation which scores 46%. Ten out of twenty-six L1 English speakers think that the speech rate is negatively correlated with comprehensibility. In other words, understanding faster speech is more difficult than understanding slower speech. Articulation rate, rather than speaking rate (which includes pauses), seems to contribute importantly to phoneme identification, which is as Derwing and Munro (*ibid.*, p.14) put it, even though it may not be the direct cause, yet, ‘rate sometimes serves as a general scapegoat for perceived comprehension difficulties’.

Given that in connected stream-like English speech, standard citation forms of words are frequently modified due to the speed of delivery, most classroom-based non-L1

English learners are easily lost in segmenting and interpreting these modified signals. In order to enable language learners to capture the real-time acoustic blurs and to be comfortable with them, as Cauldwell (2002) emphasises, language teachers should adjust their methodology to ‘spend more time with the recording in the post-listening phase’ (p.10). This post-listening phase contains the real-time blurs in the streaming speech which are very important features of everyday listening. However, it is mainly due to the relatively fast speed of delivery that students fail to have access to real features of natural spoken English. Therefore, a language learning tool would probably help students to re-listen to the materials at a speed suitable for their linguistic abilities.

‘Most of the misleading information in the signal occurs in unstressed syllables’ (Brown, 1990, p.100). Since the speakers utter their communication plans normally with ease of articulation in mind, they only assign emphasis to the parts which they think more important than others. The unstressed segments are always neglected, which results in difficulty in perceiving for language learners. Therefore, in listening and comprehension teaching, Brown (ibid.) emphasises that language learners ‘need to be taught how to understand it ... If the student does not understand what is being said the first time the sequence must be repeated until he does understand it’ (p.159). However, the point is if the sequence is repeated at the same speed, even if the learners are given a second or third chance, it is still difficult for them to become aware of the important segmental clues which they can rely on in understanding normal informal English speech.

Due to the unpredictability of natural speech mentioned above, it is very difficult for people (language learners and users, even language researchers) to perceive its features at its normal spoken speed. Yet, the practical goal for learning and researching language

is to utilize it for communication based on recognising it thoroughly. Therefore, a slow-down facility is necessary in a tailored pedagogy to facilitate language teaching and researching, especially for non-L1 language learners.

B. Advantages of slowed-down speech

The advantage of slowed-down speech (as opposed to natural slow speech), on the one hand, is that a natural speech flow can be adjusted to a slower comfortable speed by the language learner, in which the problems for non-L1 learners in their decoding process, such as unstressed syllable, assimilation, elision, vowel reduction and centralisation, can be highlighted so as to help language learners recognise or internalise these natural acoustic blurs of NS and to cope with the natural flow of English speech.

On the other hand, the slow-down technology can also help language learners follow the intonation and rhythm patterns in English. As Tatham and Morton (2006) state: ‘all speech is expression wrapped’ (p.147) (i.e., no speech is affectively neutral, but occurs within a speech envelope which may or may not be marked for any particular emotion), and prosody is ‘the vehicle of expression’ (p.133). The slow-down algorithm allows spontaneous speech to be slowed-down to an effective maximum of 40% (slower speeds sound too unnatural) without tonal distortion. Language learners can easily follow the patterns and tune themselves in to the informal, natural, target language speaking community. This technology is therefore of use in the current research, further consideration is given in Section 6.2.2. An example is demonstrated in Figure 6, which shows the same signal slowed to 40% speed, which extends the timeline by a factor of 2.5, but demonstrates clearly that each audio feature of the original is retained and highlighted in the slowed version.

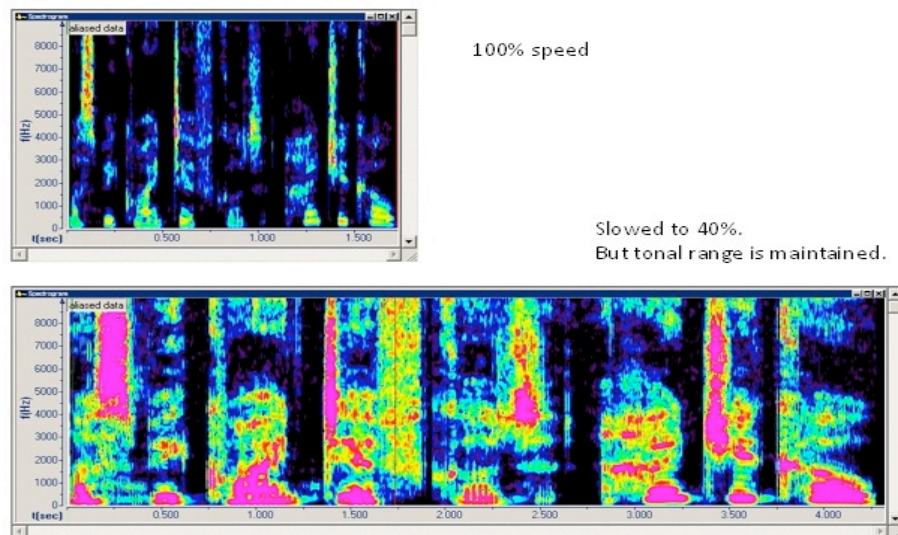


Figure 6: Slow-down algorithm

5.3.2 Vowel Pronunciation

The *Articulate!* project was funded in 2005 by Enterprise Ireland under the Proof-of-Concept Programme. The aim of the research was to create a new language learning tool, based on a novel method of giving dynamic feedback to language learners about the correct production of vowels. *Articulate!* is only at a prototype stage, and is not available to language learners, mainly because there were technical difficulties in correctly identifying subtle variations in vowels in real-time. However, the approach of *Articulate!*, by using the results of real-time digital signal processing (DSP) analysis applied to linguistics, provides learners with accurate and meaningful graphic feedback on how close their utterances are to an oral target. *Articulate!*, if fully developed, would be an objective and independent self-learning and self-assessment tool for non-L1 learners of English, without the need for language teachers or speech therapists.

A. Need for a tool for self-assessment of pronunciation

Kenworthy (1987, p.2) addresses the role of feedback for language learners in the acquisition of accurate pronunciation and maintaining active motivation. She

emphasises that learners ‘need to know what to pay attention to and what to work on’. The problem is that it is difficult for language learners to assess themselves. They therefore need help from language teachers. However, given the conditions under which English is learned and taught in China, and in view of the numerous large classes, it is not yet possible for learners to obtain information about their performance when they need it. Thus it would be useful for language learners to access a self-learning tool to let them know if they have ‘got it right’.

The faster an L1 language speaker speaks, the less time he has to produce target vowels, especially in the fast flow of informal speech when the tongue has to fly around the oral cavity. Some vowels located in more extreme positions, e.g., /i/ or /u/, are mostly reduced to schwa /ə/, which therefore often confuses language learners of that particular language – and that, in addition to the difference in vowel systems between the two languages.

A further reason for using the approach adopted by *Articulate!* for self-learning is that ‘face’ theory (as discussed by Brown and Levinson, 1987) is significant in some language learners’ L1 culture, in particular Asian users. Therefore, in order to avoid the embarrassment of making mistakes and in order not to lose face in public, Asian users, especially Chinese language learners, prefer to formulate and practise an utterance before speaking it. The non-invasive nature of *Articulate!* and its endless ‘patience’ could be key advantages therefore in coping with the individual learning requirements of Asian users.

B. Approach of the tool *Articulate!*

So, what language learners need is a facility which allows them to practise and self-assess English vowels, which are essential and subtle elements of speech, in privacy, without the embarrassment of having to do so in front of others in a classroom setting. *Articulate!* is constructed so as to provide real-time feedback to language learners at any time when needed by an intuitive, interactive, visual interface. Both a target vowel and the learner's attempt at re-producing it are mapped to (currently, in the prototype) a reference vowel quadrilateral. The learner's attempts to improve his performance are traced in real-time on the reference framework so that he can see which articulatory gestures improve or disimprove performance.

Even if perfected, however, *Articulate!* will at best address segmental problems and would be of limited use in providing feedback at supra-segmental level. Therefore, it is of little use in the current research. Further consideration is given to this in Section 6.2.2.

5.3.3 *Access to and Use of Natural Dialogic Native English Speech Assets*

The development of the Dynamic Speech Corpus (DSC), guided in part by the linguistic research of the present author, is being undertaken in DIT within the Enterprise Ireland-funded FLUENT project, which began in January 2008 and is due to be completed by mid-2010. When implemented, it is anticipated that the DSC will be a database of dynamic audio assets, combined with various types of tagging to mark significant linguistic features, and which aims to expose non-L1 language learners of English to real, natural English speech, so as to advance their linguistic abilities by the principled application of digital language learning technologies. These dialogic assets of the DSC

are of use in the current research, and further consideration is given to them in Section 6.2.2.

A. New dynamic speech concordance

The greatest advantage of the DIT dynamic speech concordance – when it has been fully implemented – is that the spontaneous phonetic environment in which the search string occurs is supplied. Clicking on one of the records returned by the search procedure will play the corresponding original WAV file. When completed, this speech concordance will sort multiple occurrences of the search string by speed of delivery and allow for the comparison of the same phoneme pronounced in different pitches, intensities and different patterns of contours from its idealised form to an extremely careless, speaker-oriented form. Multiple occurrences of even a phonetically reduced utterance can be found via the orthographic text search string. This is a citation form of transcript, which – linked to a speed index for the utterances in the corpus – will make it easier to locate samples of NS patterns. Learners can also choose any part of the audio file, listen to it and slow it down to any desired speed between 100% and 40% to meet their individual needs. Apart from accessing the audio file of an individual display line in isolation, clicking on the ‘Zoom Out’ button will allow the relevant speaker strings (before and after the central search line) to be played, then clicking on the ‘Interlocutor’ button will allow the relevant dialogue delivered by two speakers to be accessed and played, which provides language learners with access to a search string in its full phonetic, semantic and pragmatic environment. Aspects of an initial database interface are shown in Figure 7 (Campbell, et al., 2007). Such a dynamic speech corpus, by preparing language learners to cope with the peculiarities of native-to-native English

speech, is anticipated to benefit language learners and improve their communicative competence in a target speech community.

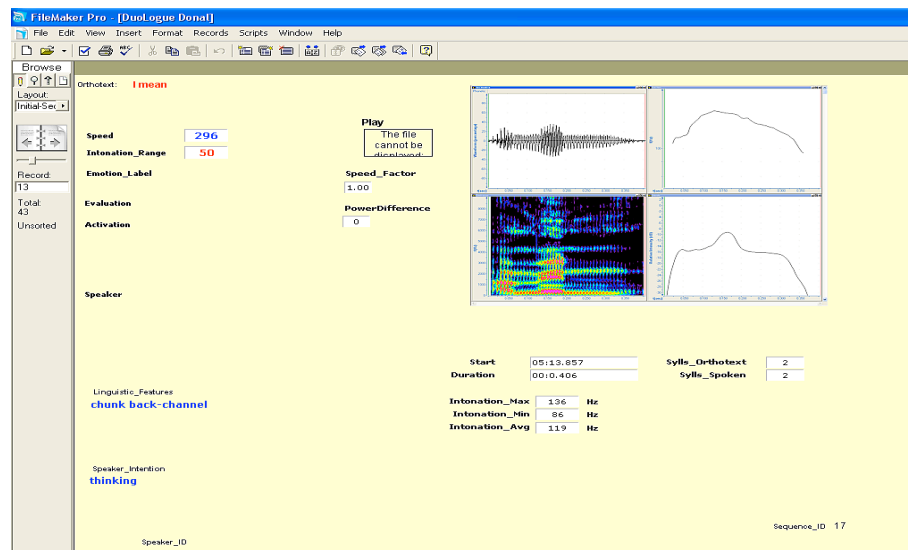


Figure 7: Initial (proposed) interface of Dynamic Speech Corpus

Source: Campbell, et al., 2007

B. Spoken corpora

Currently, there are some spoken corpora available, produced by linguistic researchers. For example, the Naturalistic Database – one of the best existing corpora – was implemented by a Belfast research group in 1999. It is a large corpus including 125 subjects, with many different kinds of emotions covered. Most of the database is taken from television chat shows, current affairs programmes and interviews conducted by the research team. Another corpus was constructed by Chung in 2000. The database is also taken from television interviews in which speakers talk on a range of topics with sad and joyful emotions involved. Of course, these databases are much better than the earlier ones, monologues or dialogues, which are devised specially for recording by actors in a studio environment. There are, however, some limitations in terms of emotion elicitation methods and natural linguistic expressions. As noted above, most of the database comes from chat shows or discussion programmes on television. The

interviews are conducted between the host and the actors. They do not know each other, at least not well, which undoubtedly results in some nervousness. Furthermore, these interviews are shown on television, so no natural emotions are expressed at all; most of the language they use is performed, more measured, often pre-planned or rehearsed. Because the ‘actors’ know that they are performing and being watched, one way or another, real conversational interaction cannot be naturally represented.

C. Speech assets from the Dynamic Speech Corpus

In contrast with the spoken copra discussed above, the essence of the authentic, spontaneous dialogues which populate DIT’s Dynamic Speech Corpus lies in its naturalness. It is deliberately called ‘speech’ corpus, the reason being that it focuses on the study of dialogic speech production. The database is collected in the form of dynamic dialogues between people who are familiar or know each other very well. Compared with those performed actions on television, these recordings have the considerable advantage of a lack of unfamiliarity or performance. It is the naturalness of this new resource which can reveal the actual usage of spoken language, rather than the tidy, scripted representations of recordings made specifically for foreign language learners. All dialogues are dynamic, unscripted, topic flowing between interlocutors, in which genuine interactive characteristics of spoken English, e.g., turn-changing, back-channelling, false starts, and hesitation, are tagged and demonstrated. Such a dynamic speech corpus could, as noted by McCarthy (2004, p.16), help language learners ‘experience authentic language’ without having to live in the target language environment.

The potential advantages of the DSC for enhancing language learning are outlined above. Given that the DSC is not fully developed, language learners can currently only access its early assets – real, natural, dynamic English dialogues. Unique advantages of these innovative audio assets are further described as follows.

Recordings in the DSC are designed and recorded using the facilities of the CSAL (Cognition, Speech and Audio Laboratory) in the DMC (Digital Media Centre), and have also contributed to the EU SALERO project. Compared with other spoken corpora recorded by other linguistic researchers, some advantages are apparent:

Firstly, as a result of the recording techniques developed by the CSAL, dialogues can be recorded which exhibit a high degree of naturalness. Isolation booths are used for recording, and apart from avoiding distracting noises, this method can also overcome the observer effect to provide natural recordings. Speakers are either family members or friends who know each other well and who therefore relax in each other's company. After five minutes or so settling down and warming up, real natural interactive English speech is obtained. This is a main objective pursued by the DIT research group in the creation of an authentic natural database.

Secondly, an industry-standard high-quality audio recording is achieved in the CSAL. Dialogues are recorded at a 24bit/192KHz professional resolution (Cullen, et al., 2006). This level of recording quality, which is four times CD quality, ensures that optimal assets are obtained for both linguistic, and potentially, later, instrumental analysis. Meanwhile, a baffle is used to prevent unnecessary sound inputs e.g., for blocking of plosive /p/, and equipped with wideband absorbers as well so as to reduce background

noises, e.g., the computer, the lights, paper being-folded, and the microphone and body movements.

Thirdly, the recordings in the CSAL produce genuine, spontaneous dialogues rather than monologues or orchestrated turn-taking and acted sequences. Participants produce genuine dialogue, including back-channelling, word-fillers, and cross-talking where both speakers are talking at the same time. In traditional recordings, these acoustic sequences would be spoiled and could not be used for analysis purposes; whereas with the advanced technologies developed by CSAL, genuine conversational interactions can be obtained without spoiled acoustic signals and collated to form a speech database. Language learners can access the dialogue as an entity, or choose to listen to the speech of each individual speaker (as two ‘semi-logues’), or switch between them, as shown in Figure 8. This is a unique approach and a unique resource, which involves real, natural, interactive dialogues while avoiding the acoustic confusion of overlapping signals. This is anticipated to benefit both language learners and researchers, and it forms the basis for the linguistic research undertaken by the present author.

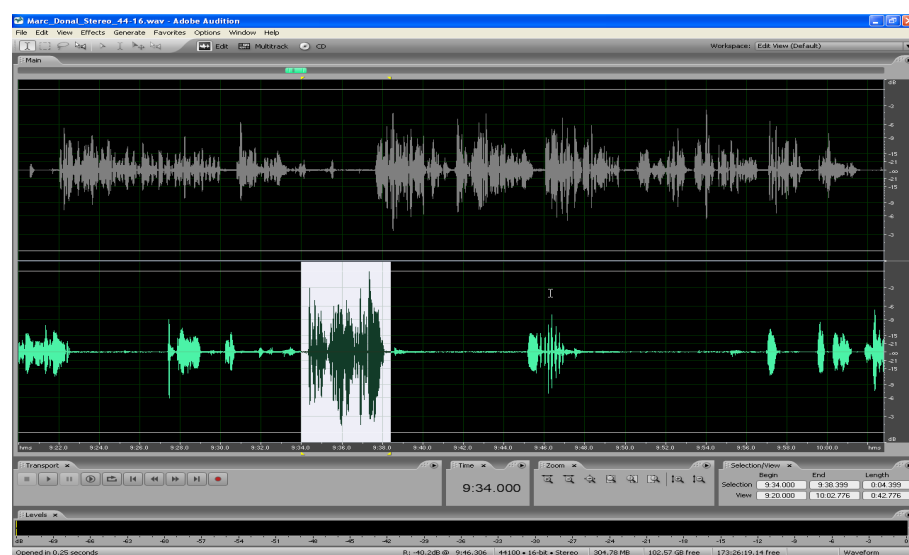


Figure 8: Waveform representation of dialogue produced in the CSAL

Given that the natural, interactive, native-to-native English dialogues demonstrate the remarkable (and for the learners: difficult to imitate) characteristics of speech production, they will be a major contribution to language learning and research. Therefore, this innovative speech corpus when fully developed and its speech assets analysed, will allow the discovery of the key features of natural English speech. This unique database might satisfy the needs of EFL learners, language teachers, especially EFL teachers, and linguistic researchers.

5.4 Conclusion

Given the advantage of computer-assisted technology in enhancing language learning, as discussed in Chapter 4, various language learning tools are widely employed by language learners. While facilitating learners of language and promoting autonomous study, these conventional language learning tools also have some disadvantages in design and usage, as discussed in the first part of this chapter. This gives rise to the need for a more innovative language learning platform which better promotes language learners' linguistic competence.

Some language learning speech tools developed by DIT, e.g., the slow-down algorithm, the approach adopted by *Articulate!* with its segmental recognition, the Dynamic Speech Corpus and its unique natural dialogic speech assets, are designed with the aim of bridging the gap in available technology, as considered in detail in the second part of this chapter. These innovative TELL technologies, when they are available, might be able to advance language learners' spoken communicative abilities by exposing them to real, natural, spoken English, so as to facilitate them in coping with authentic, informal English from a native-speaking community.

Based on the literature review in Chapters 2 and 3, and the discussions in Chapters 4 and 5, the research design of the current study is presented in the following chapter.

6. Research Design

6.1 Introduction

Spoken English and Chinese are reviewed in Chapters 2 and 3. Different pedagogies of English learning and teaching as an L1 and L2/EFL are also examined. In addition, some problems encountered by Chinese EFL learners in their native-speaking communication are described. What Chinese language learners need to cope with natural, authentic, dialogic English speech, and the pedagogical use of technologies on language learning are also discussed in Chapter 4. In Chapter 5, current conventional TELL tools are briefly reviewed. Some innovative speech tools developed by DIT are also described, which are aimed at facilitating language learning, and improving students' abilities in spoken English and promoting autonomous study.

In this chapter, based on the review and discussions above, the scope of the current research is identified, the rationale for designing the study and the procedures and methodology for its implementation are also described, which leads to the research work undertaken by the present author, as outlined in the following chapters.

6.2 Rationale and Procedures in Current Research

6.2.1 Analysis of Formulaic Language

As considered in Chapters 2, 3 and 4, given that the significant role formulaic language plays in natural, dialogic, L1-L1 English speech, and also due to the gap in the literature on the phonological realisations of formulaic language, one of the main areas of this current research is concentrated on phonological aspects of formulaic language. Considering the specific problems which Chinese EFL learners encounter in the east-

west prosody divide, phonological analysis of formulaic language is mainly focused on the speed of delivery and pitch range. This addresses RQ1.

The analysis is based on the real, natural, interactive speech recordings of the DSC. An investigation into the segmentation of the speech into viable sub-units called ‘flow units’ (as mentioned in Section 2.2.2 and detailed in Section 7.2.2) as well as undertaking work in the area of tagging these units for significant linguistic features constitute the bulk of the linguistic research of the current study.

6.2.2 Speech Technologies Employed

As discussed in previous chapters, the role of computer-assisted language learning technology, given its efficiency in language learning and teaching, especially autonomous study, has been widely acknowledged. TELL tools, therefore, play an increasingly important role in the field of language learning. Current TELL systems from different domains of language learning, e.g., listening, speaking, reading and writing, facilitate, to some extent, language learners in improving their linguistic competence. Some of them, due to disadvantages in design and usage, seem to highlight the need for a more efficient and innovative language learning platform, which can accommodate language learners by offering more efficient self-learning and self-assessment.

Some technologies developed in DIT were available when the present author was starting her research, such as the slow-down algorithm, the approach to segmental skills adopted by *Articulate!* and the assets of the developing Dynamic Speech Corpus. However, it was decided to employ only the slow-down facility and the natural, authentic, recording assets of DSC in the current research.

As considered earlier, time-scaling technology can slow any audio file down to approximately 40% without distortion which, on the one hand, can highlight reduced phonetic features which might easily lead to misunderstanding or confusion for language learners when accessed at normal speed. With this resource, it is possible for language learners to listen to NS ‘blur’ (where it naturally occurs) and contrast the blur with the citation form in the learner’s head, which is set as one of the main areas the current research aims to investigate. The slow-down helps them to concentrate on these phonetic phenomena and improve intelligibility in real target language speaking environments. On the other hand, use of a slowed-down replay speed can help language listeners to easily follow the intonation patterns of English speech by being exposed to the audio signals for two and a half times longer, at the 40% speed. This will allow learners to comfortably follow native intonation patterns, in particular learners with L1s which are tonal languages, so as to facilitate them in coping with the natural flow of connected English speech.

There is one point however which needs to be considered with this technology used at 40% speed. Due to the playback taking two and a half times longer than normal, the speech sounds unnatural. This is well acknowledged by language users and is known to the DIT linguistic researchers. The different speeds are suitable for different domains of language learning. For example, research carried out by Meinardi (2006) and Richardson (2009) demonstrates that both 80% and 60% speeds are effective for word recognition and pronunciation improvement. In contrast, the 40% speed, by providing extra exposure to the natural flow of speech, is not only anticipated to be helpful for segment recognition, but also useful for increasing ‘conscious awareness’ (Crabbe, 2003) of intonation patterns, which is also part of the research the present study aims to

investigate. If the 40% speed can be demonstrated to benefit non-L1 language learners, especially Chinese EFL learners, then the advantages will obviously outweigh its drawback of sounding unnatural and artificial. Therefore, it was decided to use the 40% slowed-down speed exclusively in the training sessions for the Test Group. This addresses RQ2.

The present research makes use of the early assets of the DSC, such as natural, authentic, interactive L1-L1 English speech, a high degree of naturalness, industry-standard audio, and a method of recording which allows the speakers to be separated so that clear signals are available even during cross-talking. The assets also provide significant linguistic features, such as formulaic language, which forms the bulk of the current research. One reason is that these unique recording assets allow the present author to analyse real, natural, dynamic English speech. The analysis of formulaic language and its phonological realisations, i.e., speed of delivery and pitch range, is based on the analysis of these natural dialogue resources. Another reason is that all materials used in the case study for testing and training sessions were taken from the same natural, spontaneous, L1-L1 speech dialogues, in which significant linguistic characteristics and sociocultural knowledge are embedded. As anticipated, there are strong indications that the authentic, dynamic English speech may facilitate learners of English in improving their pragmatic competence in using language by exposing them to a real target language speaking community. This addresses RQ5.

Articulate!, as discussed in Section 5.3.2, is another pilot language learning tool developed by DIT which facilitates language learners in self-practising and assessing their vowel production. However, it was decided not to include this technology in the current study. The main issue is that the programme is still at the prototype stage, and is

not available for language learners. Another issue is that *Articulate!* would be used for vowel practice and recognition which occur at segmental level in an isolated production environment. While the present study concentrates on the natural flow of authentic speech – the intelligent ‘blur’ and intonational patterns – this technology is of relatively minor interest to the current study.

6.2.3 Research Undertaken in Application of Speech Technologies to Acquisition of Formulaic Language for Chinese EFL Learners

In view of the limitations acknowledged by Meinardi (2006) and Richardson (2009) in their studies, it was decided to carry out both tests and the training sessions in China rather than Dublin as the test students would be uninfluenced by an English-speaking environment and therefore the advantages gained by using the training materials could be more objectively judged. All the training materials were based around formulaic sequences, given their importance in native-like prosody.

By using test subjects in three Chinese school settings, homogenous bodies of students could be formed to act as test and control groups. The control group had access to the training materials at 100% speed only, whereas only the test group was allowed to use the slow-down facility. The efficiency of technology-assisted language learning could be verified by comparing test results between the test group and the control group, each having the same teacher, thus further reducing the number of variables in the tests.

Test 1 was designed as a benchmark test with the aim of assessing the basic linguistic proficiency of the participants. This would be used as the baseline to compare with the results in Test 2. After Test 1, and before Test 2, there was a gap of ten months, which

was deliberately designed to allow for an adequate number of training sessions. In this period, training materials taken from natural, dynamic, native English speech were devised and delivered to the participants. The slow-down technology was used to provide the test group with audio assets at 40% speed. This was the main difference between the test group and the control group. After giving participants enough exposure to authentic English speech, Test 2 was planned to verify the benefits of the application.

The slowed materials were available only to the test group and were used only in the training sessions. Both Test 1 and Test 2 employ normal (100%) speed, since testing the actual linguistic proficiency of language learners has to be undertaken in a natural environment.

As examined in Chapters 2 and 3, given the differences between English and Chinese, the specific problems encountered by Chinese learners of English are, on the one hand, pronunciation, especially a staccato spoken delivery and a flat prosody leading to confusion and misunderstanding by L1 English speakers. On the other hand, the messy flow of connected English speech frequently causes Chinese learners to get lost in the decoding process. This is a difficulty experienced by many Chinese language learners and prevents them from engaging with a real, informal target language speaking community. In order to cross the east-west divide in pronunciation and prosody, Chinese learners need to notice and to tame the natural acoustic ‘blur’ of English speech, so as to cope with authentic English communication. This is the central concept guiding the present experiment.

The hypothesis for these tests, therefore, is that the formulaic language, in conjunction with the speech technologies (i.e., slow-down facility and the natural, authentic,

dynamic English speech) will be of help for Chinese learners of English when involved in a target language speaking environment. The aim of the tests is to identify Chinese language learners' ability to imitate and re-produce a native-like acoustic blur as exemplified in formulaic language. The methodologies employed are designed to distinguish between Chinese learners' accurate perception of the original NS recording and their production of 'intelligent blur' – i.e., a principled imitation of the NS reductions present in many formulaic sequences.

Proposing this aim for the current exploratory study does not support the proposition that non-L1 language learners have to learn native-like pronunciation and prosody so as to arrive at successful communication. In reality, as Brown and Yule (1983a, p.27) point out, 'the aim of achieving native-like pronunciation is not only unattainable but unreasonable'. Non-L1 language learners are most likely to communicate with other non-L1 language learners (Jenkins, 2005, p.145).

The motivation for the current study is simply, as noted by Wray (2002) that, '... if the speaker has non-native-like phonology, the hearer will need to engage in extra processing for the phonological decoding' (p.99). Also as Robin Walker (Personal communication, 15 October 2008) points out, 'From an ELF (English as a Lingua Franca) perspective, users need to be able to deal with NS blur on those occasions that they are saddled with an L1 speaker who is unable or unwilling to accommodate to their non-L1 speaker interlocutor'. Given the considerable proportion of formulaic language produced in natural L1-L1 English speech and given that its more significant role is realised by phonological delivery rather than its lexical structure, training particularly Chinese EFL learners' ability to appreciate and produce native-like acoustic blur is an ambition to get them to a level of communicative competence aligned with C Level of

The Common European Framework of Reference (CEFR), or, at least, to quote McCarthy (1988, p.198), for the purpose of ‘preparing the learner to tackle without panic the harsh realities of natural talk’.

The initial work to be carried out in the current research is the analysis of formulaic language, bearing in mind the potential users of this material and using DIT speech technologies, the research work to be carried out is as follows.

The next step is to design an initial test (Test 1) to establish the existing level of students’ linguistic ability. Then comes the implementation of Test 1, followed by evaluation and analysis. Next comes the design of training materials and their delivery over six months to prepare for Test 2, and carrying out of evaluation and feedback from these training materials. There then follows the design and implementation of Test 2, and its evaluation and analysis. The next step is the comparison of Test 1 and Test 2, and the evaluation of the application of the assets and technologies. Finally, the discussion of tests and further research is considered.

An evaluation methodology, by which the natural flow of speech production is evaluated, is to be investigated in this case study. This addresses RQ4. The pedagogical effectiveness of using natural, authentic, dialogic English speech with technological support is also to be demonstrated in this case study; which addresses RQ3.

6.3 Conclusion

Based on the literature review in Chapters 2 and 3, and the discussion in Chapter 4, and also the review of TELL tools in Chapter 5, the research design of the current study is presented in this chapter.

The five research questions identified inform the remainder of the research work, which is further outlined in the following chapters. Formulaic language and the analysis of its phonological realisations is detailed in Chapter 7. Application of speech technologies to the acquisition of formulaic language for Chinese EFL learners is further investigated in Chapter 8.

7. Analysis of Formulaic Language

7.1 Introduction

As described in Chapter 6, the area of the present research is largely focused on the analysis of formulaic language. Therefore, the analysis of formulaic language, the link between various categories of formulaicity and their phonological realisations, the correlation between various positions of formulaic sequence within each speech unit and their different speeds of delivery, forms the bulk of the investigation in this chapter.

The analysis of formulaic language is mainly dealt with in 7.2, and the conclusion is covered in 7.3.

7.2 Analysis of Natural Authentic Interactive English Speech

The process undergone in dealing with the analysis of formulaic language is as follows. After listening to the natural, spontaneous NS recording, the first step is to transcribe the dialogue as orthographic text (described in 7.2.1), and then to segment and tag the speech units for further linguistic analysis (described in 7.2.2). Some significant linguistic characteristics of spoken English, i.e., formulaic language and its phonological realisation, with respect to speed of delivery and pitch range, are investigated based on the current data (detailed in 7.2.3). Some questions are left open for further research (discussed in 7.2.4).

7.2.1 Inclusive Transcription via Orthographic Text

As described above, the recordings to be used in the current research are all taken from the Dynamic Speech Corpus (DSC). A total of 20 hours recordings are included in the DSC, with roughly 50,000 words transcribed so far. The DSC covers a wide range of

topics, from everyday life to social, economic, cultural and religion. Speakers are from Ireland, England, Scotland, America and Canada. All these recordings are native-to-native speech. Several salient linguistic features, for example, formulaic sequences and their speed of delivery, cross talk, back-channelling, and word-fillers, are revealed in these recordings. Some sequences were delivered at considerable speed, and even though they could be listened to again and again, their significant linguistic characteristics could still not be captured without access to a slowed-down version.

A. Transcription as orthographic text

The transcript aims at making accessible the natural, interactive features of normal English speech using an idealised written form. The transcript does not attempt to emulate the phonetic realisation of the utterances, but rather to make available their meaning and to allow the relevant speech sequences to be located. These transcript versions are essential for linguistic analysis. Based on the orthographic transcription the present author can segment the speech recording according to natural pauses and communicative pitch changes, and tag it so as to mark the significant linguistic features. All these procedures are based on phonologically significant phenomena in the signal but made available via the basic written form – the transcription.

Many key features of natural, relaxed native-to-native dialogue cannot easily be accurately represented in written form. Acoustic blurs, for example, occur only in real-time spontaneous speech. The transcriptions, therefore, are not phonemic or phonetic transcriptions, but represent the idealised or citation form.

B. Features and advantages of orthographic text

Based on the acoustic signals, the written version of the dialogue is presented in citation form rather than reduced form. For example, ‘it’s’ is rendered as ‘it is’, and ‘how dju know’ is transcribed as ‘how do you know’. It is these contracted forms in spoken speech that may make language learners confused in interpreting the speech signal. Highlighting these variations between informal speech and formal written form facilitates language learners in dealing with real, spoken English communication.

The second feature which the orthographic text can help with is cross-talking in the recordings. As mentioned earlier in Section 5.3.3, since high-quality technology and equipment are employed in the CSAL, a dialogue can be separated into two ‘semi-logues’, therefore, there are no gaps left in the transcription, i.e., no cross signals which impede analysis. Cross-talking frequently correlates with back-channelling when the listener gives responses to the speaker, or with turn-taking when the listener wants to take over the conversation. These are significant characteristics in natural conversational interactions, which other transcriptions fail to demonstrate, and they are also key to the linguistic analysis, such as in the case of chunks, and their delivery speed and pitch range, as detailed in Section 7.2.3.

Another advantage of the orthographic text which is also important, is that it is an inclusive transcription, i.e., all speech actions with communicative value are transcribed, including truncations, false starts, hesitation, interruptions, repairs, and also non-linguistic features, e.g., non-word fillers, like ‘um’, ‘erm’. A sample is shown below in Figure 9, in which two speakers are transcribed in two separate columns, and the different colours are used for highlighting various linguistic features. Brown and Yule (1983a, pp.86-87) also emphasise that it is unhelpful to present language learners with a tidied-up version of natural spoken language, since this mismatch between the

real messy speech and the neatly transcribed written-language-like version will undoubtedly be detrimental to language learners in perceiving the real characteristics of natural, interactive conversations and in improving their pragmatic competence in using the target language. This is also another gap that the current research methodology aims to bridge.

what do you	I thought it was OK
	I am not a good judge
	the restaurant was nicer than
	than the food
Yes	Yeah
	Physically
I	Which
i guess you are right	You
	that is
Yeah	that is quite common here these days
you were sitting next to	The
to whom	
	I
	I was

Figure 9: Transcription of natural conversational speech

At present, all the processes of transcription and segmentation are implemented manually. That is to say, after listening to the original WAV file, the present author has to separate it into small segments and transcribe them. In the light of reducing the great deal of detailed specialised work and to speed up the process, a transcribing tool would be desirable. The orthographic transcription described above forms the basis for the analysis of speech in ‘flow units’, as detailed below.

7.2.2 Segmenting and Tagging the Units

As reviewed in Section 2.4.2, in order to fill the identified gap in the literature on formulaic language, the present study aims to link different categories of formulaic language with their phonological realisations, e.g., speed of delivery and pitch range.

Segmenting and tagging the speech sequences is an essential step in this process. The first decision to be made concerns how to ‘segment’ the sequences.

A. The segmenting process

As reviewed earlier in Section 2.2.3, connected speech is a flow of signals in which there are actually no clear-cut segments in the stream. The ‘segment’ meant here, as Tatham and Morton (2006, pp.13-14) put it, refers to ‘an abstract label which is assigned to a portion of signal’. What the present author is trying to do is to assign useful labels to the sequences for linguistic and acoustic analysis in units which are determined by the speaker.

a) defining ‘flow unit’

In contrast to the linguistic units discussed in Section 2.2.2, a new concept, the ‘flow unit’ is set as the minimal production unit for labelling and tagging the sequences. The flow unit was originally called ‘timed unit’ by Campbell, et al. (2006) because the segmentation facilitated the recognition of speed-induced phonetic ‘distortions’. A working definition of this unit is described as ‘a segment which has its own flow and pragmatic integrity. It is speaker dictated and can be timed, and is bounded by pauses – no matter how brief – or marked by a sudden change of tack/thought’ (ibid.). In the flow unit, it was later refined by the FLUENT research team as ‘a speaker-determined production with tonal coherence or ended by a perceptible pause’. See Figure 10 below for some examples.

b) features of the flow unit

A key feature of the flow unit is its shortness. On average, the experience of the present author and the FLUENT research team (Campbell, et al., 2006) shows that flow units tend to be three or four words in length. Given that informal, unscripted conversation is real-time speech delivery, in which speakers mainly concentrate on communication, rather than producing grammatically correct sentences, therefore the speaker has little time to think through what is to be put into words, and so consequently the delivery is often interrupted by pauses, filled pauses, repetitions, or sudden shifts in intonation, etc., as reviewed in Chapter 2. That is to say, spontaneous speech delivery is almost always characterised by imperfect productions consisting of short phrases, a fragmentary chunk of words, or even incomplete syllables, which contribute to the messy nature of everyday English speech. Another reason which contributes to the brevity of the flow unit is due to the dialogic interactions in natural native-to-native English conversations. Dialogic speech is highly sensitive to interlocutor reaction and feedback. It can be micro-adjusted almost instantly by the speaker, who is at the same time his own listener, monitoring how the utterance might sound to the interlocutor. As natural L1 English dialogue is a dynamic flow process in which there is interactive engagement on the part of both speakers, therefore, it is highly marked with authentic speech features, such as back-channelling, discourse markers, which are more likely to be shorter chunks, rather than longer complete sentences. That is to say, this shorter-than-usual interactive unit better matches the production unit of informal speech than longer, semantically-driven, elegant formulations.

c) advantages of the flow unit

Tagging based on the flow unit can enable L1-speaker phonetic phenomena to be examined in their natural, immediate, phonetic environment. Firstly, it allows a more

accurate speed of delivery to be calculated for the short sequence in which an utterance was spoken, thus better capturing the high speed of unstressed parts of the speech flow and therefore making it more likely that those speech sequences can be found and retrieved which characterise the difficulties of native-to-native speech and have to be addressed by the learner. Secondly, apart from the tagging based on speed of delivery, the flow unit can also allow a more subtle perception of pitch contour to be evaluated. As considered in Chapter 2, English is an intonation language, in which intonation is widely employed by speakers to express their attitude, to direct the listener's attention, and to mark salience in the speech, etc. Therefore frequency in the change of intonation patterns would indicate that a smaller, more subtle unit might be required, especially in the context of a dialogic flow of speech interaction.

d) segmentation of natural conversational speech into flow units

Therefore, the sequences in speech are characterised by short runs, bounded by pauses, hesitations, subtle changes in intonation, and even, on occasion, completed word-final consonantal stops. Based on this working definition, it is clear that the flow unit is a natural production unit which can be labelled in two ways. On the one hand, it is determined by speakers by inserting a silent pause which functions like punctuation in written language to indicate the sequences of language which need to be co-interpreted by the listener. The minimum threshold for pause recognition, based on Goldman-Eisler's (1968, p.12) investigation, is 0.25 seconds. On the other hand, it is possible that there is no physical pause between sequences and that the sequence boundary is marked by a sudden change occurring in the mind of the speaker, indicated by a sudden change of pitch direction. Sometimes the semantic content of one sequence changes suddenly, clearly distinguished from its neighbouring sequences, even though there is no

straightforward phonetic evidence for segmenting them, especially in less-organised informal speech. The sudden change in speakers' tack can also be marked by changes in other linguistic aspects, such as speed of delivery or pitch contour, see, for example, training sample No. 42 in Appendix 16. The purpose of segmentation is to facilitate tagging the duration and intonation contour variations of each spoken sequence, excluding gaps, to discover the features of natural L1-L1 English speech. This flow unit, therefore, seems to be potentially, at least, an ideal unit for linguistic analysis, even if sometimes there is only one syllable or one word involved in an individual flow unit.

For example, in one of the recordings, when one speaker was setting up the recording session with two microphones, he said: 'The ... lapel mike is on', followed by an uninterrupted utterance 'The front mike is on'. These two sequences obviously paralleled each other. Each of them was a complete semantic sentence with an intact intonation pattern. However, distracted by the physical checking process, the speaker slowed the delivery of the first sequence with a slight gap after the initial 'the'. For speed-checking purposes, the sequence was segmented into two flow units – 1a and 1b, as shown in Figure 10. The same case applies to the flow units 6a, 6b, and 6c. The absence of internal breaks in a single flow unit will supply a more accurate speed of delivery indicator.

	A	B	C	D	E	F	G	H	I
1	Speaker	Unit	Orthotext	SigSyll	Start	Duration	OrthoSyll	Speed	Emotion?
2	Marc	1a	The	1	0:0.000	0.363	1	165	
3	Marc	1b	lapel mike is on	5	0:0.419	1.041	5	288	
4	Marc	2	The front mike is on	5	0:1.916	1.028	5	292	
5	Marc	3	#erm#						
6	Marc	4	I shall leave that for the moment	7	0:3.870	0.863	8	556	
7	Marc	5	#erm#						
8	Marc	6a	Pat	1	0:6.517	0.327	1	183	
9	Marc	6b	Patrick	2	0:6.873	0.366	2	328	
10	Marc	6c	Paddy	2	0:7.234	0.366	2	328	
11	Pat	7	P	1	0:8.104	0.359	1	167	

Figure 10: Segmentation of natural conversational speech into flow units

B. The tagging process

Following upon the segmentation of speech into flow units, tagging these flow units is the next step. The sequences in the speech flow can be tagged for many features which would be of interest to the learner, such as gender, age, variety of English, regional accents and emotion. But at this early stage of the current research, tagging was primarily concerned with the process of timing the WAV file so as to calculate the delivery speed of each sequence, and also to get the highest and lowest pitch value so as to determine the pitch range of each sequence. While researchers, such as Laver (1994, p.158), use articulation rate as opposed to speaking rate in calculating speed of delivery, the present author uses communicative speed, as detailed below. The rate used is based on the number of syllables per minute (Towell, et al., 1996), rather than number of words per minute, since the number of syllables within a word varies. Pitch value is measured within a normal range of 75 Hz to 600 Hz (the default in *PRAAT*) which sufficiently covers all speakers, since the average fundamental frequency for men is around 120 Hz, for women 220 Hz, and for children 265 Hz (Cruttenden, 1997, p.3).

a) tagging for speed of delivery

The programme *Speech Analyzer* was used to facilitate manual tagging in the present study. After loading the WAV file, the boundaries of a sequence are determined, based on the acoustic signals displayed in the graph forms: waveform, spectrogram, intensity and pitch. For example, in one of the recordings one speaker said: ‘that is my full name’. It starts at 0:6.925, timed from the beginning of the recording; the duration is 0.914 ms, and the number of orthographic syllables is 5 (c.f. 4 syllables in the actual speech signal!). Thus, the duration divided by the number of orthographic syllables, and multiplied by 60, gives the number of syllable per minute for each flow unit (see Figure 11).

	A	B	C	D	E	F	G	H
1	Speaker	Line	Orthotext	SigSyll	Start	Duration	OrthoSyll	Speed
2	Aelish	185	probably	1	6:38.162	0.185	3	973
3	Aelish	151	I do not know	3	5:33.390	0.259	4	927
4	Marc	307	yeah	1	10:37.009	0.065	1	923
5	Marc	15	I am not giving	4	0:28.801	0.359	5	836
6	Marc	254	Yeah	1	8:59.738	0.072	1	833
7	Marc		Yeah	1	3:14.785	0.075	1	800
8	Marc		Yeah	1	6:22.970	0.075	1	800
9	Aelish	317	and	1	10:55.488	0.076	1	789
10	Marc		yeah	1	8:59.958	0.079	1	759
11	Marc	186	That is right	2	6:0.976	0.243	3	741
12	Marc	109	That is right	3	4:12.225	0.250	3	720
13	Marc	256	yeah	1	9:3.070	0.084	1	714
14	Marc	10	originally	4	0:19.304	0.424	5	708
15	Marc		very good	3	9:33.502	0.258	3	698
16	Marc	65	Right	1	2:33.489	0.087	1	690
17	Marc	208	That is right	2	7:32.976	0.266	3	677
18	Marc		Yeah	1	8:1.450	0.092	1	652
19	Marc	240	engineer	3	8:32.314	0.281	3	641
20	Marc	65	right	1	2:33.749	0.095	1	632
21	Marc	111	that is right	2	4:15.730	0.287	3	627
22	Marc	22	how are you doing?	4	0:40.679	0.482	5	622
23	Marc	252	That is correct	3	8:55.194	0.387	4	620
24	Marc		right	1	5:5.544	0.097	1	619
25	Marc	81	Civil War?	3	3:14.307	0.294	3	612
26	Marc		isn't that right?	4	6:37.768	0.393	4	611
27	Aelish	206	he was in the	4	7:27.517	0.393	4	611
28	Marc	153	That is where Daddy is from	5	5:37.579	0.692	7	607
29	Marc	127	He was a Taylor	5	4:41.263	0.498	5	602
30	Marc		the	1	6:24.456	0.100	1	600
31	Marc	130	Israel	3	4:45.854	0.310	3	581
32	Aelish	66	but it does not mean	4	2:37.791	0.520	5	577
33	Marc	153	is not it	2	5:38.272	0.312	3	577
34	Aelish	271	sure	1	9:31.063	0.105	1	571

Figure 11: Segmenting and tagging the sequences based on flow unit

The reason why the number of orthographic syllables is used, rather than the syllables as spoken, is that this is a more objective and accurate way to present the phonetic reductions common in informal native-to-native English speech which are useful measures of communication rate, as opposed to speaking rate or articulation rate (as shown in Table 3). This spoken speed, therefore, is also termed by Dermot Campbell (Personal communication, 9 July 2008) as communicative speed, i.e., the elevated speed

which would be required for the citation forms to be spoken in the same time frame as the reduced syllables. At the moment, the process of tagging for speed is carried out manually, but it is hoped to develop an automatic calculation of speed of delivery.

	Speed at which ...	Pause
Speaking Rate	phones actually spoken (not citation version)	included
Articulation Rate	phones actually spoken (not citation version)	excluded
Communicative Speed	citation syllables spoken (re-constituted meaning, if any element of word present in speech signal)	excluded

Table 3: Definition of speeds of delivery

One of the initial aims of segmentation is for speed calculation. The reason is that many of the significant features of native-to-native speech are reflected by delivery speed. By segmenting the sound file and providing a speed index, reductions or weak forms in spoken language can be more readily found, such as ‘dju remember’, and also other characteristic features of informal speech, such as formulaic language.

b) tagging for pitch range

Apart from speed of delivery, the pitch range is also tagged based on the flow unit. Given that formulaic language is one of the main domains investigated in the present research, there is some evidence to show that there is a correlation between formulaic language and its intonational pattern. Pitch range therefore is the second objective in segmenting and tagging the sequences. *PRAAT* – a professional software for acoustic analysis – is currently used for pitch value evaluation. After opening a WAV file, locating in and out points in the signal to correspond with the flow units, the pitch is then analysed for maximum and minimum values. The difference between the maximum and the minimum is therefore the pitch range of this sequence. Unlike

tagging for speed of delivery, pitch range does not necessarily have to be tagged for each flow unit. It is currently calculated mainly for sequences that might have significant linguistic features, e.g., formulaic language. By the process of segmenting and tagging, different patterns of contours can be discovered with contrasting communicative values, which are among the goals of the current research.

c) tagging for speech attributes

Another key aim of segmentation is tagging for speech attributes, which includes tagging for marked features in each turn, flow sequence and flow unit, under the headings of turn behaviour, speaker intention, formulaic sequences, phonetic features, and discourse function, etc. Given that the tagging for speech attributes does not directly correlate to the PhD research questions of the present author, but follows on naturally from them, further research in this area is reported in Appendix 1.

7.2.3 Initial Findings – Formulaic Language

Most of the present author's research, as described so far, has been in the analysis of recorded speech. The main domain being investigated in this study is formulaic language. Formulaic language, given its frequency of occurrence and prominent position in native-to-native English speech, is well documented in the literature, as considered earlier in Chapter 2. Based on the multiple functions of formulaic language, various categories and terms are defined and assigned to this linguistic phenomenon, e.g., collocations, chunks, lexical phrase, fixed expressions and idioms. There are also many phonological features, such as precise articulation, speed of delivery, stress, pause and intonation patterns, which are identified as correlating with formulaic language. However, as the present author pointed out earlier, there is no research carried out so far based on real, dynamic dialogue to demonstrate the link between different categories of

formulaic language and their relevant phonological realisations. The considerations given below by the present author are aimed at bridging this gap and adding to the body of knowledge, which addresses RQ1.

A. 'FS≠FS'

Firstly, not all formulaic sequences (FSs) are equal. The same sequence of formulae, depending on different speeds of delivery and intonation patterns, can realise different communicative functions. According to the analysis given by McCarthy and O'Keeffe (2006), the chunk 'you know' makes up 60% of the word 'know' in their corpus analysis. Tagging the audio recording made in the CSAL between the speakers Marc and Donal (who are colleagues and friends), there are seventeen sequences of 'you know' occurring in Marc's database, and eighteen in Donal's. The fastest and slowest speeds among them are shown below in Table 4.

Sample	Speaker	Orthographic Text	No. Flow Unit	Speed (syll/min)
1	Marc	you know	1124	694
2	Donal	you know	158	674
3	Marc	you know	892	266
4	Donal	you know	509	302

Table 4: The fastest and slowest speeds of delivery of FS 'you know'

Comparing these, it is clear that the speed of delivery in Samples 1 and 2 are impressively faster than the medium articulation speed of 5.3 syllables per second (= 318 syllables per minute) (Laver, 1994, p.158). The reason is that these two chunks uttered here by both speakers are stored in the mind as a unit and uttered unconsciously and automatically just to keep the conversation going. At the slower speeds shown in

Samples 3 and 4, however, the chunks take on a more interactive characteristic, displaying more attitudinal features, rather than merely filling the gaps.

Acoustic signal analysis based on pitch contour also shows the relationship between the various functions of formulaic sequences and the speed of delivery.

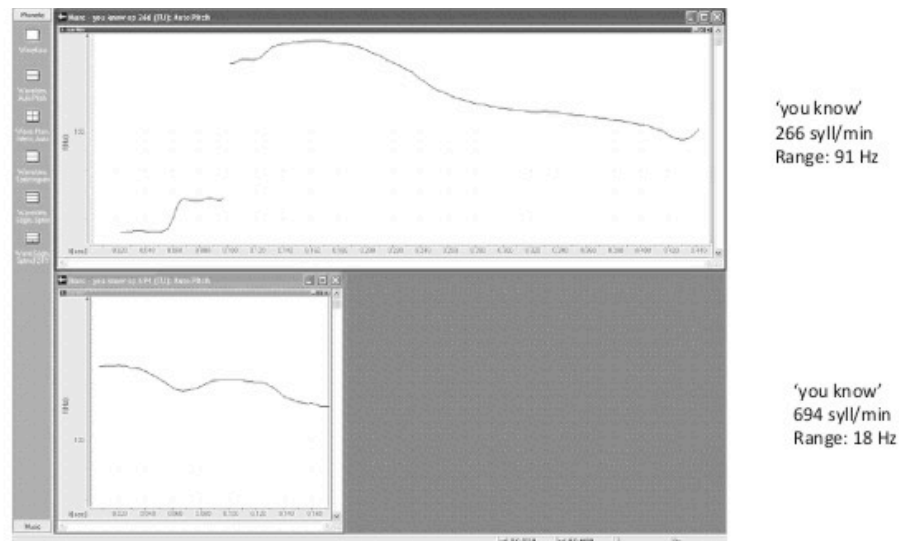


Figure 12: Pitch contours of FS 'you know' uttered at different speeds of delivery

The above graphic (Figure 12) is a representation of the chunk 'you know' uttered by the same speaker, in the same dialogue, but at two different delivery speeds. The intonation contour with the faster delivery speed of 694 syll/min shows a flatter curve than that with the slower speed of 266 syll/min. The difference in the lowest and highest value in the intonation curve is also given. This conforms to the interim finding that a faster rate means fewer or less obvious interactive features, while retaining a suitable intonation pattern. In addition, the intonation contour produced at the faster speed is more intact than the slower one. It suggests that a formulaic sequence delivered at a faster speed is more preassembled as a coherent unit in the mind than when uttered at a slower speed, and that there is more semantic and emotional engagement in the slower utterance.

Another example of the link between formulaic language and its speed of delivery and pitch contour is based on the analysis of the chunk ‘I mean’. As shown in Figure 13, based on the transcription, both occurrences of ‘I mean’ look the same. However, with access to the audio file, it is easy to hear the difference between them. The iterations of two ‘I mean’, spoken by the same speaker, are delivered within three seconds of each other, but occur in different environments and are produced differently. The first one is spoken in isolation and slowly, with the speed of 296 syll/min and with a pitch range of 50 Hz, in order to create thinking time and to prepare the next sequence. The terminal nasal ‘-n’ in ‘mean’ is drawn out. By contrast, the second one is part of a longer sequence, with a remarkably fast delivery speed of 624 syll/min and a flatter pitch contour of 15 Hz, and in isolation it is barely distinguishable. Therefore, it seems reasonable to conclude that not all formulaic sequences are equal, as addressed by Campbell, et al. (2008). The potential communicative functions cannot be accurately identified and understood without considering their phonological features, especially speed of delivery and pitch range.

	as you probably remember yourself from teaching it is
	i mean
	you learn more
[laughs]	that is one of the wonderful things about teaching
	i mean
	you learn you learn you learn as you are teaching
you do yeah	but then it becomes difficult to keep things simple
yeah	so erm
yeah	but ah
ah	yeah yeah

Figure 13: Examples of FS ‘I mean’

Based on the above basic principle, all thirty-three samples of ‘I mean’ spoken by both speakers in the half-hour dialogue (called ‘Marc-Donal’ for convenience; all conversations are labelled based on the names of the speakers) are analysed and graphed as shown in Figure 14, in which the vertical axis shows the speed of delivery (upper blue line) and pitch range (lower red line), while the horizontal axis represents the samples. From the graph, it can be seen that, although there are clear deviations in the middle speed range, the trend lines indicate that there seems to be an inverse relationship between speed of delivery and the pitch range.

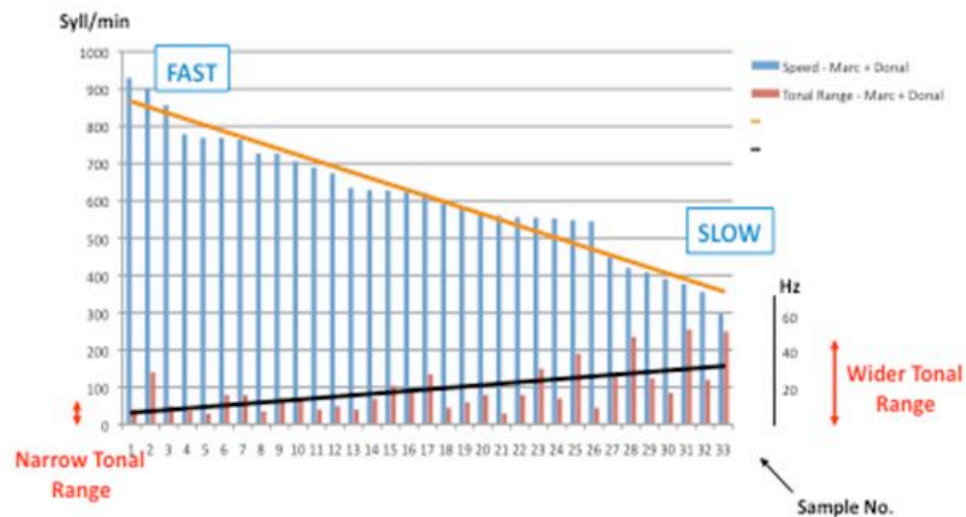


Figure 14: Different realisations of speed and pitch range of FS 'I mean'

B. Categories of FSs and their speed of delivery and pitch range

In order to fill in the gap of phonological realisations of various categories of formulaic language as stated earlier, the analysis of speed of delivery and pitch range is incorporated into the framework investigated in the literature. The calculation of speed, at the moment, is done manually with *Speech Analyzer* based on the dialogues between Marc and Aelish (Marc's mother), and Marc and Jenny (Marc's niece). The average speed of delivery for Marc is 400 syll/min, for Aelish 287 syll/min, and Jenny's average speed is 305 syll/min. Thus, the medium speed range in the present study is defined

between 300-400 syll/min. Pitch range analysis is carried out automatically with *PRAAT* based on the dialogues between Marc and Donal (Marc's colleague and friend), and Marc and Darragh (Marc's friend). The average pitch range for Marc is 74 Hz, for Donal 72 Hz, and for Darragh 78 Hz. The medium pitch range in this analysis is considered to be between 50-100 Hz.

The categories below, used for the current analysis, are based on the literature. The specific samples selected are informed by the audio assets available to the present author at that stage, and chosen to exemplify the categories emerging from the literature.

a) collocations

Firstly, the analysis of collocations is based on the formulaic sequences 'of course', 'and so on', 'kick out' and 'even though' drawn from the dialogue Marc-Donal. As the data in Table 5 show, the average speed of delivery is 377 syll/min, and the average pitch range is 39 Hz. Six out of ten samples are with a medium speed (between 300-400 syll/min) and eight are with a narrow pitch range (below 50 Hz). Five out of ten (50%) conform to both of these parameters. Collocations tend to be short and have a high frequency of co-occurrence, but can be discontinuous. Collocations function as semantic units.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
of course	(Donal) of course	392	37
	(Donal) and and of course our	337	50
	(Donal) he promised not to mention where it came from of course	323	30
	(Donal) of course there is a bad	256	90

	side to that as well		
	(Marc) of course I had a	302	62
and so on	(Marc) and so on	400	13
	(Marc) and so on	356	21
kick out	(Donal) I was kicked out after a year	446	32
	(Donal) he was kicked out of a pub	448	38
even though	(Donal) even though the ideas were not there or something	506	18

Table 5: Phonological realisation of collocations

b) semi-fixed frames

The next category examined is semi-fixed frames, i.e., templates which include variables. Five samples are chosen, as shown in Table 6. The average speed of delivery is 400 syll/min, and the average pitch range is 65 Hz. Three out of five samples are within the medium speed range (between 300-400 syll/min), and four of them are within a medium pitch range (between 50-100 Hz). Three out of five (60%) conform to both parameters. Semi-fixed frames, due to consisting of a fixed frame plus a variable, tend to be longer. As with collocations, semi-fixed frames are also semantic units, delivered at a medium rate. However, given that tailor-made elements need emphasis, the pitch range tends to be wider – up to a medium level (over 50 Hz).

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
here we go	(Donal) so here we go	367	58
on your left hand side	(Marc) on your left hand side	376	64
in good form	(Donal) he was in good form	334	73
a year and a half	(Donal) a year and a half ago	481	35
	(Donal) a year and a half I think	442	96

Table 6: Phonological realisation of semi-fixed frames

c) idioms

The third category analysed in this study is idioms. Given that idioms occur relatively infrequently (Strässler, 1982), only two examples (see Table 7) are chosen from those spoken by Marc in the dialogue Marc-Donal. The average speed of delivery is 186 syll/min, and the average pitch range is 150 Hz. As most of the idioms are opaque expressions to some extent, a slow speed of delivery (below 300 syll/min) and wide pitch range (above 100 Hz) are needed to highlight the metaphorical implications of the expressions. Both of the two samples (100%) conform to these two parameters. Idioms have semantic meaning as well, they can be any length, but mostly tend to be longer. Idioms, even although prefabricated, stored and retrieved as a unit, can be used either wholly or in part, but the elements used need to remain frozen.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
from pillar to post	(Marc) from p- pillar to post	240	113
out of a seed	(Marc) out of p- ah ah ah a seed	132	187

Table 7: Phonological realisation of idioms

d) chunks

The analysis of chunks is more complicated than the other categories. As discussed earlier, there are two completely different kinds of phonological realisations of delivery speed and pitch range in the chunks ‘you know’ and ‘I mean’. These chunks differ from semantic units, which mainly focus on expressing semantic meaning (either transparent or opaque), in that they are of low semantic value, and mainly serve as an interactive unit by providing various discourse markers to progress the natural flow of the

conversations. Depending on the different functions they realise, chunks can be defined as unmarked or marked. Unmarked chunks, e.g., back-channelling, characterise most of the chunks produced unconsciously and automatically and are more likely to be used for filling in gaps. Marked chunks, in contrast, tend to show more engagement and more interactive features in the conversations. Therefore, the same orthographic chunk of a sequence can be marked or unmarked, depending on its phonological realisations.

i) unmarked chunks

Table 8 below shows the data of unmarked chunks based on fifteen samples. The average speed of delivery is 427 syll/min, and the average pitch range is 29 Hz. Ten out of fifteen are within a fast speed range (over 400 syll/min), and all these fifteen samples are within a narrow pitch range (below 50 Hz). Ten out of fifteen (67%) conform to both these parameters.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
or something	(Donal) or something	437	28
	(Donal) even though the ideas were not there or something	674	29
	(Donal) or something	471	24
I suppose	(Marc) I suppose	348	23
	(Marc) I suppose	358	43
	(Donal) particularly for OLDer people I suppose	410	26
I must say	(Marc) I must say	433	13
you see	(Donal) you see	421	22
	(Donal) you see	407	37
	(Marc) well you see	432	38
	(Marc) you see Donal	453	30

	(Marc) you see	612	38
I guess	(Marc) I guess you are right	292	26
I see	(Marc) I see	264	35
	(Donal) aw I see what you mean	396	20

Table 8: Phonological realisation of unmarked chunks

ii) marked chunks

Eight samples (see Table 9) are chosen for the analysis of marked chunks. Given that unmarked chunks are delivered at a fast speed and a narrow pitch range as discussed above, it is anticipated that marked chunks, in order to perform more interactive functions, will have a wider pitch range. Seven of the samples are clearly within the medium pitch range. The situation with the speed parameter is less clear, however. The average speed is 375 syll/min, i.e., within a medium speed range. However, not one sample falls within this range (300-400 syll/min). It would therefore seem advisable to separate the samples into a fast group (over 400 syll/min) and a slow group (below 300 syll/min), as shown in Table 10 and Table 11.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
I suppose	(Marc) I suppose a a a big thing for	295	87
	(Marc) I I suppose there is one thing that really	432	90
	(Donal) well I suppose you can	759	60
I must say	(Marc) I must say	271	77
I guess	(Marc) so I guess we are in	474	94
I see	(Marc) I see	249	65
	(Marc) I see	256	54
	(Marc) I see	264	35

Table 9: Phonological realisation of marked chunks

Table 10 shows five samples with a slow speed of delivery of an average 267 syll/min. The average pitch range is 64 Hz. All the five samples are within a slow speed range (below 300 syll/min), four of them are within a medium pitch range (between 50-100 Hz). Four out of five (80%) conform to both parameters.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
I suppose	(Marc) I suppose a a a big thing for	295	87
I must say	(Marc) I must say	271	77
I see	(Marc) I see	249	65
	(Marc) I see	256	54
	(Marc) I see	264	35

Table 10: Phonological realisation of marked chunks with a slow delivery speed

Three samples with a fast speed of delivery are shown in Table 11. The average speed is 555 syll/min, and the average pitch range is 81 Hz. All three samples are within a fast speed range (over 400 syll/min) and a medium pitch range (between 50-100 Hz). All the samples (100%) conform to both parameters.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
I suppose	(Marc) I I suppose there is one thing that really	432	90
	(Donal) well I suppose you can	759	60
I guess	(Marc) so I guess we are in	474	94

Table 11: Phonological realisation of marked chunks with a fast delivery speed

In short, as discussed above, the phonological realisation of chunks is rather complex. Depending on the different functions in the discourse, chunks can be defined as

unmarked or marked. Unmarked chunks normally are of a fast speed of delivery (over 400 syll/min) and a narrow pitch range of below 50 Hz. By contrast, marked chunks are identified as having a medium pitch range, and can be realised at either a slow speed of below 300 syll/min, or a fast speed of over 400 syll/min. Chunks, characterised by short sequences, high frequency of co-occurrence and not amenable to unpacking, play an important role in formulaic language.

e) grammatical paradigms

The last category examined here is grammatical paradigms. Grammatical paradigms are grammatical units which tend to be longer, with a relatively frozen grammatical frame plus a variable, depending on various tenses and registers, to structure an utterance and build up an expression. Nine samples are drawn from the dialogue Marc-Donal, as shown in Table 12. The average speed of delivery is 694 syll/min, and the average pitch range is 46 Hz. Eight out of nine samples are within a fast speed range – over 400 syll/min or more, which is in line with Cruttenden's (1997, p.17) analysis of the most common vowel reductions occurring in auxiliary verbs. Five of them are of a narrow pitch range of below 50 Hz. That is to say, five out of nine samples (56%) conform to both parameters.

Example	Flow Unit	Speed (syll/min)	Pitch Range (Hz)
was to	(Donal) the plan was to be	383	52
	(Donal) my my plan was to be a political cartoonist	573	75
were not able to	(Donal) I was and you you were not able to continue doing them	629	30
must have	(Donal) I must have sent you	591	81
	(Marc) it must have been	577	41

is going to	(Donal) it is going to be	729	37
	(Donal) January is going to be spread out	1043	29
	(Donal) which is going to make life much easier for	870	14
was going to	(Donal) but ah naw I was going to	854	56

Table 12: Phonological realisation of grammatical paradigms

In conclusion, based on the analysis of natural interactive dialogues, some links between different categories of formulaic language and their realisations of relevant phonological characteristics, e.g., speed of delivery and pitch range, are investigated in the present study. A comprehensive tabulation is laid out in Figure 15. This is the investigation the present author aims to add to the study of formulaic language, which is suggested as only a starting point. More data is needed to facilitate the analysis of formulaic language and prosody.

Category		Example	Speed (syll/m)			Pitch Range (Hz)			Function
			0-300	300-400	400+	0-50	50-100	100+	
Collocations		of course		✓		✓			Semantic
Semi-fixed Frames		on your left hand side		✓			✓		Semantic
Idioms		from pillar to post	✓					✓	Semantic
Chunks	Unmarked	I must say / or something			✓	✓			Interactive
	Marked	I must say	✓		✓		✓		
Grammatical Paradigms		is going to			✓✓	✓			Grammatical

Figure 15: Towards inclusion of prosody in FS typology

C. Position of FSs in flow units and their speed of delivery

Apart from the investigation of various categories of formulaic language and its phonological realisations, another analysis correlates to the relationship of the positions of formulaic sequences and their speeds of delivery. Various locations of a formulaic sequence within a flow unit seem to trigger different speeds of delivery. This phenomenon is examined on the basis of the following three positions – FS followed by a variable, FS preceded by a variable, and FS preceded and followed by a variable. Fifteen samples respectively consisting of the chunk ‘you know’, ‘I do not know’ and ‘do not worry’ are chosen from the dialogues Marc-Donal, Marc-Aelish, and Marc-Jenny. Two kinds of delivery speed are given for each sample for comparison purposes. The first speed indicates the delivery speed of the whole flow unit, and the second one is calculated deliberately for the formulaic sequence part in these samples.

Seven samples are structured with a formulaic sequence followed by a variable, as shown in Table 13. By comparing the speeds of delivery of the chunk ‘you know’, ‘I do not know’, and the delivery speeds of the whole flow unit, it can be seen that the flow units with formulaic sequences followed by a variable are delivered at a much slower rate than the individual formulaic sequences on their own. The current investigation leads the present author to believe that it is due to the appended terminal variable, which is specifically generated in a rule-based fashion, and therefore slows the delivery of the whole flow unit.

Sample	Speaker	Orthographic Text	Speed (syll/min)
1	Donal	You know I had a fif...	455
		‘you know’	571
2	Donal	You know something that I noticed when we overlapped at the very beginning	472
		‘you know’	619
3	Donal	You know a fear of dealing with a	514

		'you know'	678
4	Donal	You know the mistakes that come up during the exam	507
		'you know'	779
5	Aelish	I do not know why	444
		'I do not know'	684
6	Aelish	I do not know too much about	351
		'I do not know'	407
7	Marc	I do not know it at all	463
		'I do not know'	478

Table 13: Position of FS and its delivery speed (FS followed by a variable)

For the structure of formulaic sequences preceded by a variable, two samples are given in Table 14. In contrast to the results of Samples 1-7, the formulaic sequence preceded by a variable in Samples 8-9 slightly speeds up the whole flow unit. A study, carried out by Underwood, et al. (2004, p.162), on the eye movement control in the processing of formulaic sequences, shows that the terminal word of a formula is processed more quickly than the equivalent word in a non-formulaic text, due to the context provided by the formulaic sequence facilitating the processing. Even though their study is based on the reading out of written text, as opposed to speech production, there seems to be a link between them, which indicates that the formulaic sequence is stored and retrieved holistically and thereby allows for a faster utterance overall.

Sample	Speaker	Orthographic Text	Speed (syll/min)
8	Donal	Ah you know	313
		'you know'	291
9	Jenny	But I do not know	577
		'I do not know'	544

Table 14: Position of FS and its delivery speed (FS preceded by a variable)

Table 15 shows the samples of formulaic sequences preceded and followed by a variable. The analysis turns up contradictory results. Samples 10-13 and 15 show that the speed of delivery of the formulaic sequence on its own is faster than the speed of the whole flow unit. While Sample 14 shows the opposite result, in that the whole flow unit is delivered slightly faster than the individual formulaic sequence. Other factors are anticipated to contribute to the realisation of speed of delivery, e.g., the individual delivery speed of the variables, which require further investigation.

Sample	Speaker	Orthographic Text	Speed (syll/min)
10	Donal	Well you know more than me	418
		'you know'	645
11	Marc	b- you know you	400
		'you know'	545
12	Marc	Which you know was a v-	434
		'you know'	1176
13	Marc	Do you know what I mean	629
		'you know'	1111
14	Marc	So you know the notion of community	581
		'you know'	531
15	Marc	So do not worry about it	563
		'do not worry'	676

Table 15: Position of FS and its delivery speed (FS preceded and followed by a variable)

Even though the analysis of the position of formulaic sequences and their realisations of speed of delivery is inconclusive to some extent, however, early indications from the data collected are that speed-indexing might shed an interesting light on the link between the different positions of formulaic sequences and their speed of delivery. A wider range of samples is needed, and other elements which might affect the speed of delivery also need to be considered in further investigation.

7.2.4 *Limitations of Analysis*

In analysing authentic, dynamic speech assets, some natural interactive features in L1-L1 English speech have been discovered and highlighted, such as formulaic language and its various phonological realisations. However, there are also some limitations in the analysis which require for further investigation.

First of all, the main analysis of formulaic language is based on the speed of delivery and pitch range. Given that the speed calculation is mainly done manually, this inevitably introduces a degree of subjectivity.

The analysis given above is restricted to the limited data resources available, and mainly focuses on one prominent speaker – Marc, and his family members, his friends, his colleague and student. The advantage of this is to highlight the different reactions of the same speaker with different relationships. The disadvantage, however, is that it might result in the range of samples being relatively narrow, either with regard to linguistic characteristics, or phonological realisations, i.e., the setting of various ranges of speed of delivery (slow, medium and fast) and pitch range (narrow, medium and wide). While the present findings are indicative, a larger data set and a wider range of samples would be required to confirm the initial findings.

The analysis of the various positions of formulaic sequences and their different realisations of speed of delivery is inconclusive. More elements are assumed to be linked with the speed realisation, e.g., the individual delivery speed of the variables within the flow unit, or other flow units before or after the current unit but within a larger speech envelope. Apart from the linguistic perspective, other disciplines, e.g., psychology or analysis of neural activity, might also throw light on the subject.

Therefore, more evidence is needed to demonstrate whether the position of the formulaic sequence affects the realisation of speed of delivery, what factors contribute to this realisation and how their effects influence the speed of realisation. This also forms the basis of further research as detailed in Chapter 10.

In general, initial investigations on formulaic language based on the analysis of some early dialogue assets of the DSC, as pointed out earlier, are limited in scale. As the development of the speech corpus progresses, more data will be available, and more evidence will lead to an objective and reasonable conclusion.

7.3 Conclusion

This chapter mainly details the progress to date by the present author in researching and analysing natural, authentic, dynamic English speech. Based on transcribing the dialogue as orthographic text, segmenting and tagging the sequences, initial findings on formulaic language and its phonological realisations of speed of delivery and pitch range are investigated. Some limitations are left for further research.

As reviewed in Chapters 2 and 3, formulaic language is frequently employed and produced by L1 language speakers in native-to-native informal English speech, and it is also particularly subject to the phonetic reductions and prosodic variations. In addition, a key problem for most L2 learners, especially Chinese EFL learners, is to cope with the phonetic and prosodic realisations of formulaic sequences in natural, casual conversation. Therefore, the current research focuses on the investigation of formulaic language and its variations as characterised by speed of delivery and intonation patterns.

Based on the research work undertaken above, and in order to demonstrate how the speech technologies, e.g., the slow-down facility, combined with real, natural, interactive L1-L1 English speech, can be of help, the application and evaluation of these technologies to the acquisition of formulaic language by Chinese language learners is further investigated in the next chapter. The design and implementation of tests and the methodologies adopted are also discussed in Chapter 8.

8. Speech Technologies Applied to Acquisition of Formulaic Language

8.1 Introduction

In the previous chapter, on the basis of tagging and analysing real, natural, interactive English speech, initial research on formulaic language is investigated. Some questions for further research are also discussed. The communicative value of formulaic language cannot be simply interpreted by its written form; its communicative value often lies in its prosody. For Chinese EFL learners, influenced by their mother tongue as reviewed in Chapter 3, their English speech often sounds flat and staccato. Therefore, by using formulaic language as exemplar material and combining this with speech technologies, is anticipated that this will lead to more efficient L2/EFL learning and an enhanced pedagogy.

In this chapter, in order to evaluate the use of speech technologies in language learning, especially the slow-down algorithm, and particularly when applied to the natural, authentic English in the nascent Dynamic Speech Corpus, the application of these technologies to the acquisition of formulaic language by Chinese language learners is described. The design and procedures for both tests and training period, and the methodologies for evaluation are also detailed in this chapter. The chapter concludes by considering the considerable improvement and positive feedback from the Test Group.

8.2 Application and Evaluation of Speech Learning Technologies

Research of the application and evaluation of innovative speech technologies which happened to be available to the author in DIT was carried out in four phases – specification and design (described in Chapter 6), Test 1 (detailed in Section 8.2.1),

training sessions (described in Section 8.2.2) and Test 2 (detailed in Section 8.2.3). Specification and design were carried out between March and May 2007. Test 1 was implemented in June 2007, followed by training sessions over six months from September 2007 to April 2008 (excluding January 2008 and February 2008 which was the Winter Holiday period for the Chinese students), and the last phase was mainly for Test 2 which was carried out in May 2008.

The design and implementation of the tests, procedures for the training sessions, and methodologies for evaluation are described in detail in this section.

8.2.1 Test 1

Test 1 is the initial benchmark test, carried out in June 2007. One hundred Chinese students participated in three different schools in Anshan. The result of Test 1 was used as the baseline for comparison with Test 2. The detailed procedures for the design and implementation of the test, and evaluation of the test data are described as follows.

8.2.1.1 DESIGN OF TEST 1

A. Design principles and requirements

Choosing an appropriate speech testing sample which ‘adequately characterizes the overall potential language use of the individual’, as Bachman (1990, p.11) points out, is one of the primary premises needed to be considered when designing a test. The sample must be an indication of language in a natural environment, rather than specially devised for the test. In order to arrive at useful data, validity needs to be built into the design stage. Therefore, two key principles are considered when designing and choosing samples, for both Tests 1 and 2, and for the training sessions.

The first principle is that all samples are taken from the natural, dynamic English speech recorded for use in the Dynamic Speech Corpus (DSC). As discussed earlier in Chapter 5, these recordings are resources which can demonstrate the natural usage of L1-L1 English speech. All the authentic, interactive characteristics of speech are reflected in these dialogues, rather than scripted and rehearsed dialogues or monologues. The second principle is that, apart from natural, interactive dialogues, the slow-down algorithm was also employed in the experiment (with the test group only), given that the need for the slow-down is because (speed-induced) reductions and various phonological realisations are the main domains investigated in this study. In view of the fact that formulaic language is prone to reductions and modified phonological characteristics (which will be demonstrated below), formulaic sequences are the subject matter of the investigation, and form the bulk of the testing and training samples.

In order to reflect precisely the language ability of participants in using the target language, vocabulary frequency is another factor which needs to be considered when choosing test samples. Consequently government issued vocabulary lists for the test students were consulted before drawing up the test materials. Out of seventy-six items (i.e., words) used in Test 1 samples, seventy-one items (93%) had been acquired by seventy (70%) of the participants (100 participants coming from three different educational levels; see Section 8.2.1.2). Only 7% of the items are new words for thirty (30%) of the participants. Thus, in practice, because of care taken in the selection of materials, low frequency words and unfamiliar vocabulary were kept to a minimum.

With these pragmatic considerations underpinning the choice of samples, the design of Test 1 is, on the one hand, to show the significant, easily ‘missed’ linguistic features in L1-L1 English speech, and on the other hand, to demonstrate initially to non-L1 English

speakers, especially Chinese EFL learners, what formulaic language is, in real informal English conversations.

B. Choice of test snippets

In all, forty-six snippets with significant linguistic features were chosen from four authentic, interactive dialogues (totalling approx. 2 hours of recording) which had been recorded at that stage. As shown in Appendix 2, all these snippets contain at least one formulaic sequence. Apart from formulaic language, all these forty-six snippets exhibit important linguistic characteristics which Chinese learners of English need to learn in order to cope with informal dynamic speech, i.e., lively intonation patterns, weak forms and reduced forms, as reviewed and discussed in Chapters 2, 3 and 4. The remainder may also have these features, but were not used, either because the vocabulary involved was low frequency, or because they were grammatically too complex or the sequences were too long for oral recall.

Based on the testing principles and the procedures for the training sessions, these forty-six snippets, along with some other new snippets taken from new recorded data, were divided into three parts – respectively used as testing samples for Test 1 (details given below), as testing samples for Test 2 (see Section 8.2.3.1), and as target materials for the training sessions (see Section 8.2.2.3). Ten of them were chosen as the testing samples in Test 1 (see Table 16), as justified below. All these ten samples set out to exemplify salient linguistic characteristics, such as elision, assimilation, formulaicity. All of them were good examples to illustrate real, natural, informal native English speech to non-L1 language learners. The detailed characteristics of these samples are laid out as follows. The items highlighted in red and bold indicate formulaic sequences

exhibiting natural blur and meriting a score of ‘4’ in the evaluation system (see details in Section 8.2.1.3).

Sample	Orthographic Text	Linguistic Features	Speed (syll/min)	Pitch Range (Hz)	Length (no. items)
1	Which one are you thinking of?	1. weak form of 'are', produced as /ə/, 2. reduced 'you' to /je/, 3. grammatical paradigm FS ‘are you’, with the speed of 714 syll/min and pitch range of 13 Hz, 4. clearly pronounced /v/ in ‘of’, 5. collocation ‘think of’, with the speed of 437 syll/min and pitch range of 73 Hz	467	77	6
2	From pillar to post.	1. idiom with lively intonation pattern, 2. slow speed and wide pitch range	241	113	4
3	Certainly made up for that.	1. lively stress intonation pattern, 2. collocation ‘make up for’, with the speed of 457 syll/min and pitch range of 29 Hz	458	88	5
4	I moved out of home when I was eighteen.	1. reduced ‘moved out of’, produced as /mu:vdaʊ də/, 2. collocation ‘move out of’, with the speed of 471 syll/min and pitch	401	97	9

		<p>range of 83 Hz,</p> <p>3. weak form of 'I' in 'I was', produced as /ə/,</p> <p>4. grammatical paradigm FS 'I was', with the speed of 670 syll/min and pitch range of 9 Hz,</p> <p>5. double word stress on 'eighteen'</p>			
5	<p>Do you know</p> <p>what I mean?</p>	<p>1. reduced 'do you know',</p> <p>2. weak form of 'what I mean',</p> <p>3. chunk with fast speed of delivery</p>	595	71	6
6	<p>Fifty sixty</p> <p>percent of the people.</p>	<p>1. barely distinguishable 'of the' (= uh th),</p> <p>2. collocation 'percent of', with the speed of 655 syll/min and pitch range of 46 Hz,</p> <p>3. intonation indicates a range or a more exact figure,</p> <p>4. big intonational change</p>	416	95	6
7	<p>You are scraping the bottom of the barrel sir.</p>	<p>1. idiom 'scrapping the bottom of the barrel' with hyper-articulation at the speed of 485 syll/min and pitch range of 58 Hz,</p> <p>2. weak form of 'of', produced as /ə/,</p> <p>3. fast speed in unstressed elements,</p> <p>4. expressiveness of this</p>	499	59	9

		idiom fits into the expressive envelope: ‘Aw, Donal - How could you tell such a sick joke?!?’			
8	You have a list of questions?	1. reduced 'you have a', produced as /juvə/, 2. huge reduction on the non-lexical word ‘of’, 3. collocation ‘a list of’, with the speed of 732 syll/min and pitch range of 16 Hz, 4. question intonation pattern marked with a rising tone at the end, rather than a grammatical structure	565	57	6
9	Well that is exactly what the Italians would have been doing.	1. weak form of 'what', produced as /wə/, 2. reduced 'would have been' to /wudəbin/, 3. grammatical paradigm FS ‘would have been’, with the speed of 1047 syll/min and pitch range of 10 Hz, 4. really fast speed of delivery, 5. huge intonational change	788	155	11
10	I cannot walk down the street without having to walk on the	1. it is hard to hear that 'can't' is negative (and not ‘can’. Negativity is made clear to the L1 listener by	535	114	14

	road.	extending the ‘-n-‘ and following with ‘without ...’, NOT by articulating the ‘-t’ in ‘can’t’), 2. extreme reduction of ‘without’ to /wiə-/, 3. reduction of ‘having to’ to /havn to/, 4. grammatical paradigm FS ‘having to’, with the speed of 1071 syll/m and pitch range of 9 Hz, 5. intonation and extension of ‘road’ indicate further phrase to follow, 6. huge intonational change			
--	-------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--	--

Table 16: Testing samples in Test 1

C. Justification for samples chosen

For Chinese learners of English, since they are more accustomed to the clearly pronounced citation forms produced by their language teachers (as discussed in Chapters 2, 3 and 4), weak forms or reduced forms of words always raise problems in listening and decoding native English speech. Samples 1, 4, 5, 6, 7, 8, 9 and 10 all have contracted linguistic characteristics. In Sample 1, reduced ‘you’ is shown to be compared with the citation form /ju:/ which is the most likely accepted pronunciation by Chinese students. The grammatical word ‘are’ is also reduced in Sample 1. Samples 4 and 9 give an example of how a pronoun word, e.g., ‘I’, ‘what’, can be reduced in informal English speech. Sample 5 has both a reduced form of ‘do you know’ and a

weak form of ‘what I mean’ which speed up the production to 595 syll/min. Non-lexical words are very much reduced in the rapid flow of informal speech. Samples 6-10 demonstrate the various reduced forms of different particle words which are important segmental information Chinese learners rely on in interpreting the speaker’s utterance.

Secondly, Chinese speakers, influenced by their mother tongue, tend to utter their English speech with a flat intonation pattern and a narrow pitch range, which often causes problems for L1 English speakers (as examined in Chapters 2, 3 and 4). Samples 2, 3, 6, 8, 9 and 10 are very rich in their intonation and tonal expressions. Sample 2 is a formulaic sequence with a wide pitch range of 113 Hz. Sample 3 consists of a collocation with lively stress intonation pattern. The intonation of Samples 6, 8 and 10 all communicate more than the constituent words. For example, Sample 6 gives a more exact figure, rather than merely a number, Sample 8 shows a question intonation pattern instead of a statement, and Sample 10 indicates there is a further phrase to follow. In addition, the pitch range in Sample 9 is a very large 155 Hz. All these are very significant supra-segmental linguistic features absent in Chinese language learners’ English speech.

Thirdly, as the phoneme /v/ does not exist in the Chinese pinyin system as considered in Chapter 3, most Chinese speakers tend to replace it with /w/ or /f/. Sample 1 provides a clear example of the production of the consonant /v/.

Finally, as reviewed and discussed in Chapters 2, 3 and 4, very few Chinese learners have any knowledge of formulaic language use in English and this frequently results in problems in their spoken English communication. Therefore, one of the aims of Test 1

is to demonstrate formulaic language to Chinese language learners. All of these samples consist of different types or sub-types of formulaic sequences. Some of them are easier to access by Chinese learners either in their language classes or course-books, such as, the sequences in Samples 1, 3, 4, 6, 8 and 10. A few of them, however, are very difficult and cannot be acquired naturally and automatically, such as those in Samples 2, 5, 7 and 9.

D. Anticipated difficulty of test samples

Considering all these linguistic features, Sample 9 is anticipated to be the most challenging one due to its very fast speed of 788 syll/min, wide pitch range of 155 Hz and a longer string of 11 items, which overloads the storage capacity of the Short Term Memory (STM) of seven (plus or minus two) pieces of information (Miller, 1956). Sample 10 is also considered to be of great difficulty with an even longer sequence than Sample 9, relatively fast speed and wide tonal range. Sample 7 is graded as the third most difficult one. The main reason is its formulaicity which is not familiar to Chinese language learners, its nine-item length and fast delivery of unstressed elements. The next one thought to be challenging to imitate is Sample 5 which is a formulaic sequence with a really fast speed of delivery made up of large chunks of reduced and weak forms. Among these samples, Snippet 2 is considered to be the easiest. Even though it is a formulaic sequence with a relative wide pitch range of 113 Hz, yet it is a short string with only four items and a really comfortable delivery speed of 241 syll/min, which is easy to be perceived and imitated by participants.

8.2.1.2 SUBJECTS FOR TEST 1

A. Composition of Test Group and Control Group

The choice of subjects in Test 1 is a population of convenience available to the present author. After she had contacted some of her former colleagues with whom she had previously taught English in China, 100 Chinese learners of English were made available for the tests. They are either studying at a third level college or at a junior secondary school. Three different levels of linguistic competence among these participants were deliberately chosen for comparison of the evaluation of the application of speech technologies and also for providing more information for further study in this area.

- a) Test Group A (GpAt) and Control Group A (GpAc) are all first year students at a third level college in which there are fifteen students in each of the groups.
- b) Test Group B (GpBt) and Control Group B (GpBc) are third year students at a public junior secondary school in which there are twenty students in each group.
- c) Test Group C (GpCt) and Control Group C (GpCc) are students chosen from the first year in a private junior secondary school, and fifteen students are in each group.

The students in the test group and the control group within the same level (GpA, GpB and GpC) are taught by the same teacher.

Before the tests, an ethics form was given and signed by all the subjects in both Test Group and Control Group.

B. Questionnaire 1 (Control Group and Test Group)

Questionnaire 1, written in Chinese (see Appendix 3 for an English version), was implemented prior to Test 1 to assess the basic background information of the

participants. There was a 100% rate of return of questionnaires distributed. An analysis is shown as in Table 17.

Group	Gender		Age (avg.)	Mother tongue	Fluent in any other language	No. Yrs learning English (avg.)	No. Hrs English class /wk (avg.)	No. Mins listening to English /wk (avg.)	No. Mins speaking English /wk (avg.)
	M	F							
GpAt	0	15	20	Chinese	No	8	12	200	260
GpAc	0	15	20	Chinese	No	9	12	230	267
GpBt	6	14	15	Chinese	No	6	4.5	109	76
GpBc	8	12	15	Chinese	No	5	4.5	79	60
GpCt	5	10	12	Chinese	No	5	4.5	106	106
GpCc	5	10	13	Chinese	No	6	4.5	127	127

Table 17: Questionnaire 1 of participants in Test 1

From the questionnaire, it can be seen that all of the participants are Chinese L1 speakers without the influence of any other foreign languages, and all of them have the same exposure to English at school as their fellows within the same group. Both GpAt and GpAc have fifteen female students with an average age of 20. GpAc, however, has been learning English longer than GpAt, and they also have more exposure to English listening and speaking than GpAt. The situation for Group B is a little different. There are twenty participants in each group with an average age of 15. There are fourteen females and six males in GpBt, but twelve females and eight males in GpBc. The participants in GpBt started to learn English somewhat earlier than their peers in GpBc, and also have more practice in English listening and speaking than GpBc. As with GpCt and GpCc, there are fifteen students in each of the groups, ten females and five males. The average age for GpCt is 12, but 13 for GpCc. Participants in GpCc have more

experience in learning English and spend more time exposed to English listening and speaking exercises than those students from GpCt.

8.2.1.3 METHODOLOGY FOR TEST 1

Given that the aim of the tests is to evaluate the correct imitation and production of the snippets produced by the participants, therefore, verbatim recall (Field, 2004, p.318) was adopted as the method of eliciting of test data for both Test 1 and Test 2. The reason for using verbal retrieval is to minimise the interference of irrelevant factors so as to increase the validity of data collection. Dictation or cloze test, which is a common means for listening tests, is avoided in this experiment. On the one hand, since the wrong answer may arise from a failure to understand the listening contents or to transfer the understanding into written form, i.e., wrong spelling, dictation cannot correctly indicate the ability of language learners in using English. On the other hand, dictation or cloze tests themselves are more suitable for testing listening comprehension, rather than spoken production and intelligibility. Whether or not intelligibility was achieved by the test subjects imitating and re-producing the exact utterance cannot be ascertained by this means of dictation or cloze testing.

By using verbal recall to elicit testing data, an important time factor was also considered. Based on Field (2004), accurate verbal recall is relatively ‘short-lived’, especially with auditory input in which the trace of the spoken utterance decays in 1-2 seconds (Field, 2003a, p.111). Therefore, before the test, an instruction was given to the participants. All subjects were given a constant elicitation input – listen to the recording, then imitate immediately. Time pressure can also lead to another advantage by ensuring that the performances produced by participants come from input recall, rather than meaning reconstruction.

Therefore, ten samples were played one by one. Samples 1-8 were played three times in quick succession, Samples 9 and 10 were played five times due to their length (based on a pilot test by using an L1 speaker of English as the testee). Since this was a testing rather than a training exercise, all of the samples were played at normal (100%) speed only. After listening to the recording, the participants were encouraged to imitate and record what they heard immediately, even if they could only capture the intonation pattern or some of the phonemes, rather than the words or the whole utterance. Each student was tested under the same conditions, and tested individually so that no student could hear the performance of any other student. All the participants in each group were tested on the same day. Each response produced by the participant was recorded as a WAV file and was given a unique name, for example,

File A_T1_1: Test 1 data produced by Test Student 1 from GpAt (the third level college)

File B_C1_10: Test 1 data produced by Control Student 10 from GpBc (public junior secondary school)

File C_T1_15: Test 1 data produced by Test Student 15 from GpCt (private junior secondary school)

All the data in Test 1 were collected by the present author using the same methodology.

After the data collection stage, the following step describes how data were evaluated so as to gain reliable and relevant information. This is a procedure based on the assessment of required information, rather than error detection, and aims to measure the participant's overall ability with respect to intelligibility by imitating and re-producing native-like formulaic speech. Therefore, the evaluation of Test 1 was based on the methodology of listening to and manually scoring the productions, which includes the

measure of the participant's performance on both lexico-grammatical accuracy and prosodic pragmatics, e.g., intonation patterns.

The first stage in evaluating student performances is as follows. Firstly, the native speech of each snippet is written out in an Excel spreadsheet with one word per column. The full citation form of a word is used if any of its phonemes are present in the speech signal. The analysis of the data was carried out both qualitatively and quantitatively. Given that the existing evaluation method in the literature, such as Jones and Haywood (2004), Kim (2006), does not fit the aim of this study, the present author had to develop a new intelligibility rating scale for this experiment. Therefore, a 5-point Likert evaluation scale has been devised which is capable of providing the greatest amount of information, in which

0 = word not spoken or indecipherable; no segment perceived

1 = word indecipherable; some segments correct

2 = word decipherable; some segmental errors

3 = citation word identified correctly – or at least repeated by the participant. This score can also represent a students' comprehension of NS blur, but avoidance of or failure to produce the blur themselves

4 = native speech flow re-produced successfully (principled, 'intelligent' blur, as opposed to accidental production of segment or blur). Surrounding words must attract a positive score of at least '2'.

The 5-point evaluation system adopted in this study consists of a five point scale, ranging from 0 to 4 without midpoints. Score 0 indicates no participant performance, or no correct phonemes produced by the participant. Score 1 means at least one phoneme is correct in isolation, but the word as a whole is not recognisable. In contrast, Score 2

indicates that the whole word is identifiable, but entails individual phoneme errors. Score 3 means a correct citation form of the word is imitated and produced by the participant. Score 4 is defined so as to represent a native-like speech production with an intact intonation contour and an intelligent blur, rather than an accidental parroting. This productive level of performance is achieved based on a top-down approach. To distinguish between parroting and a principled production of native-like blur, and in order to achieve a score of '4', it is necessary that there is corroborative evidence of intelligibility left and right of the blurred sequence, in the form of convincing segmental performances scored as either a '2' or a '3'. A '4' is a qualitative score, and meant to separate out participants capable of producing principled blur. It is not intended as a reward for a 'better' performance, but is one of the aims of the present study to help Chinese learners to overcome the east-west divide (as discussed in Section 3.2.3) in pronunciation and prosody and learn to apply the top-down approach.

One of the interesting things about this evaluation scheme is that it allows for a distinction to be made between a low-level parroting and a principled production of native-like blur. The fact that weaker students scored lower grades in the lexical items surrounding the blur highlights the fact that intelligibility is low overall and any coincidence of NS and NNS blur is fortuitous. Where, on the other hand, the 'intelligibility grades' of words surrounding the NS blur show more positive values, this is a strong indication that any '4' awarded is the product of principled production on the part of the participant.

During the process of evaluation of the data, two issues needed to be considered. One is the constant assessment criteria. Based on the 5-point scoring system, each word is assessed and given a score between 0-4, the important thing is to apply constant criteria

to all the data. Therefore, some data were listened to and re-evaluated several times so as to enhance their validity. Another issue, as argued by the present author, is on the method of evaluating the correctness of individual phonemes. The aim of this study is to investigate intelligibility by means of imitation and re-production of a natural flow of English speech, especially embedded within the acoustic blur and other phonological variations of formulaic language, rather than the evaluation of discrete segmental sounds of English. A general evaluation method is thus adopted based on a comparison between the master sample and the participant's performance in terms of overall intelligibility, with regard both to segmental phonemes and intonation pattern (blur). This is deemed to be a suitable evaluation method to correctly assess the real language ability of the participants in non-test situations.

Based on these central principles and methodology, after the participant's production was listened to, an appropriate score between 0-4 was assigned. The evaluation was done word by word, but also assessed within the speech envelope encapsulating the prosodic characteristics under investigation. The maximum score for each word was 4, and a score from 0 to 4 was given to each word produced by the participant. A sum was got for each snippet by adding all the scores of the items in this snippet. Then a correct percentage was calculated depending on the sum achieved. Based on this methodology, all the data was evaluated in accordance with the first version of the evaluation (see Appendix 4).

However, this first version was later refined in preparation for Test 2 for two reasons. Firstly, Score 4 should only be given to those items with 'intelligent' blur. Compared with the master samples, in which the exemplar speaker did not produce a blur for every word, most of the items given were citation forms. It would therefore be unreasonable

to set the maximum score for each item at ‘4’. Secondly, given that the aim of the present research was to validate the usefulness of the slow-down facility for enhancing the intelligibility of Chinese English learners when producing formulaic language (which is consistent with the second research question), thus, Score 4 should only be applied to the formulaic sequences themselves, rather than extensively assigned to the whole sample. Due to these two considerations, then, Score 4 was re-set only to those items which were formulaic sequences and produced with real natural blur.

For example, there are six items in Sample 1 – ‘which one are you thinking of’. There are two formulaic sequences embedded – ‘are you’ and ‘think of’. The blur, however, only occurred on the former grammatical paradigm. Therefore, the maximum Score 4 was only given to this item, leaving Score 3 for the remaining items. The formulaic sequence was evaluated as a unit. That means, if a deliberate blur was produced intelligently (as opposed to ‘parroting’), all the items within this envelope were assigned a ‘4’, other items were judged word by word and received a maximum score of ‘3’. In contrast to the first assessment version, the refined version (see Appendix 5) is more accurate and legitimate, and this is also used as the evaluation methodology for Test 2 (detailed in Section 8.2.3.2).

The evaluation for Test 1, based on the revised version, was done by the present author. A panel was engaged to further validate the methodology in Test 2 (see Section 8.2.3.4).

8.2.1.4 ANALYSIS OF TEST 1

Interpretation of test data is an indispensable procedure for a test, and how to achieve meaningful and appropriate inferences based on test scores is even more important for

validity. In the present study, the validity of Test 1 is demonstrated on the basis of the analysis of both test samples and test results.

A. Analysis of sample snippets

Given that all of the participants are L1 speakers of Chinese, and that most of time they are exposed to a non English-speaking environment and learn English in a typically Chinese manner, these authentic English snippets (as shown in Table 16), especially the Irish accents, are very difficult for them and pose problems in perceiving and imitating the sounds. Based on the average scores, various results emerged for these snippets, as shown in Figure 16. The percentage scores shown are derived from the average student performance per snippet.

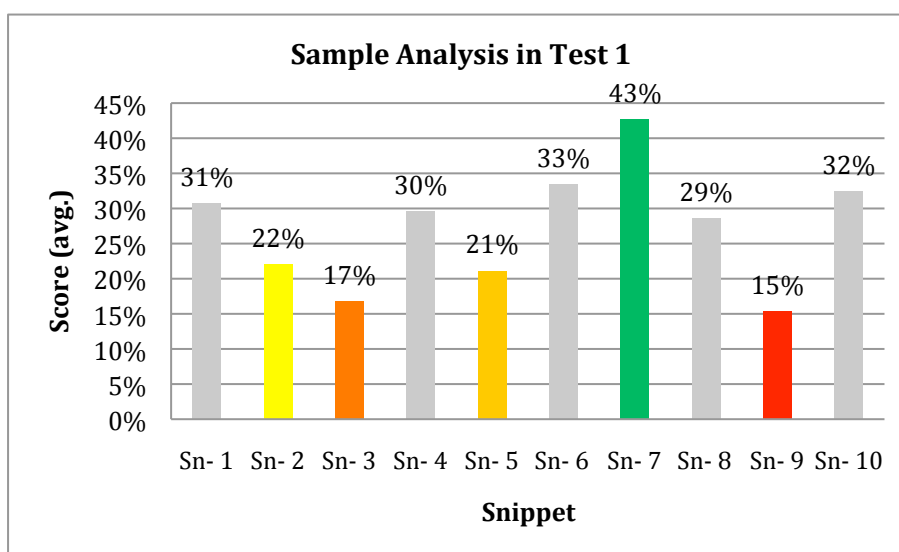


Figure 16: Sample analysis in Test 1

Contrary to expectation, Snippet 7 ('You are scraping the bottom of the barrel sir.') is the best perceived and produced of these samples. In addition to the salient linguistic feature – reduced form of the non-lexical word 'of', Snippet 7 is an idiom which is not familiar to most of the Chinese learners, spoken at a very fast speed and including unstressed elements as well; so this snippet was anticipated to be one of the most

difficult for the participants. However, it proved to be much easier. The best student gets 74% correct and only 6% of the students appear unable to produce a single phoneme, which gives an average score of 43%. The main reason for this may be that the stressed elements in this snippet are easily captured by students. For example, 93% of the participants can perceive and produce the word ‘you’, 78% of them get the word ‘are’ correct and 87% of them can produce the word ‘scraping’ at different levels. Another reason may be that, although the students have no knowledge of the formulaic sequence ‘scrape the bottom of the barrel’, however, they are so familiar with the basic structure of the idiom – ‘the ... of the ...’. Therefore, Snippet 7 scores much better than expected. In contrast, Snippet 2 (‘From pillar to post.’) was anticipated to be the easiest one. However, it proves to be one of the most difficult. Although it is only a four-item snippet and delivered at a really slow speed, it is a pure formulaic sequence without other novel elements which can be predicted, thus, it is not easy to be perceived and acquired by Chinese learners. With a score of only 22% correct responses, this snippet is positioned fourth from the bottom.

Of all these samples, Snippet 9 (‘Well that is exactly what the Italians would have been doing.’) returned the worst performance with an average score of only 15%. The length of eleven items in combination with a really fast delivery speed of 788 syll/min hinders imitation and production by students. The best participant, who gets 39% correct, can only capture the phonemes at the beginning and at the end of the snippet, due to the position effects (Field, 2004, pp.175-76, p.216). In contrast with Snippet 9, Snippet 10 (‘I cannot walk down the street without having to walk on the road.’) did much better than expected. Even though there are three items more in Snippet 10, however, since there are some high frequency words, e.g., ‘walk’, ‘street’, ‘road’, the collocations ‘down the street’, ‘on the road’, and the parallel structure(s) – ‘walk down the street’ and

‘walk on the road’, all these elements make the sequence easier to capture and therefore the participants achieved relatively higher scores than in Snippet 9. Besides Snippet 9, Snippet 3 (‘Certainly made up for that.’) was also done badly. One of the reasons may be the interesting stress intonation patterns which raise the problem for participants in ‘segmenting’ the signals. Another reason may be the formulaic sequence – ‘make up for ...’ which is not easily ‘decoded’ by students either. The same applies to Snippet 5 (‘Do you know what I mean?’). It is formulaic as well, delivered with a fast speed and including the hugely reduced and weak forms ‘do you know’ and ‘what I mean’. All these important linguistic features prevent the participants from arriving at a correct imitation and production.

The other samples, Snippets 1, 4, 6 and 8 are performed relatively well as anticipated. One of the biggest problems in these samples is reduction and weak forms. In Snippet 1 (‘Which one are you thinking of?’), the syllable ‘one’ cannot be perceived by 98% of the participants, and the syllable ‘are’ cannot be perceived by 94% of the students. In Snippet 4 (‘I moved out of home when I was eighteen.’), 89% of the students cannot capture the weak form ‘I’ in the grammatical paradigm formulaic sequence ‘I was’. The non-lexical word ‘the’ in Snippet 6 (‘Fifty sixty percent of the people.’) is neglected by 100% of these participants. The second problem is the formulaicity. Prepositions in collocations are ignored by almost all the students. For example, 10% of the students get the syllable ‘of’ in the collocation ‘move out of’ in Snippet 4 and in ‘a list of’ in Snippet 8 (‘You have a list of questions?’), and only 1% of them succeed in perceiving it in the collocation ‘percent of’ in Snippet 6. This is in agreement with the findings of Spöttl and McCarthy (2004) who notice a bias towards ‘heavy’ lexical items in processing strings. Thus, in Test 1, the important linguistic features are presented to the

participants, and the basic linguistic ability of the participants in imitating and reproducing the test snippets is recorded.

B. Analysis of results within the three groups

There are one hundred participants in Test 1. Most of the participants were very positive about the test. They very much concentrate on listening to the samples and try to imitate them as correctly as possible. Although they are from three different educational levels, an equivalent result emerges, as laid out in Figure 17.

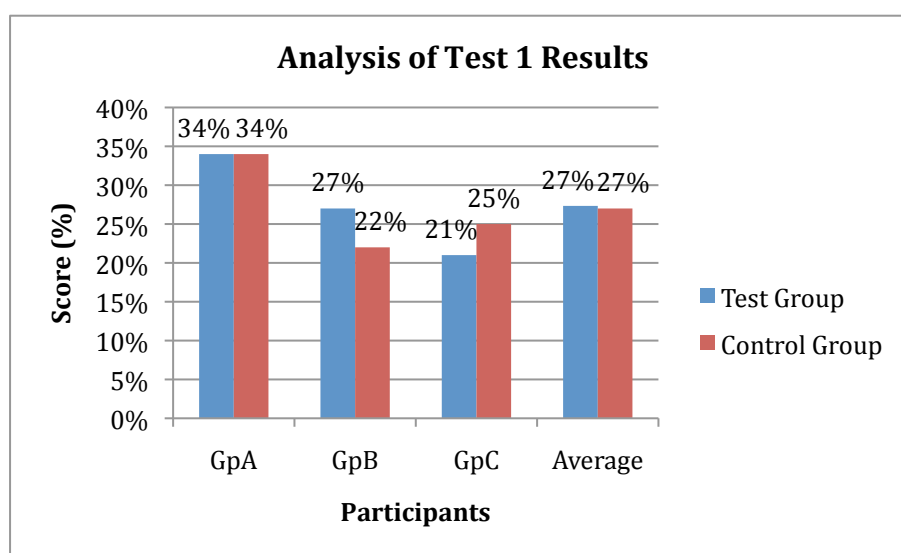


Figure 17: Analysis of Test 1 results

From the test results, it can be seen that there is no difference between the test group and the control group in GpA (the third level college). However, for GpB (public junior secondary school), the overall result of participants in the test group is slightly higher than those in the control group. The result between test group and control group in GpC (private junior secondary school) shows a slight imbalance as well, with a 4% average higher score in the control group. Even though there are some discrepancies between test group and control group in each of the groups GpA, GpB and GpC, the equivalent average result between Test Group and Control Group overall shows that there is no gap

between these two groups in this benchmark test. Both Test Group and Control Group are homogenous and have approximately the same proficiency in English, which verifies the validity of Test 1 and provides a reasonable baseline for Test 2.

In addition to the overall balanced level between Test Group and Control Group, Test 1 also seems to indicate a direct relationship between the results participants get and the numbers of years they have been learning English and the amount of their exposure to English listening and speaking practice. Based on Questionnaire 1 (see Section 8.2.1.2), of the six groups, GpAc, GpBt and GpCc have much more exposure than the other groups at the same level, and meanwhile the participants in these groups also perform a little better than the other group, except for the identical result between GpAc and GpAt. A comparison between these factors, as shown below in Table 18, seems to indicate a linear relationship between higher scores and students' exposure to spoken English. This demonstrates, as anticipated, the significance of exposure in the learning and acquisition of target language.

Group	Age (avg.)	No. Yrs Learning English (avg.)	No. Hrs English class /wk (avg.)	No. Mins listening to English /wk (avg.)	No. Mins speaking English /wk (avg.)	Score (avg.)
GpAc	20	9	12	230	267	34%
GpAt	20	8	12	200	260	34%
GpBt	15	6	4.5	109	76	27%
GpBc	15	5	4.5	79	60	22%
GpCc	13	6	4.5	127	127	25%
GpCt	12	5	4.5	106	106	21%

Table 18: Analysis of correlation between score and amount of English exposure

8.2.1.5 OVERVIEW OF TEST 1 AND SUGGESTIONS FOR TRAINING AND TEST 2

A. Overview of Test 1

In Test 1, ten snippets, taken from natural authentic L1-L1 English speech, were used as elicitation of test data. These ten samples are very rich in demonstrating the significant characteristics in real, native, spontaneous speech.

Subjects in Test 1 were one hundred Chinese students from three different schools, and comprising the Test Group and the Control Group. Questionnaire 1 was implemented before the test for both Test Group and Control Group. The feedback from the questionnaires made clear that the participants had various exposures to native spoken English. This provided a representative sample for the analysis of Test 1 and the design for Test 2.

The aim of Test 1 is to evaluate the intelligibility and native-like speech re-production, rather than listening for comprehension. Therefore, as long as the participants can perceive the sounds and re-produce them correctly, it shows that they achieve intelligibility. The evaluation was based on the methodology of listening to each student individually and manually scoring their oral re-production of the snippets played to them. A 5-point scoring system of 0-4 was applied to the participants' responses. All the data in Test 1 was evaluated by the present author using the same methodology.

The evaluation of Test 1 was done in terms of sample analysis and results analysis. While the ten samples turned up different results, all of them reflected the identified problems for Chinese EFL learners. Two samples were too long to accommodate the

ability of short term memory. As for the test results, two points emerged. Firstly, the analysis of the six groups showed that there was no gap between Test Group and Control Group in this benchmark test, which provided a reasonable baseline for Test 2. Secondly, Test 1 also seemed to indicate a direct relationship between the learning result participants achieve and the number of years participants had been learning English, and the amount of their exposure to English listening and speaking practice.

In order to ensure the suitability of the test theory and design, the present author, along with her supervisor Dermot Campbell, consulted a speech and language therapist – Dr. Paula Bradley (Personal communication, 5 March 2008). Positive feedback was given by her, which further confirmed the methodology of Test 1 and the feasibility of Test 2.

B. Planning for training sessions and Test 2

From Test 1, some suggestions were gained for the better implementation of Test 2. Firstly, training materials were needed in order to give the participants enough exposure to the identified linguistic characteristics, so as to achieve an improvement in their linguistic competence in Test 2. It was therefore decided that every month, ten samples of authentic English snippets would be delivered by the present author to the participants (via three Chinese teachers of English). Every sample was accompanied by three versions of spoken text: the snippet, the speaker sequence in which the snippet occurred and the interactive dialogue in which the speech occurred. In addition, the orthographic text, a description of the topic covered and an explanation of relevant, important linguistic characteristics were also supplied. Audio files of these samples were to be supplied at normal (100%) speed for the Control Group, and with both 100%

and 40% speeds for the Test Group. The training session was planned to be carried out by three Chinese teachers from September 2007 until Test 2 which was to take place in May 2008. Since the present author could only support and monitor the training session from a distance, a questionnaire for and feedback from the Chinese teachers and students was needed to identify the problems and issues so as to improve the effectiveness of subsequent training session. These are discussed in detail in Section 8.2.2 below.

Secondly, in view of the extensive work involved in the evaluation of the test data in Test 1 (i.e., 10 snippets for each of the 100 participants. Each snippet has at least four items and contains several linguistic features, which adds up to a lot of detail.), it seemed advisable to reduce the scope of Test 2. In order to retain maximum validity, it was deemed advisable to re-test as many students as possible and to stay within the time framework of the PhD by reducing the number of test samples. Therefore, the same number of students would be re-tested, and the overall number of samples would be reduced from ten to six. Two of the samples in Test 1 – Samples 2 and 8 – were particularly good (as discussed in Section 8.2.1.4) and would be retained for re-use in Test 2 to evaluate any improvement. All the samples would be chosen based on the same linguistic features as in Test 1, in particular, all the six samples were to contain formulaic sequences.

The same methodology for testing and evaluation would be employed in Test 2. All the participants were to be exposed to and be tested with real, informal English speech at normal speed. The 5-point evaluation system was applied to the participant's response, but Score 4 would be highlighted to evaluate any improvement.

In Test 2, it was anticipated that both groups would perform better than in Test 1, since they would both have had more exposure to native English speech than before (due to the effect of the training sessions with the training materials). The Test Group was also expected to achieve a higher level of intelligibility than the Control Group due to the effect of the slow-down technology. Furthermore, questionnaires would be devised to ascertain if both groups reacted positively to the training session, and if the Control Group had improved because of the increase in exposure, and whether the Test Group benefits were achieved due to the availability of slowed-down speech.

8.2.2 Training for Test Participants

8.2.2.1 PURPOSE OF TRAINING

Test 1 was designed to ascertain what difficulties Chinese learners of English had when presented with native-to-native English speech. The results showed no difference between the Test Group and the Control Group. The test also revealed that both groups had a very low ability in coping with informal English NS. In order to give the participants enough exposure to selected linguistic characteristics so as to achieve an improvement in Test 2, training materials were designed from September 2007 to April 2008, and training sessions were implemented in the six month period from September 2007 to April 2008, excluding the January 2008 and February 2008 Winter Holiday for Chinese schools.

8.2.2.2 EXPLANATION BEFORE TRAINING

Before the training sessions started, a written explanation of the training purpose and the training materials (see Appendix 6 for Control Group and Appendix 7 for Test Group) was sent to the Chinese teachers who conducted the training activities on behalf of and under the direction of the present author. Firstly, Chinese English speakers'

problems when they are involved in English communication were mentioned. Then there was a discussion on what Chinese English learners need and why they need to be exposed to real, authentic, native English speech. Finally, for the Test Group, the potential advantage of using slow-down technology in helping Chinese learners of English to capture the important segmental and supra-segmental levels of linguistic features so as to be comfortable in real target language community was emphasised.

8.2.2.3 TRAINING MATERIALS

All the training samples were taken from authentic, dynamic English dialogues in the DSC. The samples were chosen based on the same linguistic features as in Test 1 (detailed in Section 8.2.1.1). There were in total 60 samples provided within the six months' training sessions. There was often more than one linguistic feature in each sample. Therefore there was a total of 214 linguistic characteristics demonstrated within 60 samples, as shown in Table 19. Among them there were 103 reduced and weak forms, 72 formulaic sequences, 32 rich intonation patterns displayed and 7 specific phonemes.

Linguistic Feature	No. of Features
reduced and weak form	103
formulaic sequence	72
rich intonation pattern	32
specific phoneme	7
total	214

Table 19: Linguistic features in training materials

Every month, ten samples were provided to the participants. Every sample was accompanied by three versions of the spoken text: the snippets, the speaker sequence in which the snippets occurred, and the interactive dialogue in which the speech occurred. In addition, the orthographic text, a description of the topic, and an explanation of

relevant, important linguistic characteristics were supplied (see Appendix 8 to Appendix 19). Along with the document files, audio files of these samples were supplied as well for both the Test Group and the Control Group. Three Chinese English teachers (the same teacher taught both test group and control group at the same level) could download these audio files from the DMC website. Both 100% and 40% speeds were made available to the Test Group, and 100% speed only to the Control Group.

Most of the training activities were done in class, controlled by the teachers. After class, the Control Group could access the training materials freely. For the Test Group, however, the teachers controlled the 40% speed audio files – the participants could only access 100% speed of audio files – to ensure that no 40% speed audio files were available to the Control Group.

8.2.2.4 MONITORING THE TRAINING PROCESS

Since the training was mainly implemented in China, in order to monitor the training sessions, regular telephone conversations (every two or three weeks) were carried out between the present author and the three Chinese teachers. Apart from the phone call communications, Questionnaire 2 (see Appendix 20 for Control Group and Appendix 21 for Test Group) was also sent to three different Chinese teachers at different stages for monitoring the training process and obtaining feedback. Suggestions were given to them from the pedagogical perspective. Meanwhile, some useful feedback was sent by the teachers (details shown in Appendix 22 to Appendix 25), which led to improvements in the training materials. Details are discussed below.

Questionnaire 2 (Feedback on training from Chinese teachers of English)

Feedback was given by three Chinese English teachers on training conditions, training process and training materials.

A. Training conditions

First of all, the basic training conditions were ascertained by means of the questionnaires. For example, Question 2 was about how long the students were exposed to the training materials per week. The participants in GpA (third level students, both test and control) were exposed to the training materials for about 20 minutes a week. GpC (beginner junior secondary) spent 45 minutes a week. GpB (advanced junior secondary) was exposed to the training materials up to one hour a week. Questions 3-5 were about where the students undertook the training activities, what equipment the teacher used and if the students wore headphones or not. These questions showed the differences in training conditions between the three groups. GpA accessed the training process in the language lab, the teacher used a PC and the students wore headphones, which ensured that all the participants could access the same level of good audio quality. For GpC, the students did the training practice in the classroom, where there was a PC, but no headphones. Since GpB was a public secondary school, there was no PC available in the classroom. Thus, the teacher used a tape recorder instead. Since the audio files were delivered via the Internet, the teacher had to transfer the CD format into a cassette tape. In order to make sure there was no change in the audio quality, samples of the cassette tape were sent back by the teacher to the present author.

B. Training process

Secondly, some suggestions were given to the teachers from a pedagogical perspective. Questions 6-20 were about how the teachers presented the training materials to the

students, and what the procedures were. By means of questionnaires for the Chinese teachers and also by means of telephone conversations (made by the present author), it was shown that every teacher had her own methodology in training students' listening and speaking practice, based on the students' ability. For example, Question 9 was about the order in which the students listened to the recorded materials. The three teachers applied slightly different methodologies, but generally speaking, there was no statistically significant difference between them. The teachers were free to decide which approach best suited the students. However, from the pedagogical point of view, some suggestions were recommended to the teacher in GpC. A procedure of beginning with snippets, then sequences, and finally moving to the wider context – i.e., dialogues – was recommended, and also for the order in which the students were exposed to the different speeds: starting with the normal speed first, then on to slow speed, then finally back to the normal speed again.

C. Training materials

Meanwhile, some useful feedback was sent by the teachers, which led to improvements in the training materials. Question 21 was about which specific phonetics the students found difficult. Weak forms and reduced forms, and some particular phonemes, e.g., /v/, /n/, /ŋ/, /θ/, /ð/ and /ʒ/, were suggested by the teachers, which conformed to the linguistic features targeted in Test 1.

In addition to the phonetic aspect, the teachers also gave some helpful feedback on Questions 22-26. All of the samples used for training were from the dynamic dialogues recorded in the DMC, which were mainly free conversations between adult speakers. The first feedback from GpAt (the test group of the third level cohort) (see Appendix

22), which was sent in October 2007, gave the comments that some of the training materials were not suitable for the students, and students would prefer materials which were closer to their studies and their life experiences. The first change was made to the materials in November 2007 (see Appendix 12 and Appendix 13), in which 3 out of 10 samples occurred in the environment of two speakers playing video games, which was a popular topic among the students.

After the first modification in materials, the questionnaires Feedback 2 (see Appendix 23) and Feedback 3 (Appendix 24) were given to GpBc (the control group of the advanced junior secondary cohort) and GpBt (the test group of the advanced junior secondary cohort) in November 2007. Feedback 2 showed that only some of the students were interested in the topics. Comments from Feedback 3 stated that the topics in the materials were too removed from what the students learned in class, and some suggestions were given that the students would like to listen to topics, such as going shopping, making a phone call, booking a ticket, seeing a doctor and asking the way. Therefore, the second change was in the materials of December 2007 (see Appendix 14 and Appendix 15), March 2008 (see Appendix 16 and Appendix 17) and April 2008 (see Appendix 18 and Appendix 19). 3 out of each set of 10 samples were in the context of doing shopping, making a call to a travel agency to book flights and seeing a doctor. Based on the contents in course-books provided by the Chinese teacher (see Appendix 26), new recordings were made by the linguistic researchers in the DMC. There were two levels in the new recordings. The first level was reading-out the dialogues taken from the students' course-book, but with more natural, native-like English production. The second level was an unscripted, interactive conversation based on the same topic as Level 1 so as to expose students to more real, dynamic native-to-native English speech.

Then, Feedback 4 (see Appendix 25) was sent to GpCt (the test group of the beginner junior secondary cohort) in December 2007, in which there were no negative comments on the training materials. The improvement in training materials ensured that the training proceeded efficiently.

8.2.2.5 STUDENT REACTION TO TRAINING SESSION

After the training was done, Questionnaires 3 and 4 were given in Chinese to the students to gain feedback about the training process. In practice this was actually done on the same day as Test 2 was delivered.

The questionnaires were designed in line with the principles of questionnaire theory (Dörnyei, 2003). All the variables were made up of multiple items, which are on a continuum of five values from very negative to very positive, including neutral (see Appendix 27 and Appendix 28 for English version). The results showed a positive attitude to the training activity and the slow-down facility.

A. Questionnaire 3 (Test Group)

Q1: Do you like listening to the training materials?

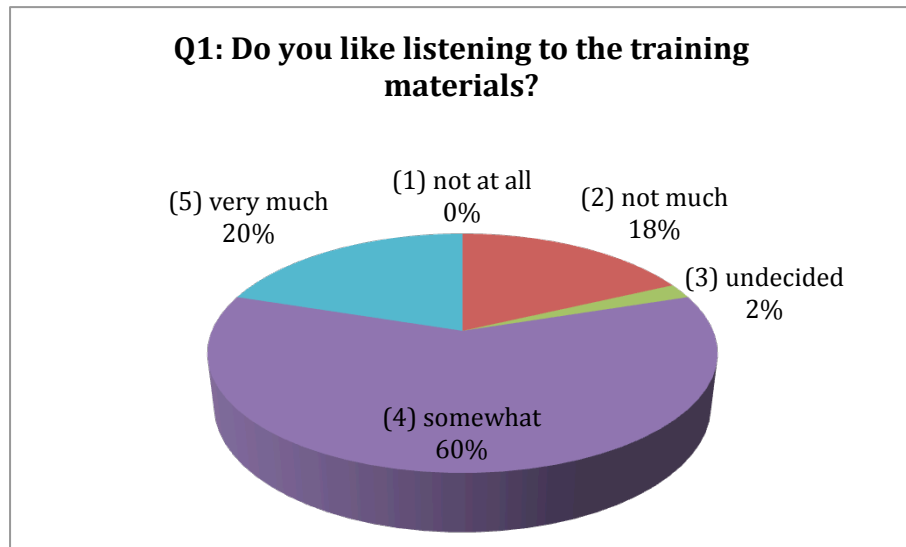


Figure 18: Q1 in Questionnaire 3 (Test Group)

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
GpAt (15)	0	5	1	8	1
GpBt (20)	0	4	0	10	6
GpCt (15)	0	0	0	12	3
Total (50)	0	9	1	30	10

Table 20: Q1 in Questionnaire 3 (Test Group)

From Q1 as shown in Figure 18, it can be seen that 80% of the participants reported that they liked listening to the training materials, and 20% of them were extremely positive. Only 18% of the participants were a little negative. In general, students liked to be exposed to real native English speech, especially the participants in GpBt (see Table 20).

Q2: Can you understand the materials when listening at normal speed?

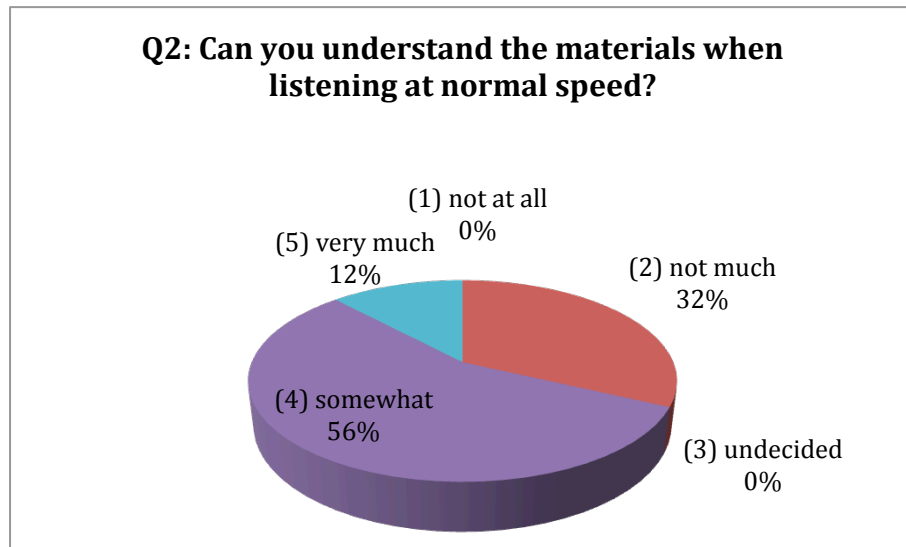


Figure 19: Q2 in Questionnaire 3 (Test Group)

In Q2 (see Figure 19), 32% of the participants thought that they could not understand the materials so well when played at normal speed. 56% of them said that they could only understand some of the materials. It was indicated that exposure to normal speed alone could not really help the listeners capture all the information they need in the rapid flow of connected speech.

Q3: Does the slow-down help you hear what was said?

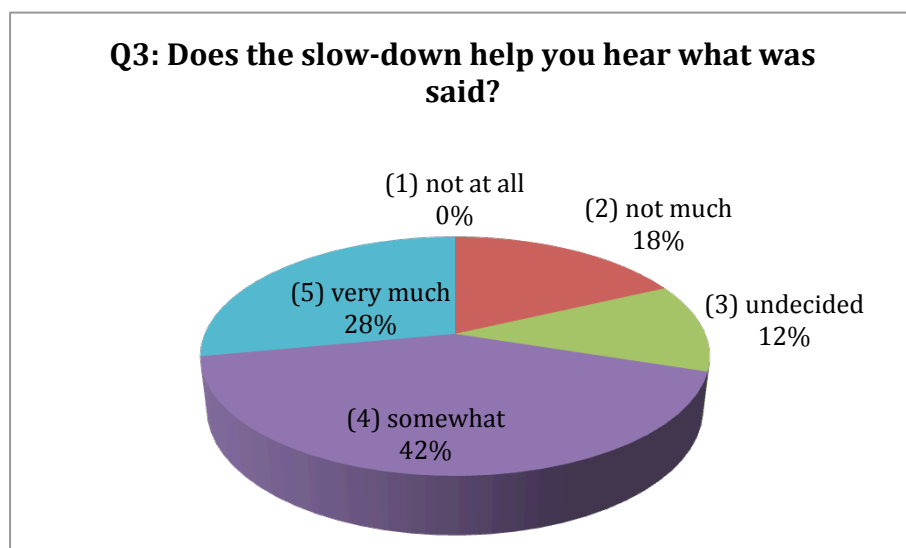


Figure 20: Q3 in Questionnaire 3 (Test Group)

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
GpAt (15)	0	3	2	7	3
GpBt (20)	0	1	2	11	6
GpCt (15)	0	5	2	3	5
Total (50)	0	9	6	21	14

Table 21: Q3 in Questionnaire 3 (Test Group)

In Q3 as shown in Figure 20, it can be seen that, accessing the slowed speech, 70% of the participants stated that the slow-down really helped them hear what was said in the training snippets. Of the three groups, GpBt was the happiest with the slow-down facility, and 85% of this group were positive (see Table 21).

Q4: Did the training period help you understand L1 speakers better?

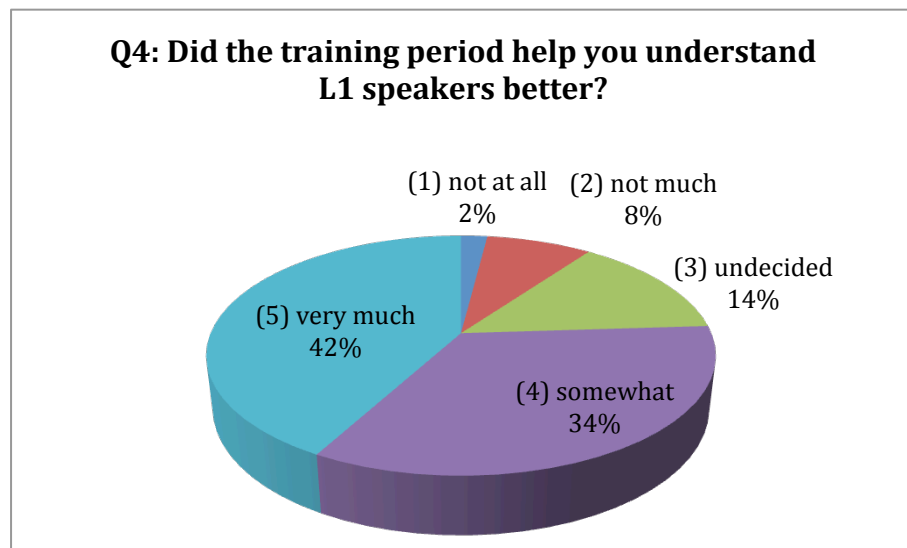


Figure 21: Q4 in Questionnaire 3 (Test Group)

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
GpAt (15)	0	1	4	8	2

GpBt (20)	1	3	2	5	9
GpCt (15)	0	0	1	4	10
Total (50)	1	4	7	17	21

Table 22: Q4 in Questionnaire 3 (Test Group)

The advantage of exposure to real native English speech is reflected in Q4 (see Figure 21). 76% of the participants thought that their ability to cope with native English speech had been improved, especially GpCt with 93% positive (as shown in Table 22).

Q5: If the answer to Q4 is '(4)' or '(5)', was the improvement due to the slow-down facility?

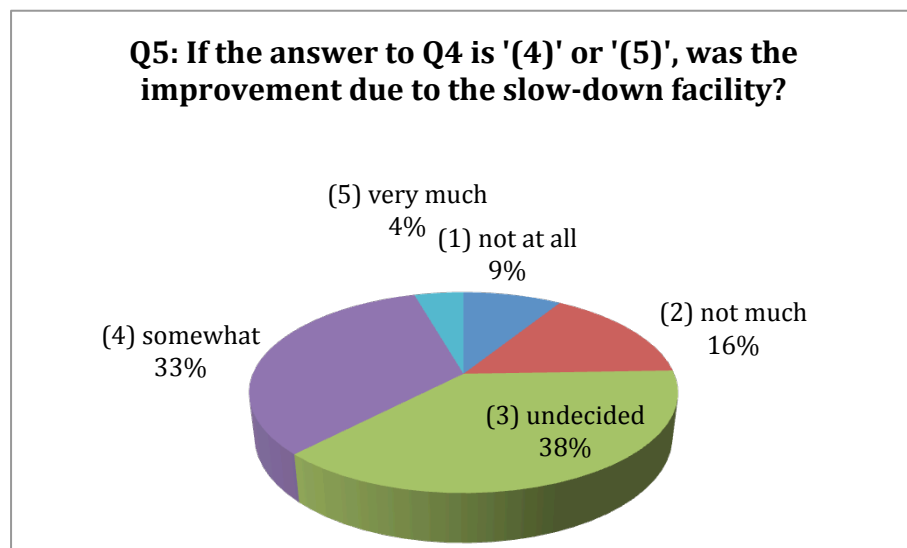


Figure 22: Q5 in Questionnaire 3 (Test Group)

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
GpAt (15)	0	1	3	7	1
GpBt (20)	3	1	9	5	0
GpCt (15)	1	5	5	3	1
Total (50)	4	7	17	15	2

Table 23: Q5 in Questionnaire 3 (Test Group)

In Q5 as shown in Figure 22, the question as to whether the slow-down contributed to the participants' improvement was posed. 37% of the participants gave a positive response. Only 25% of the participants were negative. 38% of them were not sure. This would seem to indicate that students (particularly in GpBt and GpCt, see Table 23), while acknowledging there was an improvement due to the training period, were unclear as to whether the slow-down was the reason for this improvement.

Q6: Other comments:

As shown in Appendix 27, six choices were given to the participants in order to facilitate them in giving their comments on slow-down facility. More than one possibility could be chosen.

The slow-down facility:

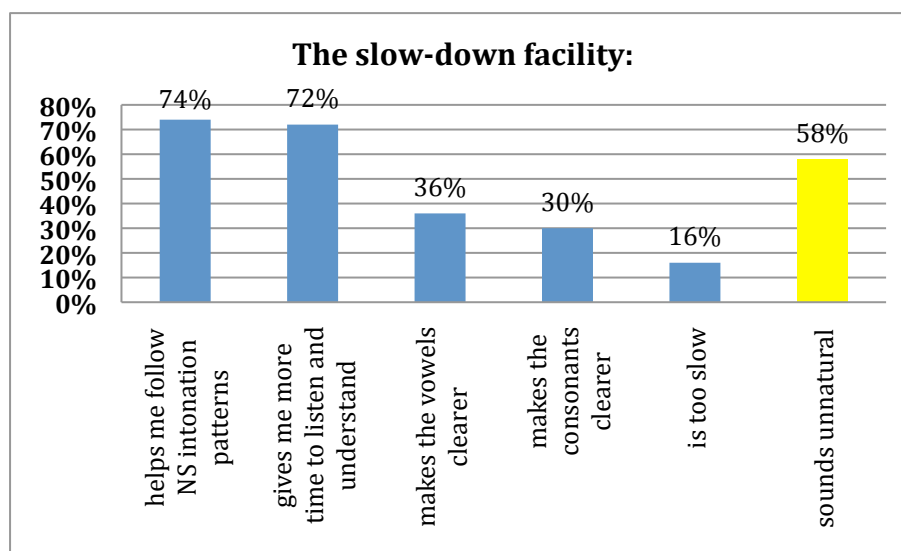


Figure 23: Q6 in Questionnaire 3 (Test Group)

Comments on the slow-down facility (see Figure 23) showed that most of the participants in the Test Group were positive. After using the slow-down facility in their training period, 74% of the participants thought the slow-down gave them the chance to

follow native intonation patterns in English, and 72% of them stated that the slow-down could allow them to capture more linguistic information and help them understand native English speech. There were respectively 36% and 30% of the participants who thought that the slow-down could make either vowels or consonants clearer. Only 16% of the participants thought that the speed used (40%) was too slow. This meant that most of them liked the slowed speed which is the slowest, practicable (and acceptable) speed for language learners to improve their ability to perceive accurately. There were also 58% of the participants who expressed the opinion that the slow-down speed sounded unnatural. Since the slowed speed was two and a half times slower than the normal speed, it indeed sounded unnatural. However, informal tests by Dermot Campbell (Personal communication, 23 April 2005) indicated that at 40% speed, the listeners' attention was drawn more to the prosody of the utterances, which was what they were expected to listen to, rather than the semantic content. Nevertheless, the respondents were very positive about the slow-down, despite it sounding 'unnatural'.

B. Questionnaire 4 (Control Group)

Q1: Do you like listening to the training materials? (same as Q1 in Questionnaire 3 for Test Group)

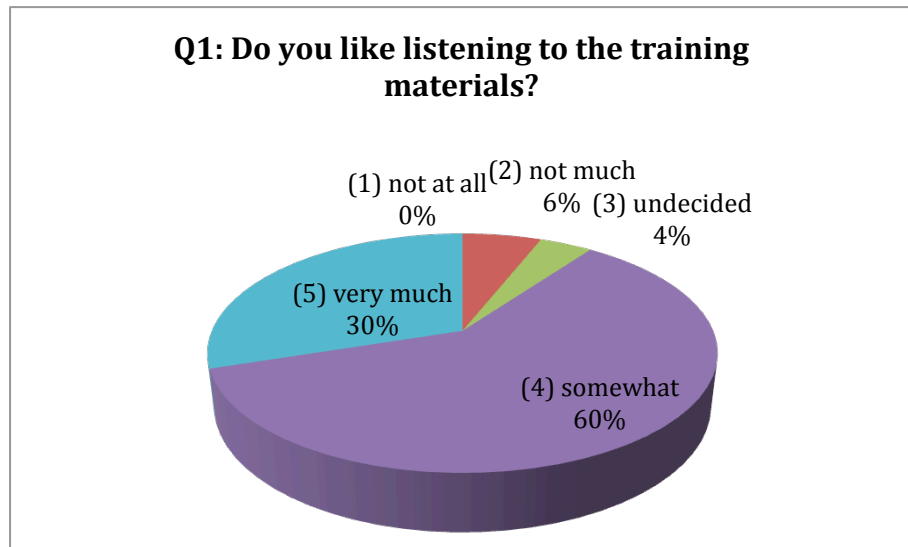


Figure 24: Q1 in Questionnaire 4 (Control Group)

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
GpAc (15)	0	2	1	8	4
GpBc (20)	0	1	1	10	8
GpCc (15)	0	0	0	12	3
Total (50)	0	3	2	30	15

Table 24: Q1 in Questionnaire 4 (Control Group)

A very positive attitude can be seen in response to Q1, as shown in Figure 24 and Table 24. 90% of the participants in the Control Group (and particularly 100% in GpCc) stated that they liked listening to and being exposed to real, natural native English. Only 6% of them were undecided.

Q2: Did the training period help you understand L1 speakers better? (same as Q4 in Questionnaire 3 for Test Group)

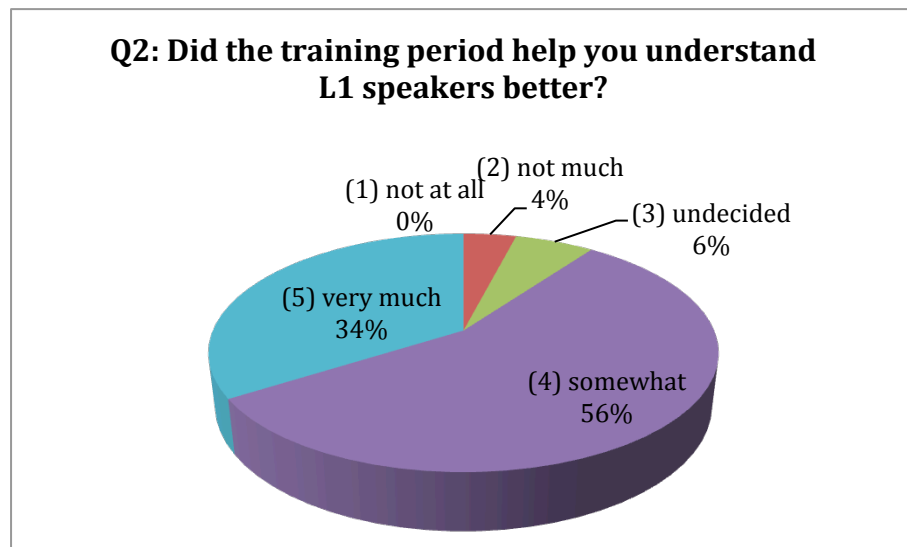


Figure 25: Q2 in Questionnaire 4 (Control Group)

To Q2 as shown in Figure 25, there was also a very positive response. 90% of the Chinese students thought that they benefited from the training activity, which helped them understand real native English speech better and improved intelligibility rates. 34% of them were very positive, 6% of them were not sure, and only 4% were a little negative.

Q3: Other comments:

After Q1 and Q2, the participants in the Control Group were invited to give their comments. In contrast to Q6 in Questionnaire 3 for the Test Group, Q3 in this questionnaire (see Appendix 28) was open-ended, so as to ensure that responses emerged from participants' own perspectives. 32 out of 50 gave their suggestions on the training as follows:

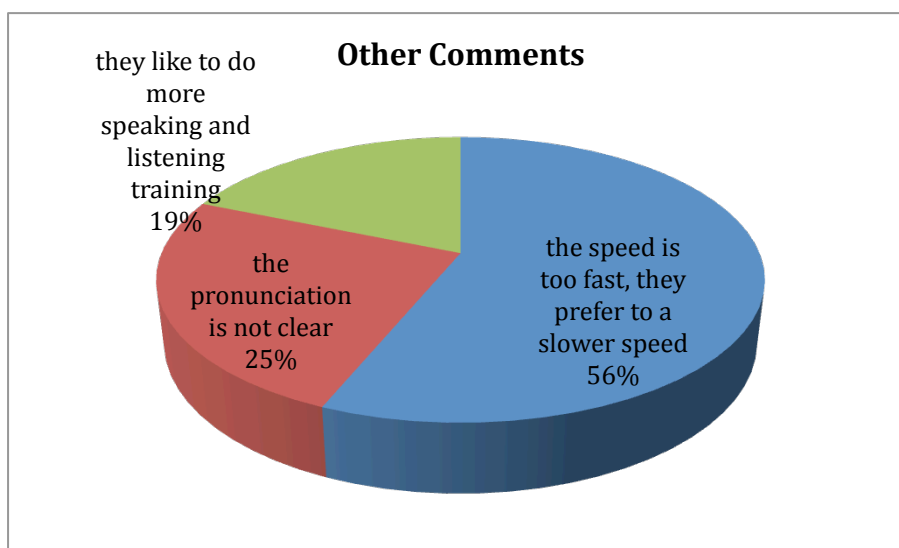


Figure 26: Q3 in Questionnaire 4 (Control Group)

From Q3 (see Figure 26), it can be seen that 19% of the participants expressed that they would have liked to do more training. 56% of the participants thought that the 100% speed was too fast for them, and they preferred a slower speed. 25% of them commented that the pronunciation was not clear enough to capture what was said. Compared with the comments given by the Test Group in response to Questionnaire 3, the problems of speed of delivery and pronunciation recognition could be compensated by the advantage of the slow-down facility.

8.2.3 Test 2

After the students had spent six months training in listening to real natural interactive English speech, Test 2 was carried out in China, in May 2008, by the present author to ascertain whether a change had occurred in the participants' ability to imitate and reproduce native-like, authentic English speech. The detailed procedures for the design and implementation of the test, and evaluation of the test data are described as follows.

8.2.3.1 DESIGN OF TEST 2

A. Design principles and requirements

All the training activities had finished by the end of April, 2008. By the time Test 2 started on May 16, 2008, any direct memory of the training materials was minimised (Hulstijn, 2003). The same principles as in Test 1 were employed in Test 2. In order to reduce the influence of unfamiliar words, only 3% of the items in Test 2 samples were new words for thirty (30%) of the participants, which could better reflect participants' actual language ability.

B. Choice of test snippets

There were ten samples in Test 1 and evaluating one thousand snippets was a very time-consuming task; thus, in Test 2, the overall number of samples was reduced to six to facilitate the evaluation, rather than reduce the number of participants. Two of the samples – Samples 2 and 8 in Test 1 – were re-used in Test 2 to help evaluate any improvement. The other samples were different from both Test 1 and the training materials. All the samples were chosen based on the same linguistic features as in Test 1. The detailed characteristics are laid out in Table 25. As in Test 1, the items highlighted in red and bolded indicate formulaic sequences exhibiting natural blur and meriting a score of '4' in the evaluation system (see details in Section 8.2.1.3).

Sample	Orthographic Text	Linguistic Features	Speed (syll/min)	Pitch Range (Hz)	Length (no. of items)
1	It is bananas.	1. slightly reduced 'it is' to /its/, 2. grammatical paradigm formulaic sequence 'it is', with the speed of 745 syll/min and pitch range of 11 Hz, 3. rich intonation pattern on	426	175	3

		<p>‘bananas’, with speed of 335 syll/min and pitch range of 96 Hz,</p> <p>4. idiom with lexico-grammatical problems,</p> <p>5. fast speed and wide pitch range</p>			
2	<p>From pillar to post.</p>	<p>1. idiom with lively intonation pattern,</p> <p>2. slow speed and wide pitch range</p>	241	113	4
3	<p>When did you move out of home?</p>	<p>1. emphasis on the questioning word ‘when’,</p> <p>2. slightly reduced ‘did you’ to /di dje/,</p> <p>3. grammatical paradigm formulaic sequence ‘did you’, with the speed of 632 syll/min and pitch range of 79 Hz,</p> <p>4. weak forms on ‘out of’, produced as /əudə/,</p> <p>5. collocation ‘move out of’, with the speed of 541 syll/min and pitch range of 194 Hz</p>	519	196	7
4	<p>You have a list of questions?</p>	<p>1. reduced ‘you have a’, produced as /juvə/,</p> <p>2. huge reduction on the non-lexical word ‘of’,</p> <p>3. collocation ‘a list of’, with the speed of 732 syll/min and pitch range of 16 Hz,</p> <p>4. question intonation pattern marked with a rising tone at the end, rather than a</p>	565	57	6

		grammatical structure			
5	He had no idea what it meant.	1. emphasis on the negative word 'no', 2. unstressed form of 'idea', 3. collocation 'had no idea', with the speed of 395 syll/min and pitch range of 29 Hz, 4. reduced 'what it', produced as /wədə/, 5. chunk 'what it meant', with the speed of 393 syll/min and pitch range of 61 Hz	401	48	7
6	Which makes it, erm, you know , quite difficult.	1. chunk 'you know', followed by a pause, used as a word-filler, with the speed of 330 syll/min and pitch range of 49 Hz	275	98	8

Table 25: Testing samples in Test 2

C. Justification for samples chosen

First of all, all the six samples consisted of different types or sub-types of formulaic sequences as in Test 1. There were three collocations in each of Samples 3, 4 and 5. Both Samples 1 and 2 were idioms. Two grammatical paradigms with a fast speed were delivered in Samples 1 and 3. There were also two chunks embedded in Samples 5 and 6. Three collocations were relatively easy to learn by the participants either in class or by self-study. However, the other types of formulaic language, especially the chunks, could only be acquired via their phonological realisations within the context in which they occurred, as discussed in Section 7.2.3.

Secondly, a very rich intonation expression with a wide pitch range was another significant feature in these samples. Both Samples 1 and 3 had a huge pitch range, with the highest at 196 Hz and a very fast speed of delivery at over 400 syll/min. Sample 2 also had a wide pitch range, but with a slower delivery speed of below 300 syll/min. Sample 4 was a question intonation pattern marked with a rising tone at the end, rather than a grammatical structure. Sample 6, overall, had a wide pitch range of 98 Hz, but contained a really flat chunk with a range of only 49 Hz.

Thirdly, as with Test 1, and in response to feedback from Questionnaire 2 by the Chinese teachers, reduced forms and weak forms were included in Test 2. For example, a huge reduction of the non-lexical word ‘of’ and ‘you have a’ speeded up the delivery of Sample 4 up to 565 syll/min. In Sample 3, ‘did you’ was slightly reduced, and weak form ‘out of’ occurred. An unstressed form of ‘idea’ and a reduced form of ‘what it’ also occurred in Sample 5.

The last feature in Test 2 was on the production of the phoneme /v/, which is absent in Chinese Pinyin and often causes problems for Chinese learners of English. Samples 3 and 4 displayed the /v/ sound for evaluating the improvement.

D. Anticipated difficulty of test samples

Comparing all the linguistic characteristics in each of the six samples (see Table 25 above), it was anticipated that Sample 1 would be the easiest for the participants. Although Sample 1 was delivered at a fast speed of 426 syll/min, there were only three items in the snippet, which was suitable for processing in working memory. Furthermore, the three elements were also high frequency items. In addition to Sample 1,

it was anticipated that Sample 2 would be the second easiest one to produce. Sample 2, like Sample 1, was a short snippet with only four items. The whole snippet was also delivered at a really slow speed of 241 syll/min. Sample 6, however, was expected to be the most difficult. One reason was that, the sample was a longer sequence with eight items. More importantly, there was a chunk embedded in the snippet, which was at a fast speed and flat pitch range. Based on the present author's teaching experience, it was considered that Chinese EFL learners would have difficulty in coping with different delivery speeds and pitch ranges within a single snippet and at the same time capturing the lexical units in an eight-item sequence.

8.2.3.2 SUBJECTS AND METHODOLOGY FOR TEST 2

In order to stay within the scale of the PhD framework, the same participants as in Test 1 were re-tested in Test 2.

In Test 2, the same methodology as in Test 1 was employed. All the participants were exposed to and tested with real informal English speech at normal speed. The participants were asked to imitate native speech production to test for and demonstrate intelligibility. The same evaluation methodology as in Test 1 was also applied.

The same 5-point scoring system of 0-4 was kept for this second test. A greater number of students scoring a '4' highlighted the improvement achieved in Test 2. A score of '4' was given only for formulaic sequences exhibiting a real, fluent, native-like speech production; thus all the 'blur-words' within an envelope were given the same score, rather than judged word-by-word for citation form. Based on this methodology, all the data were first evaluated by the present author. Sample data would also be evaluated by a panel afterwards for further validation of methodology and marking scheme.

8.2.3.3 ANALYSIS OF TEST 2

As for Test 1, the evaluation of Test 2 was carried out in terms of the analysis of the sample snippets and the analysis of the test results, as discussed separately below.

A. Analysis of sample snippets

There were six samples in Test 2, as shown in Table 25. After evaluation, the six samples showed various results (see Figure 27) for average student performance (as a percentage of the maximum score per snippet):

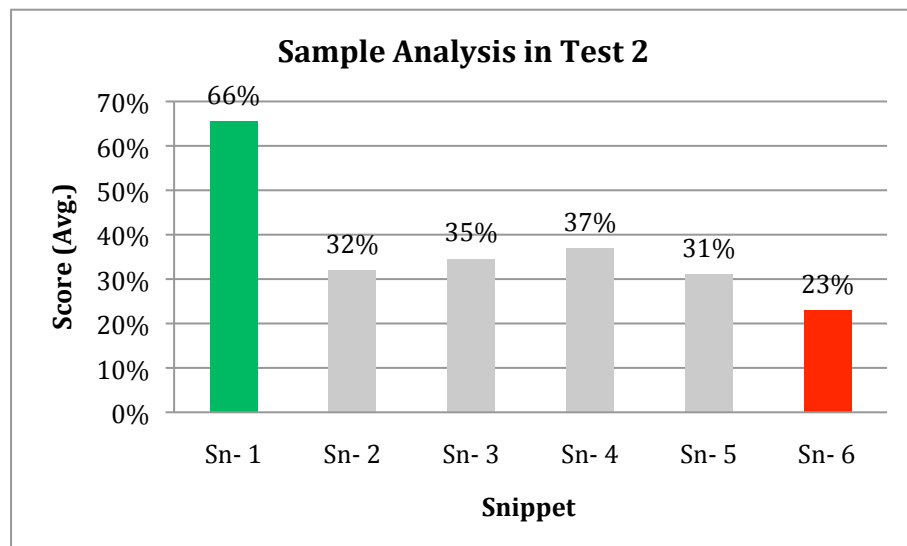


Figure 27: Sample analysis in Test 2

As anticipated, Sample 1 ('It is bananas.') was the one where students performed best. 65 out of 100 participants got '3' for all three items in this snippet. Only 3 of the participants got 25% correct, which was the lowest score for this sample. However, no participant achieved a score of '4'. The most likely reason for this performance was that there were two phenomena in this idiom. The first was the grammatical paradigm 'it is' with fast speed of 745 syll/min and flat pitch range of 11 Hz. The second was the rich intonation pattern on 'bananas', with a speed of 335 syll/min and pitch range of 96 Hz.

Most of the participants could capture the phonemes, but not the intonation patterns. As a result, only a flat imitation was given, rather than a native-like speech production. The other reason may be that this sample posed a lexico-grammatical problem. After recognising the sound and re-constructing the snippet, students provided what they imagined was a ‘correct’ version: ‘It is a banana’. It can be seen here that most of these Chinese English learners seem to mainly depend on the bottom-up processing method, concentrating on the phonemes and the syntactical structure, and failing to tune in to the intonation patterns.

Sample 6 (‘Which makes it, erm, you know, quite difficult.’), as expected, produced the worst performances, with an average score of 23% correct. The best participant got 81% correct, and 5 out of 100 failed to produce any correct phoneme at all. The most likely reason was the chunk ‘you know’ embedded in the middle of the snippet, which caused great difficulties for the participants in capturing the sound and following the intonation patterns at the same time. With the exception of the chunk part of the snippet, and the interjected word ‘erm’ in front of it, there were only five items in the snippet. The speed of delivery for the whole snippet was 275 syll/min and the pitch range was 98 Hz. However, the chunk part ‘you know’ was delivered at a faster speed of 330 syll/min and had a flat intonation pattern with a pitch range of 49 Hz. The large disparity in speed of delivery and pitch range resulted in perception and production problems for the students.

Apart from Samples 1 and 6, the other samples, Samples 2, 3, 4 and 5, were re-produced with a similar result. The most likely factor preventing the participants from achieving a better score was the reduced forms and weak forms in the testing samples. Sample 5 (‘He had no idea what it meant.’) presented more difficulties. 2 of the participants got an equally high score of 59% correct and another 2 of them got a score of ‘0’ due to

zero performance, which lowered the average of this sample to 31%. The reduced form ‘what it’ caused the main problem for the participants. 70 out of 100 participants got a score of ‘0’ when producing the item ‘what’, and 88 out of 100 got ‘0’ on the item ‘it’. In addition, the unstressed form ‘idea’ in the collocation ‘had no idea’ also resulted in an incorrect interpretation for the participants. There were two reduced forms in Sample 4 (‘You have a list of questions?’). 72 out of 100 participants did not catch the phoneme ‘you’ in the reduced form of ‘you have a’, and 89 of them failed to produce the non-lexical word ‘of’ in the highly reduced collocation ‘a list of’. In Sample 3 (‘When did you move out of home?’), the weak form of ‘out of’ caused 97 of the participants to get a score of ‘0’ on the item ‘of’ and 43 of them on the item ‘out’.

In addition to the contracted form vs. citation form, supra-segmental linguistic features were also ignored and were not re-produced by the Chinese participants. For example, Sample 2 (‘From pillar to post.’) was one of the two samples re-tested in both Tests 1 and 2, which was anticipated to lead to a better performance. The finding turned up a different result due to its lively intonation pattern with a slow speed of 241 syll/min and a wide pitch range of 113 Hz. 30 out of 100 participants got a score of ‘3’ on the citation part ‘from’, 2 on the item ‘pillar’, 4 on ‘to’, and 36 on ‘post’. However, no participant got a score of ‘4’ for the real NS production. The participants concentrated on catching the segmental sounds of the items, and failed to perceive the lively intonation patterns.

However, comparing the two samples used in both Tests 1 and 2, an overall improvement was evident, as shown in Figure 28.

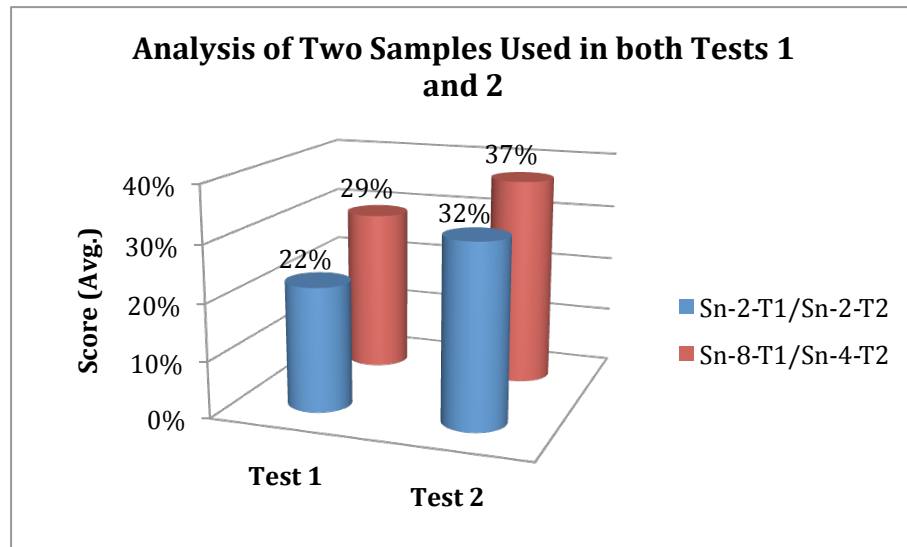


Figure 28: Analysis of two samples used in both Tests 1 and 2

Samples 2 and 8 were used in Test 1. After that, neither of the two samples was presented to the participants again. Thus, after the six months' training session, the same two samples, then re-labelled as Samples 2 and 4, were re-tested in Test 2 to evaluate any improvement. In Test 1, the best student performance for Sample 2 ('From pillar to post.') was 63% correct, and 20 of the participants got a score of '0'. In Test 2, the best participant achieved the same level of correct percentage as in Test 1; yet only 5 of the participants got a score of '0'. For Sample 8/4 ('You have a list of questions?'), the highest score for the participants in Test 1 was 67% correct. In Test 2, however, the best performance achieved a perfect native-like production with 100% correct. 4 out of 100 participants failed to produce the snippet in Test 1; but only 3 out of 100 failed to do so in Test 2. Therefore, some improvement can be seen due to exposure to real native English speech, especially in the case of these two samples, which were re-tested.

B. Analysis of results

In Test 2, the same participants were tested and the same methodology was applied, however the results were different, as shown in Figure 29.

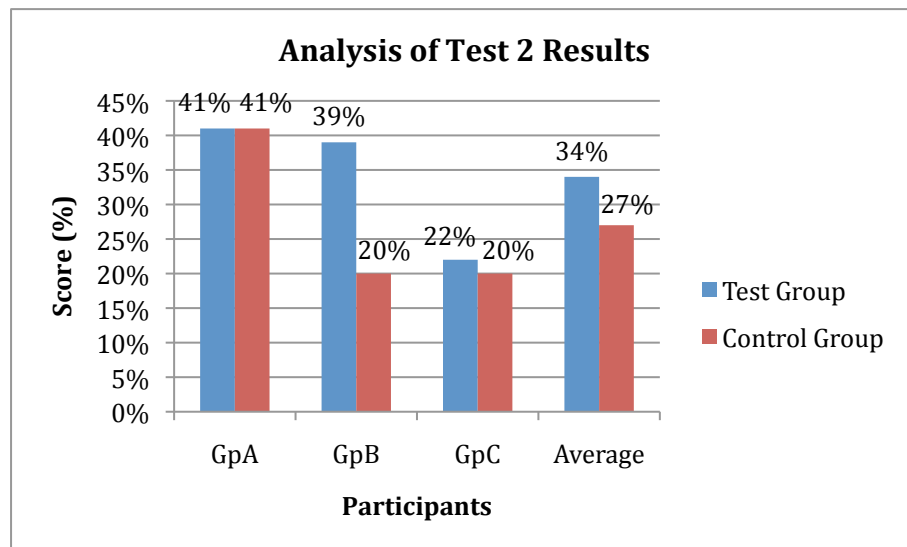


Figure 29: Analysis of Test 2 results

a) overall improvement by Test Group

Firstly, from the results, it can be seen that a higher level of intelligibility was achieved by the Test Group. The average score for the Test Group was 34% correct, but 27% for the Control Group. There was a considerable gap between Test Group and Control Group.

i) Group A

In GpA, the result was the same for both Test Group and Control Group. The first reason may be that Group A spent the least time on the training materials. Questionnaire 2 for the Chinese teachers showed that only 20 minutes a week was used for training, which resulted in less improvement in the training programme. Another reason may be that these students were at a third level college, and had to take a College English Exam in order to graduate with an honours award. Preparing for their formal written exam (which was due in April, 2008) also distracted the participants from the

training activity. Thus, there was no significant difference between the two groups in GpA.

ii) Group B

However, for GpB, there was a substantial discrepancy between the Test Group and the Control Group. The Test Group achieved 39% correct overall, yet a result of only 20% was gained by the Control Group. There may be several reasons for this. First, the greatest exposure was given to the participants in GpB. According to teacher responses to Questionnaire 2, the participants spent one hour a week on listening to and practising the training materials. A handout for all the training samples listing the relevant linguistic features was also given out by the teacher (see Appendix 29), which made it possible for the participants to access these samples and practise the characteristics whenever they wanted. The second reason may be that these participants in GpB were in their final year at junior secondary school. An Entrance Exam to the senior secondary school, oral and listening test included (which was different from the written exam in which GpA took part), would be taken in June 2008. Thus, the participants had more motivation than the other two groups. A very positive attitude can also be seen from Q1 in both Questionnaires 3 and 4. For these reasons, a 19% gap in performance emerged between the Test Group and the Control Group in GpB.

iii) Group C

GpC had an exposure to the training materials of 45 minutes a week. However, the students were in their second year at junior secondary school, and they had less linguistic knowledge than the other two groups to cope with (to them) fast, native-to-

native speech in a strange language variety. Thus, the Test Group achieved only a 2% higher score than the Control Group.

However, the overall 7% average higher score achieved by the Test Group in Test 2 indicates that it was probably the slow-down facility that allowed the Test Group to achieve a considerable improvement.

b) overall improvement by test participants

Secondly, in contrast with Test 1, it can also be seen that an overall improvement (net improvement in percent compared with Test 1, not average score) was achieved by most of the participants. 100 participants had been exposed to authentic English speech for six months. Sorted by the improvement in the participants' score, 61 of them achieved various levels of improvement (see Figure 30). The biggest improvement (achieved by Participant No. 41 from GpBt) was an increase of 28%. 39 of the participants retained the same level as in Test 1 or actually disimproved (mainly due to different testing samples being used). 11 out of 39 were from the Test Group, and 28 of them from the Control Group. Therefore, it can be seen that the overall improvement in Test 2 for most of the participants was because of the exposure to native English speech.

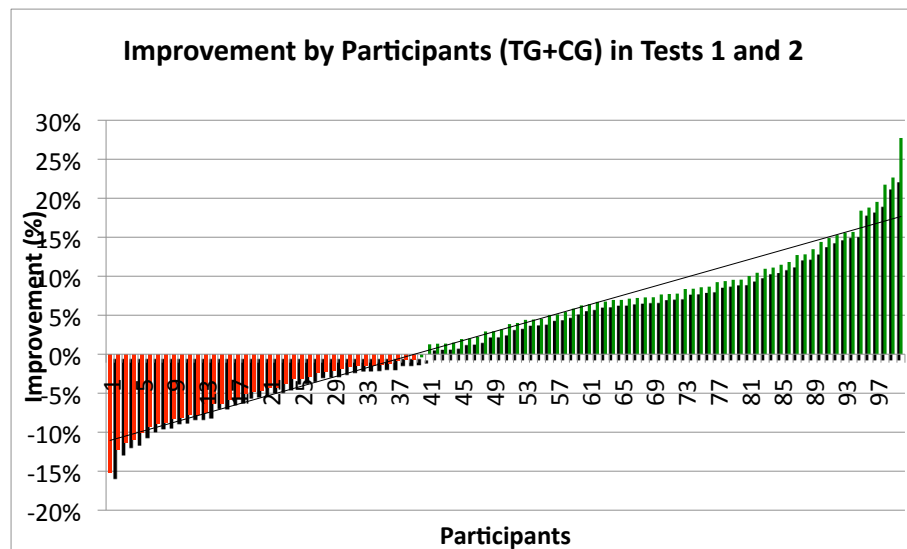


Figure 30: Improvement by participants (TG+CG) in Tests 1 and 2

c) individual improvement by participants in Test Group

Thirdly, separating the improvement by the Test Group and the Control Group, it can be seen that the Test Group performed considerably better than the Control Group. In Test 1, both groups achieved a performance of 27% correct. In Test 2, 50 participants in the Test Group returned a performance of 34% correct, with an improvement of 7% (the standard deviation is 0.08, 64% are within one standard deviation); and 50 participants in the Control Group achieved the same level on average as in Test 1 (with the standard deviation of 0.07, 58% are within one standard deviation) (as shown in Figure 31).

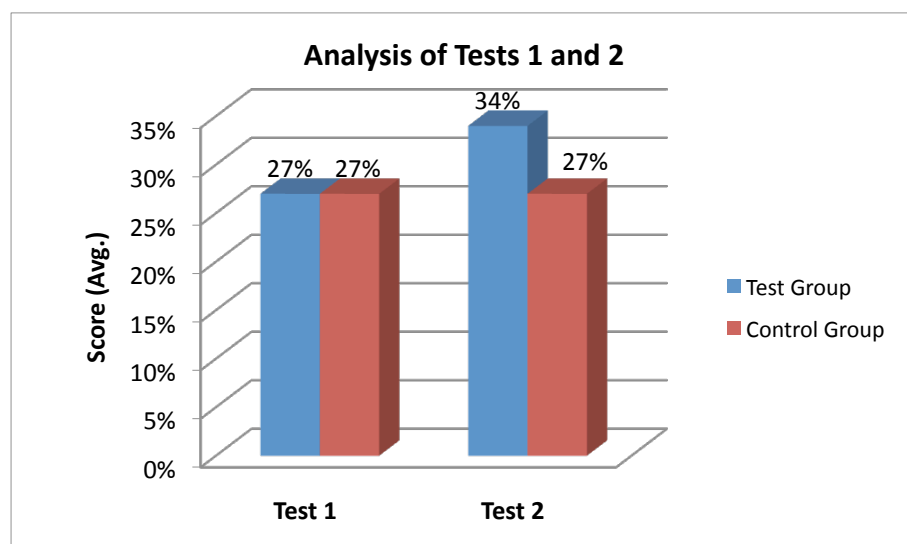


Figure 31: Analysis of Tests 1 and 2

From the analysis of individual improvement, 39 out of 50 participants in the Test Group achieved various levels of improvement. The highest improvement achieved by Participant No. 41 from GpBt was up to 28%. Only 11 of them showed no improvement. The greatest disimprovement was 12% (mainly in GpCt which had a lower linguistic ability). For the Control Group, 22 of the participants gained a certain level of improvement. The highest score arrived at was 16% up on that of Test 1. More than half of the 50 participants showed no improvement. The lowest performance among these 28 students was down by 15%. Thus, a large gap in the performances can be seen between the Test Group and the Control Group, especially in GpB. There was no disimprovement in GpBt, and the overall improvement was 12% performed by the Test Group in GpB. Both the curves and the trend lines in Figure 32 showed a consistent improvement across all the ranges between the two groups. This considerable improvement achieved by the Test Group indicates that Chinese language learners liked and benefited from not only the training activity, but also the slow-down facility.

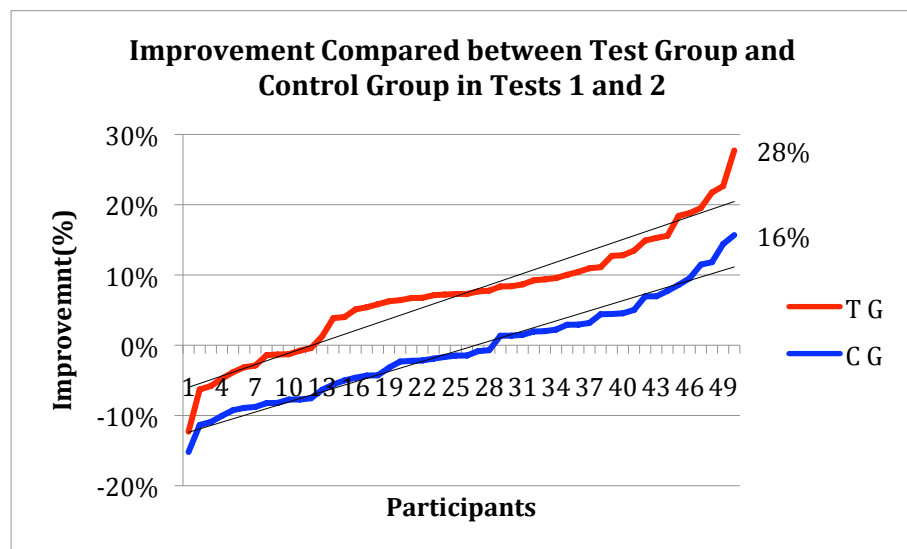


Figure 32: Improvement compared between Test Group and Control Group in Tests 1 and 2

- d) improvement in perception and re-production of native flowing English speech

The final point was that considerable improvements in phoneme recognition and native-like speech production emerged from Test 2.

- i) improvement in phoneme recognition

A number of participants achieved a score of '4'. The same methodology and the same evaluation scale 0-4 were applied in both Test 1 and Test 2. Since the Chinese participants had hardly been previously exposed to any real natural English speech, and in particular the Irish accents, no participants obtained a score of '4' in Test 1. However, after being exposed to the training materials for six months, 31 participants were first assigned a maximum Score '4' (not an average score of '4') on at least one of their native-like speech productions by the present author. To validate the scoring allocated, two months later, which sufficiently allowed all the previous memory of the 31 productions to reduce to a minimum, the present author, in conjunction with her supervisor – Dermot Campbell, reviewed all these 31 performances, and 29 of them were confirmed (see Figure 33). Two of them were excluded due to relatively weak corroborative evidence coming from the surrounding words, i.e., the scores for surrounding words were lower than '2'. 16 out of 29 were from the Test Group, and 13 from the Control Group.

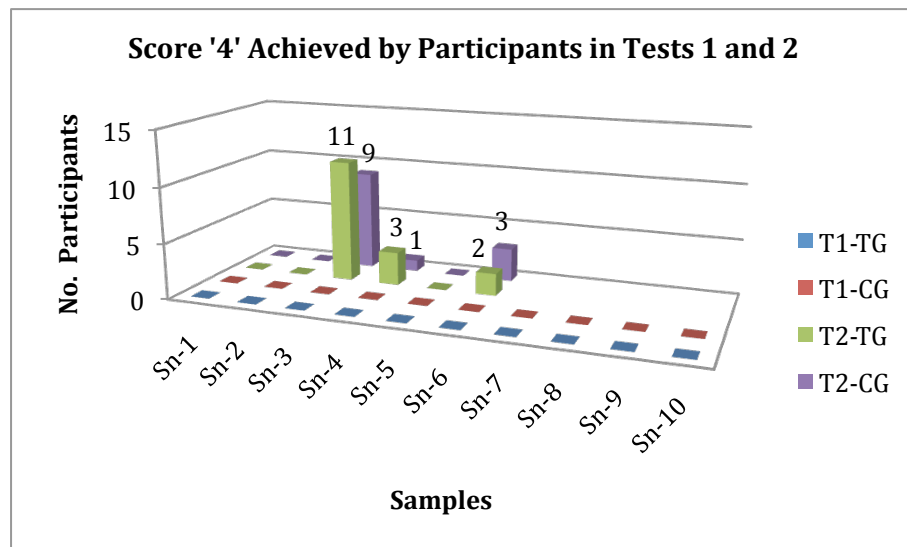


Figure 33: Score '4' achieved by participants in Tests 1 and 2

As shown in Figure 33, in Sample 3, for example, ('When did you move out of home?'), 20 of the participants achieved a perfect native-like speech production on the reduced form 'did you'. In Sample 4 ('You have a list of questions?'), there were 4 out of 100 who arrived at a native-like level of production of the 'blurred sequence' – 'a list of'. There were also 5 participants who performed a real, natural chunk 'you know' in Sample 6 ('Which makes it, erm, you know, quite difficult.'). Within the 60 training samples delivered to the participants, reduced forms, weak forms, formulaic sequences and English intonation patterns were highlighted. 13 out of 60 samples contained the reduced form 'did you', 'do you' or 'would you'. 6 out of 60 were about the practice of the reduced form of the non-lexical word 'a', and 5 of the non-lexical word 'of' as well. There were also 4 out of 60 samples on the training of the formulaic sequence 'you know'. As confirmed in both Questionnaires 3 and 4, being exposed to authentic, native English speech along with the slow-down speed allowed the participants to more effectively listen to and capture the sounds produced by the L1 English speakers.

A number of participants advanced from a lower level score to a higher level score. The improvement in phoneme recognition was not only demonstrated in respect of Score ‘4’ achieved by the participants, it was also verified by the positive achievement by the participants from lower score – i.e., Scores ‘0’ and ‘1’ to higher levels of Scores ‘2’ and ‘3’. Comparing the two samples re-tested in both Tests 1 and 2, substantial improvement can be seen in Test 2. In Sample 2 (as shown in Figure 34), there were 9 participants in the Test Group who could not produce any correct phonemes in Test 1; yet, in Test 2, the figure reduced to 1. 8 participants could get some of the phonemes produced correctly. There was a similar result for Score ‘1’. There were also 8 participants who moved up a grade. For the number of students with Scores ‘2’ and ‘3’, there were respectively 7 and 9 participants who achieved a higher level of performance, although no participants achieved Score ‘4’. For the Control Group, a similar improvement could be seen as in Figure 35, in which the number of participants in the lower level Scores ‘0’ and ‘1’ reduced and the number of higher level Scores ‘2’ and ‘3’ increased.

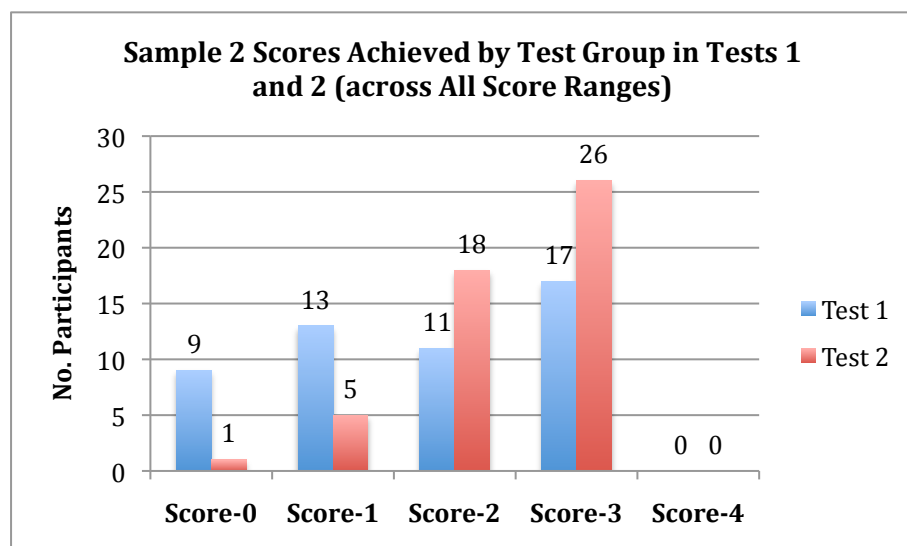


Figure 34: Sample 2 scores achieved by Test Group in Tests 1 and 2 (across all score ranges)

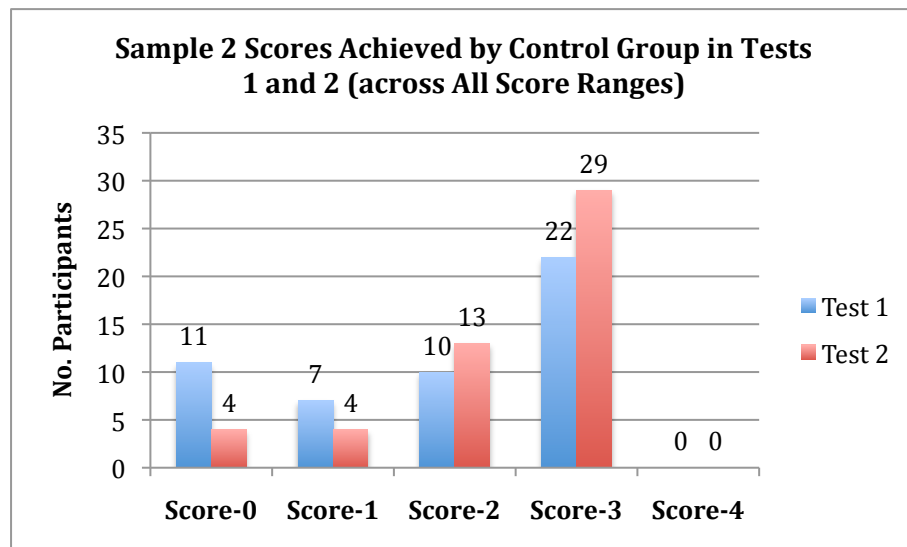


Figure 35: Sample 2 scores achieved by Control group in Tests 1 and 2 (across all score ranges)

As with Sample 2, there was an even greater improvement for Sample 4 (labelled as Sample 8 in Test 1) in Test 2. As shown in Figure 36, 3 participants in the Test Group increased their ability in word recognition and achieved a Score ‘3’, apart from 39 participants who had already got a Score ‘3’ in Test 1. An even more significant improvement was shown in respect of the achievement of Score ‘4’. There were 3 participants in the Test Group who achieved a perfect, native-like speech production. There was a similar situation for the Control Group (see Figure 37). 3 more participants improved to Score ‘3’, and 1 participant achieved a Score ‘4’ for a native-like flow of speech production.

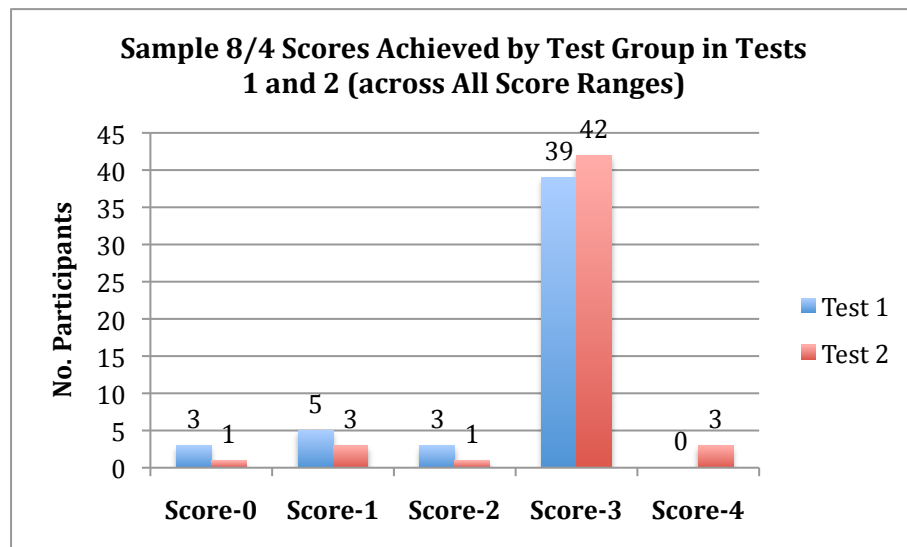


Figure 36: Sample 8/4 scores achieved by Test Group in Tests 1 and 2 (across all score ranges)

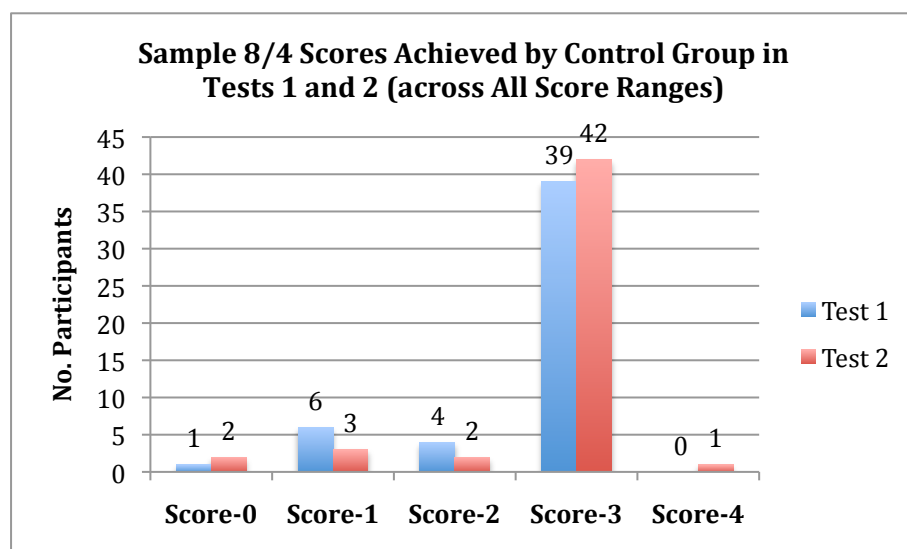


Figure 37: Sample 8/4 scores achieved by Control Group in Tests 1 and 2 (across all score ranges)

- ii) improvement in native-like speech production

Using the training materials and the slow-down facility allowed not only an improvement in phoneme recognition, but also was able to help the participants to tune themselves into the connected flow of English intonation patterns. For example, there was a chunk ‘you know’ embedded in Sample 6 (‘Which makes it, erm, you know, quite difficult.’). Compared with the complete snippet delivered at 275 syll/min and

98 Hz, the ‘you know’ part was produced at a faster speed of 330 syll/min and a flatter pitch range of 49 Hz. The chunk ‘you know’ in this snippet (see signal between the cursors in Figure 38) followed an interjected word ‘erm’ and a pause, used as a word-filler for buying time to plan and perform the subsequent utterance. A large gap in speed of delivery and pitch range could be perceived by L1 English speakers. In Test 2, a participant in GpAt achieved a perfect match for the intonation patterns on the chunk part ‘you know’ at a delivery speed of 276 syll/min and a 34 Hz pitch range (see Figure 39). Therefore, this is another indication of the benefits of exposure to a real English speaking community and access to the slow-down facility.

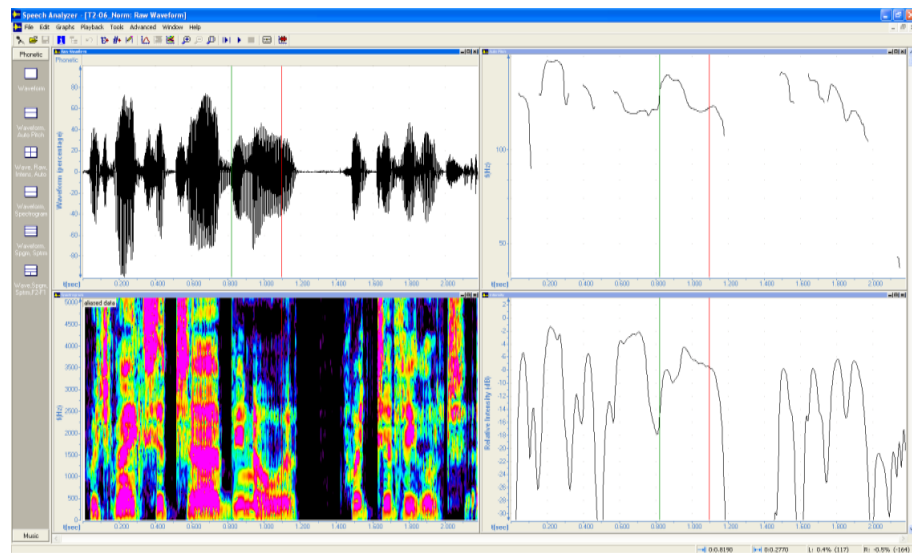


Figure 38: Exemplar production of chunk 'you know' in Sample 6

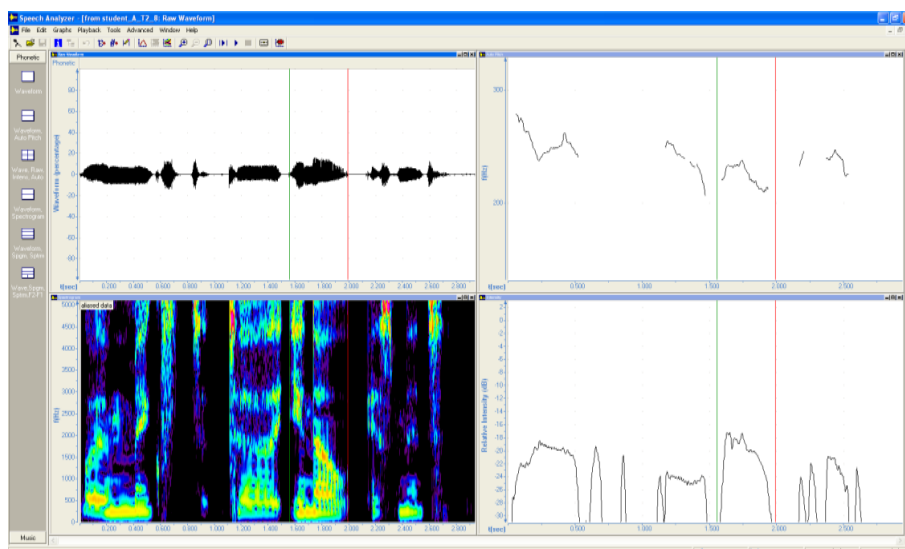


Figure 39: Participant's performance of chunk 'you know' in Sample 6

8.2.3.4 PANEL FOR FURTHER VALIDATION OF METHODOLOGY

Given that Test 1 was a benchmark test mainly used to highlight any improvement in Test 2, the evaluation of Test 1 data was done only by the present author. However, in Test 2, some improvement emerged, especially for the Test Group. In order to avoid the charge of subjectivity, the significant improvement in Test 2 required a panel to validate the evaluation methodology. This validation session is considered in detail in this subsection.

A. Rationale for choice of participant samples

There were 100 students who participated in Test 2, with each student re-producing six snippets; thus giving 600 productions in all. Only a representative number of response samples was needed for validation by the panel.

Samples which attained a rating of '4' in Test 2 were suitable candidates to be validated by the panel due to the fact that they contained native-like blur. A rating of '4' occurred

in Samples 3, 4 and 6. 20 students got ‘4’ in Sample 3, 4 students in Sample 4, and 5 in Sample 6. Sample 3 was chosen because of students’ better performance in it. Sample 4 was chosen because both Samples 2 and 4 were tested in both Test 1 and Test 2. It was thought useful to compare the differences in performance of the same sample in both tests.

For each of the two samples, all 100 student productions (an unsorted mixture of Test Group and Control Group) were ranked and divided into three bands according to the overall performance in the sequence – strong, medium and weak. Since more students gaining ‘4’ and ‘3’ would be in the strong band, thus increasing their ranking, two student samples were included from this band – one from the middle level and one from the bottom – the strongest performance was avoided. One sample only was chosen from the middle level of the medium band. There was also one sample chosen from the middle level of the weak band, rather than from the top level (which was too close to the medium band) or from the bottom level (which might contain less information). Among the response samples scored at the same rating, the one with fewer ‘0’s was chosen due to the extra information it contained. Thus, four snippets produced by the participants were chosen for each of the two exemplar samples. In all, eight student versions were chosen – four from the Test Group (i.e., No. 9, No. 34, No. 3 and No. 79) and four from the Control Group (i.e., No. 97, No. 30, No. 64 and No. 94), as shown in Table 26.

Exemplar Sample	Student Response Sample	Performance Level by the Participant Score (Avg.)	
Sample 3	No. 97 (GpCc)	Strong Band	2.1
	No. 30 (GpAc)		1.6
	No. 9 (GpAt)	Medium Band	1.3

	No. 64 (GpBc)	Weak Band	0.9
Sample 4	No. 94 (GpCc)	Strong Band	3.0
	No. 34 (GpBt)		1.3
	No. 3 (GpAt)	Medium Band	1.2
	No. 79 (GpCt)	Weak Band	0.8

Table 26: Choice of participant samples**B. Constitution of panel**

Compared with a panel of 5 used for validation in Jones and Haywood's (2004, p.280) study, for the current research, it was deemed advisable to choose more people in order to evaluate the methodology. Therefore, 11 people were chosen as a population of convenience by the present author. Since the panel were planning to give a judgment on the performances of native-like speech production, all of them were required to be L1 English speakers. Within the panel, four language teachers, with enhanced linguistic sensitivity, were used to represent the perspective of language learning and teaching. The other seven members in the panel were naïve (linguistically speaking) L1 English users chosen from a background other than linguistics, in order to obtain more information about general intelligibility of non-L1 speakers by L1 language listeners. It was anticipated that the feedback from the two groups in the panel would be interesting and informative. If the results from the two groups were similar or close to each other, that would corroborate the improvement achieved by the students. If there were substantial differences, this would require an explanation.

C. Procedures for evaluation of participant responses

Eleven members of the panel were invited by email to give their judgment on the students' performance. Written instructions were given, accompanied by the audio files of the exemplar Samples 3 and 4, and the eight response snippets produced by students.

The evaluation consisted of two steps. Step 1 involved the provision of written instructions, as shown in Appendix 30. Then the panel was asked to compare the eight response snippets (played in random order) and holistically rank them from most native-like to least native-like. After a gap of two weeks to allow sufficient time to overcome memory effects (personal communication with John Field, 11 September 2008) Step 2 was carried out. At this stage, the panel was asked to give a score to all of the snippets using the same evaluation methodology as the present author did (i.e., on a word-by-word basis). A detailed instruction on the scoring system and an evaluation spreadsheet file were given to the panel (see details in Appendix 31). The highest score for each lexical item was indicated in the spreadsheet. Meanwhile, all the student response audio files were made available to the panel as well, but in a different order to that in Step 1.

Each student sample was rated by all eleven people in the panel. Before each step, a pilot study was done using one of the panel to ensure the clarity of the evaluation instructions and rating practice. Then a formal evaluation was carried out using the whole panel. The panel was asked to remain consistent across all items and all students, especially, when they were not sure, they were asked to listen to the performance and re-score it for a more accurate rating. All the members of the panel were asked to evaluate the samples independently. After each step, the results were immediately sent back and the analysis carried out.

The evaluation by the panel was implemented in two steps. In each step, evaluation was given in terms of the panel as a whole, and also a comparative analysis between the two sub-groups involved in the panel – naïve L1 language speakers, and L1 language teachers.

a) Step 1: validating the ranking of participants' holistic performance

The aim of Step 1 was to evaluate the overall intelligibility of the student's performance in a holistic manner. After listening to the snippets as a whole, results were given by the panel.

i) Sample 3

The ranked order for Sample 3, based on the performance of four students, from the best to the worst, was given by the present author and labelled as Rank 1, Rank 2, Rank 3 and Rank 4, as shown in Table 27. The evaluation given by the panel is also shown in Table 27.

	Ranked Order (best to worst)			
Author Evaluation	Rank 1 (Student D)	Rank 2 (Student A)	Rank 3 (Student B)	Rank 4 (Student C)
Panel Evaluation				
Naïve L1 English Speakers Evaluation				
Panellist 1 (P1)	Rank 1 (R1)	Rank 2 (R2)	Rank 3 (R3)	Rank 4 (R4)
P2	R2	R1	R3	R4
P3	1	2	3	4
P4	1	2	3	4
P5	3	1	2	4
P6	2	1	3	4
P7	2	1	3	4

Language Teachers Evaluation				
P8	1	2	3	4
P9	1	2	3	4
P10	1	2	3	4
P11	2	1	3	4

Table 27: Sample 3 evaluation comparison between panel and the author (Step 1)

The evaluation of Step 1 was based on holistic listening and ranking, rather than accurate evaluation based on scores for individual words. A simple one-to-one yes/no matching between author rankings and panellist rankings would give a false picture of similarities and dissimilarities, since a single change in one ranking inevitably enforces a second ranking change. A fairly insignificant change in ranking is shown above for panellists P2 and P3. P2 ranks Student A as the best one with Student D in second place. Panellist P3 ranks Students D and A as best and second best respectively. In effect, the placement order of Students D and A has been swapped by panellists P2 and P3. A problem arises when there is a much more substantial change, or swap, in rankings between panellists. For example, supposing Panellist P3 had ranked Student D as the best student, where Panellist P2 had ranked that student as worst, then there is a significant discrepancy. Another significant discrepancy, though not quite as severe, is if the first and third rankings respectively for a particular panellist were placed in reverse order (i.e., in third and first place respectively) by another panellist.

Therefore, in order to compare the level of matching of rankings more reasonably between the panel and the present author, a swap pattern was proposed and employed in the evaluation. Table 28 shows the categorisation of all possible changes in ranking between that of the author and that of the panellists. As shown in this table, there might be some differences between the rank given by the panel and the present author, e.g.,

one swap, two swaps or three changes. Each swap is marked by the symbol \Leftrightarrow . Within each of these higher-order patterns, there can be one or more sub-patterns, depending on whether the differences in ranking are ‘adjacent’, ‘non-adjacent’, or ‘far non-adjacent’.

	Potential Pattern
1 swap (adjacent)	Rank 1 \Leftrightarrow Rank 2, or
	2 \Leftrightarrow 3, or
	3 \Leftrightarrow 4
1 swap (non-adjacent)	1 \Leftrightarrow 3, or
	2 \Leftrightarrow 4
1 swap (far non-adjacent)	1 \Leftrightarrow 4
2 swaps (adjacent)	1 \Leftrightarrow 2, and 3 \Leftrightarrow 4
2 swaps (adjacent and far non-adjacent)	2 \Leftrightarrow 3, and 1 \Leftrightarrow 4
3 changes	any 3 positions changed

Table 28: Swap patterns employed for ranking evaluation for both Samples 3 and 4 (Step 1)

Based on the swap patterns in Table 28, evaluation values needed to be assigned to each of the patterns. Comparing the ranking between the panel and the present author as shown in Table 27, it was evident that, among those panel members who did not produce a 100% match, they could easily distinguish between the better two performances and the worse two performances. Therefore, a line was drawn (conceptually) between Rank 2 and Rank 3.

The following weighting system is proposed to rank and highlight the degree of difference between the results of the panellists and those of the present author. Any swap between Rank 1 and Rank 2, or between Rank 3 and Rank 4, was evaluated as a ‘minor mismatch’, to which a score of -1 was given. Any non-adjacent swap, such as Rank 1 and Rank 3, Rank 2 and Rank 4, was evaluated as a ‘significant mismatch’, and a score of -2 was given. Any swap in the category ‘far non-adjacent’ was evaluated as a

‘major mismatch’ and a score of -3 was assigned to this mismatch. The swap between Rank 2 and Rank 3 was adjacent, however it crossed the line between the two better performances and the two worse performances. Thus, this was considered as a significant mismatch and scored as -2. Therefore, when three changes out of four of the rankings were assessed as a ‘very major mismatch’, this was marked as -4. Incorporating the equivalent mismatches and their assigned values, the assessment methodology for Sample 3 was updated as shown in Table 29.

	Potential Pattern	Equivalent Mismatch	Evaluation Value
1 swap (adjacent)	Rank 1 \Leftrightarrow Rank 2, or	minor	-1
	2 \Leftrightarrow 3, or	significant	-2
	3 \Leftrightarrow 4	minor	-1
1 swap (non-adjacent)	1 \Leftrightarrow 3, or	significant	-2
	2 \Leftrightarrow 4	significant	-2
1 swap (far non-adjacent)	1 \Leftrightarrow 4	major	-3
2 swaps (adjacent)	1 \Leftrightarrow 2, and 3 \Leftrightarrow 4	minor + minor	(-1) + (-1)
2 swaps (adjacent and far non-adjacent)	2 \Leftrightarrow 3, and 1 \Leftrightarrow 4	significant + major	(-2) + (-3)
3 changes	any 3 rankings changed	very major	-4

Table 29: Sample 3 evaluation values based on swap patterns (Step 1)

Therefore, based on swap patterns (as shown in Table 28) and their equivalent weighted values (as shown in Table 29), the evaluation for Sample 3 between the panel and the present author was assessed as shown in Table 30. The overall mismatch between the panel and the present author was very minor, with an average score of -0.5 mismatch from naïve L1 English speakers and -0.1 from language teachers. 100% match (with the

present author) was achieved by three panellists from the naïve L1 English speakers group and three from the language teachers group.

	Ranked Order (best to worst)			
Author Evaluation	Rank 1 (Student D)	Rank 2 (Student A)	Rank 3 (Student B)	Rank 4 (Student C)
Panel Evaluation (Avg.) very minor mismatch				
Naïve L1 English Speakers Evaluation (Avg.) -0.5 mismatch				
Panellist 1 (P1)	match		match	
P2	minor mismatch (-1)		match	
P3	match		match	
P4	match		match	
P5	very major mismatch (-4)			match
P6	minor mismatch (-1)		match	
P7	minor mismatch (-1)		match	
Language Teachers Evaluation (Avg.) -0.1 mismatch				
P8	match		match	
P9	match		match	
P10	match		match	
P11	minor mismatch (-1)		match	

Table 30: Sample 3 evaluation based on swap patterns and equivalent values (Step 1)

ii) Sample 4

With respect to Sample 4, the holistic ranking of four student performances given by the present author from the best one to the worst one is: Student C, Student B, Student A and Student D, respectively labelled as Rank 1 to Rank 4, which is shown in Table 31.

The evaluation given by the panel is also demonstrated in Table 31.

	Ranked Order (best to worst)			
Author Evaluation	Rank 1 (Student C)	Rank 2 (Student B)	Rank 3 (Student A)	Rank 4 (Student D)

Panel Evaluation				
Naïve L1 English Speakers Evaluation				
Panellist 1	Rank 2 (R2)	Rank 1 (R1)	Rank 3 (R3)	Rank 4 (R4)
P2	R2	R1	R3	R4
P3	1	2	3	4
P4	1	2	3	4
P5	3	2	1	4
P6	2	3	1	4
P7	2	1	3	4
Language Teachers Evaluation				
P8	1	2	3	4
P9	2	1	3	4
P10	2	3	1	4
P11	1	2	3	4

Table 31: Sample 4 evaluation comparison between panel and the author (Step 1)

The same swap patterns as shown in Table 28 are employed for the holistic ranking of Sample 4. However, different evaluation values are assigned based on the different gaps between each of the two ranks for Sample 4. As shown in Table 26, there is a large gap of, on average, 1.7 between Rank 1 and Rank 2. The gaps between Rank 2 and Rank 3, and Rank 3 and Rank 4, are much smaller, with respectively only 0.1 and 0.4. Therefore, a line was drawn (conceptually) between Rank 1 and Rank 2. Any swap across Rank 1 and Rank 2 is evaluated as either a significant mismatch, scored as -2, or a major mismatch which is scored as -3, or a very major mismatch evaluated as -4, depending on whether they are adjacent, non-adjacent, or far non-adjacent. Any swap between Rank 2, Rank 3 or Rank 4 is evaluated as either a minor mismatch scored as -1, or a significant mismatch which is marked as -2, depending on whether they were adjacent or non-adjacent. Three changes out of four rankings are assessed as a ‘very, very major mismatch’, scored as -5. Thus, various evaluation values based on different swap

patterns are used for the ranking of Sample 4, as shown in Table 32. Note that this is different to Table 29 because the holistic performance of the highest ranked student was so considerably better than the other students.

	Potential Pattern	Equivalent Mismatch	Evaluation Value
1 swap (adjacent)	Rank 1 \Leftrightarrow Rank 2, or	significant	-2
	2 \Leftrightarrow 3, or	minor	-1
	3 \Leftrightarrow 4	minor	-1
1 swap (non-adjacent)	1 \Leftrightarrow 3 or	major	-3
	2 \Leftrightarrow 4	significant	-2
1 swap (far non-adjacent)	1 \Leftrightarrow 4	very major	-4
2 swaps (adjacent)	1 \Leftrightarrow 2, and 3 \Leftrightarrow 4	significant + minor	(-2) + (-1)
2 swaps (adjacent and far non-adjacent)	2 \Leftrightarrow 3, and 1 \Leftrightarrow 4	minor + very major	(-1) + (-4)
3 changes	any 3 rankings changed	very, very major	-5

Table 32: Sample 4 evaluation values based on swap patterns (Step 1)

Therefore, according to the swap patterns (as shown in Table 28) and their equivalent values in Table 32, the evaluation of Sample 4 between the present author and the panel is shown in Table 33. A minor mismatch is arrived at between the present author and the naïve L1 English speakers (-1 mismatch), and also between the present author and the L1 language teachers (-0.9 mismatch). Two of the naïve English speakers and two of the L1 language teachers get 100% match compared with the present author's evaluation. Nevertheless, the overall evaluation of both Samples 3 and 4 achieve an overall minor mismatch.

	Ranked Order (best to worst)
--	------------------------------

Author Evaluation	Rank 1 (Student C)	Rank 2 (Student B)	Rank 3 (Student A)	Rank 4 (Student D)	
Panel Evaluation (Avg.)					minor mismatch
Naïve L1 English Speakers Evaluation (Avg.)					-1 mismatch
Panellist 1	significant mismatch (-2)		match		
P2	significant mismatch (-2)		match		
P3	match	match			
P4	match	match			
P5	major mismatch (-3)			match	
P6	very, very major mismatch (-5)			match	
P7	significant mismatch (-2)		match		
Language Teachers Evaluation (Avg.)					-0.9 mismatch
P8	match	match			
P9	significant mismatch (-2)		match		
P10	very, very major mismatch (-5)			match	
P11	match	match			

Table 33: Sample 4 evaluation based on swap patterns and equivalent values (Step 1)

iii) conclusion on validating the ranking of participants' holistic performance

The overall mismatch in Step 1 between the panel and the present author is minor, especially in Sample 3. This indicates that the evaluation given by the panel is acceptable and thus informative, and it also indicates that the evaluation done by the present author is acceptable and reliable. Another point which emerges from the validation exercise is that it can also be seen that there is a slight discrepancy between the result from the naïve language listeners and the L1 language teachers. The overall level of matching rated by language teachers is higher than that given by naïve language users. The reason is, given that the aim of Step 1 was to holistically listen to the samples and then evaluate the intelligibility, L1 language teachers had more experience in exposure to non-L1 language learners, therefore, they were more familiar with and more

tolerant of various accents which facilitated their understanding. This point is in agreement with Thompson's (1991) observation that the experienced raters in her experiment were significantly more lenient towards deviations in non-L1 language learners' production. As for naïve language users, without this professional exposure, most of time they judged the samples in comparison with a 'native' speech production, therefore, these samples produced by Chinese EFL learners caused them some difficulties.

- b) Step 2: validating the ranking of participants' performance on a word-by-word basis

The aim of the analysis in Step 2 was to evaluate student's performance by assigning an accurate score based on the same 0-4 evaluation scale employed by the present author so as to validate the methodology. A detailed instruction on the scoring system and an evaluation spreadsheet file were given to the panel (see Appendix 31), along with students' production audio files (in a different order to that in Step 1).

- i) Sample 3

First of all, for Sample 3, the average scores rated by naïve language speakers for four students' performances were 1.0, 1.3, 0.4 and 2.2, as in the rating shown in Table 34 and Figure 40. Also in Table 34, the overall scores given by the present author to Students A, B and C were a little higher. However, it can be seen that there was no significant difference. Even in the case of Student D, there was a gap of only 0.1 (2.2 - 2.1) between naïve language users and the present author, which is, in the integer scale used, totally negligible. As for the results of language teachers, as shown in Table 34 and Figure 41, some discrepancies emerged. A slightly higher score for Student A was

given by language teachers, the scores for the other three students' performances assigned by language teachers were lower than those given by the present author, especially for the performance of Student C with a gap of 0.6. However, the average scores between the panel and the present author seemed to be moderate (see Table 34 and Figure 42), with slight differences in Student Performances A and D, and a bit bigger gap in Performances B and C (with the biggest gap of 0.6).

	Evaluation Score (Avg.)			
	Author	Naïve Language Users (n=7)	Language Teachers (n=4)	Panel (Avg.) (n=11)
Student A	1.3	1.0	1.6	1.3
Student B	1.6	1.3	1.4	1.4
Student C	0.9	0.4	0.3	0.3
Student D	2.1	2.2	1.7	1.9

Table 34: Sample 3 evaluation comparison between panel and the author (Step 2)

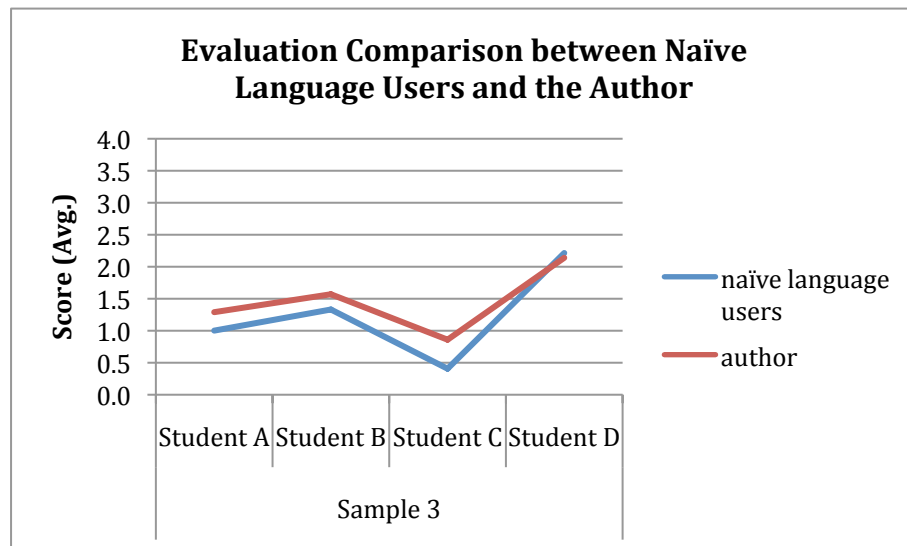


Figure 40: Sample 3 evaluation comparison between naïve language users and the author (Step 2)

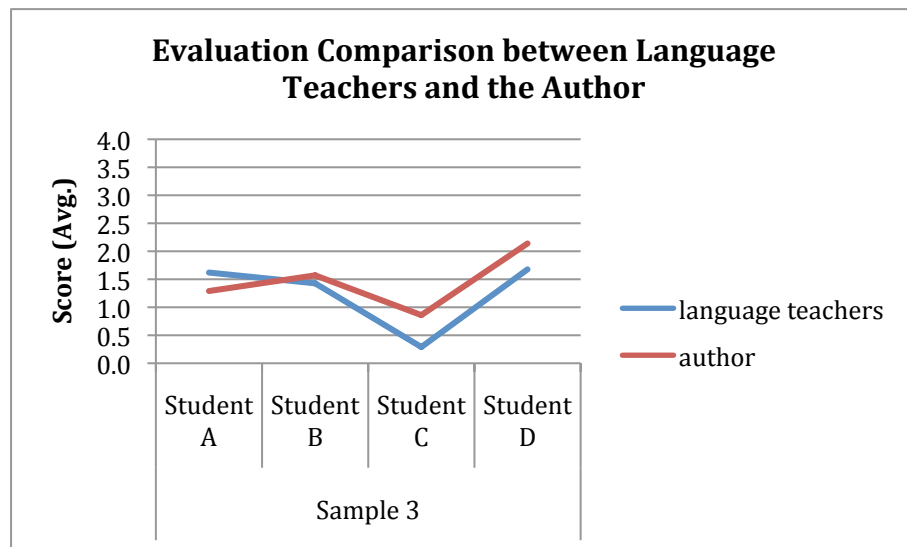


Figure 41: Sample 3 evaluation comparison between language teachers and the author (Step 2)

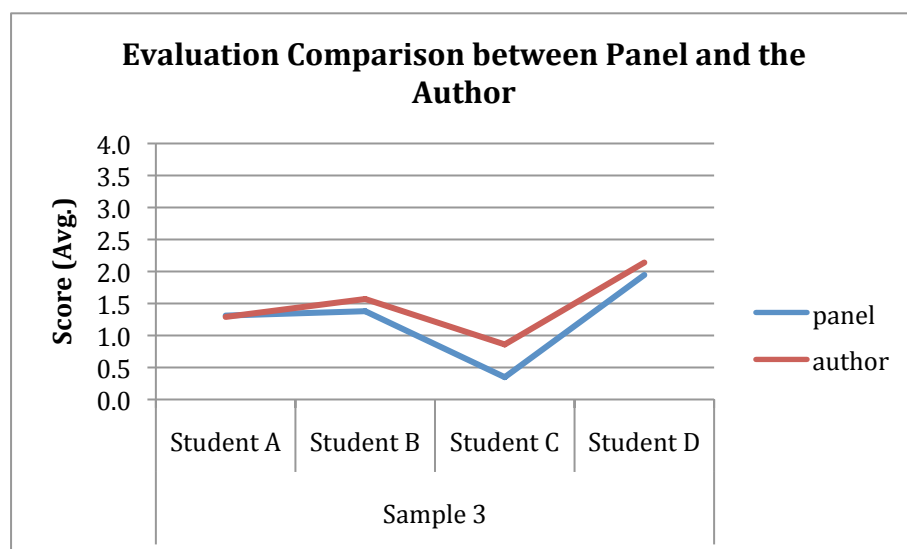


Figure 42: Sample 3 evaluation comparison between panel and the author (Step 2)

ii) Sample 4

With respect to Sample 4, the overall evaluation scores given by the present author were consistently higher than those rated by naïve language users (see Table 35 and Figure 43), especially in the performance of Student A with a large gap of 0.8 (out of a possible 4.0, meaning a 20% difference in evaluation). The comparison in scores between language teachers and the present author turned up considerable discrepancies in

Sample 4. As shown in Table 35 and Figure 44, there were slight differences in Student Performances B and C, however, there were relatively large gaps in the Performances A and D, in particular, with a considerable gap of 1.6 higher in Student D given by the present author compared with the language teachers. The average scores of the panel (see Table 35 and Figure 45) seemed to slightly reduce the gap between the panel and the present author. Apart from no gap in Student Performance C, there were some gaps in the other three performances, especially in Students A and D with a gap of 0.7 and 1.1.

	Evaluation Score (Avg.)			
	Author	Naïve Language Users (n=7)	Language Teachers (n=4)	Panel (Avg.) (n=11)
Student A	1.2	0.4	0.6	0.5
Student B	1.3	0.9	1.1	1.0
Student C	0.8	0.5	1.0	0.8
Student D	3.0	2.4	1.4	1.9

Table 35: Sample 4 evaluation comparison between panel and the author (Step 2)

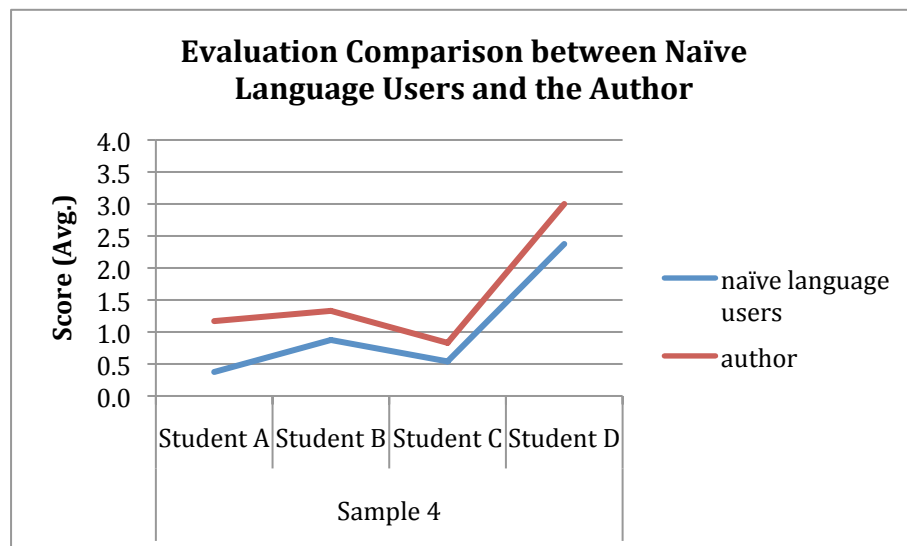


Figure 43: Sample 4 evaluation comparison between naïve language users and the author (Step 2)

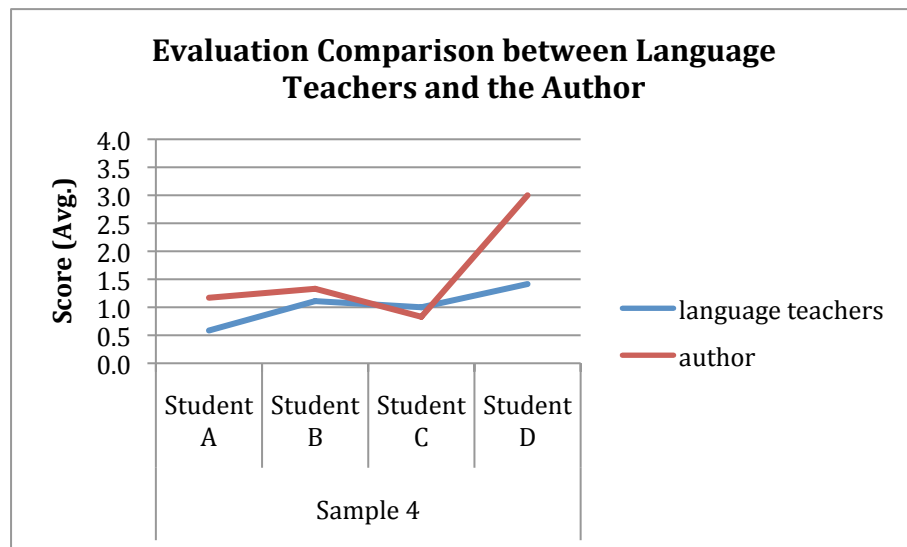


Figure 44: Sample 4 evaluation comparison between language teachers and the author (Step 2)

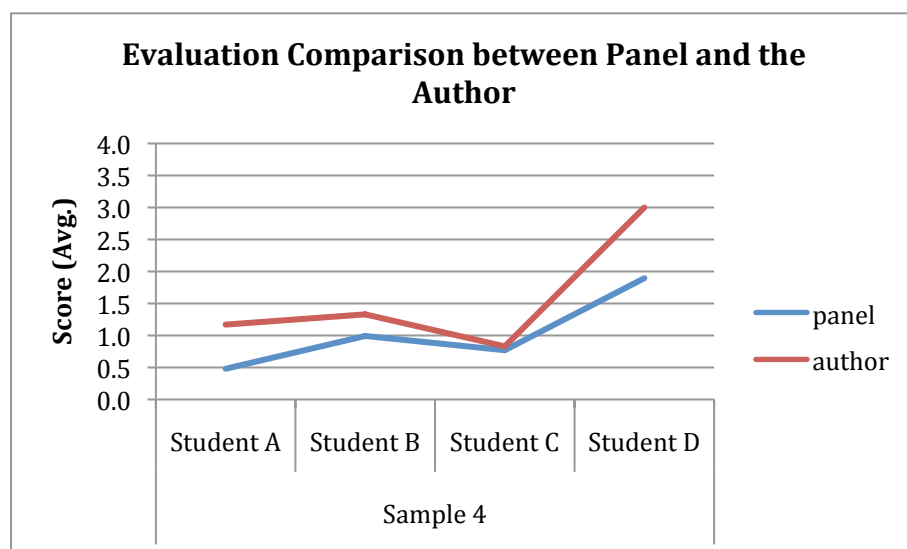


Figure 45: Sample 4 evaluation comparison between panel and the author (Step 2)

- iii) conclusion on validating the ranking of participants' performance on a word-by-word basis

The overall scores rated by the present author were slightly higher than those given by the panel. The main reason might be the evaluation by the present author was based on a word-by-word level, and also down to a phoneme-by-phoneme level, in line with the scoring system. For example, in Sample 3, Score '1' was given by the present author on

the word ‘move’, due to the fact that the individual phoneme /v/ was produced, even though the participant reconstructed another word and the word as a whole was not recognisable. However, most of the panel (7 out of 11) did not mark this phoneme, since they mainly operated at word level recognition. The same applied to the assigning of Score ‘4’. In Sample 4, ‘a list of’ was a formulaic sequence with an embedded acoustic blur. Thus, a Score ‘4’ was given to each of the words within this acoustic envelope. Compared with the evaluation by the panel, only 2 of them (18%) marked this sequence according to the same criterion. Most of them did not mark it as a formulaic sequence, i.e., they assigned different scores to these three elements instead of applying a Score ‘4’ to the whole formulaic sequence. Another reason might have been that the present author is a non-L1 English speaker, and shares the same mother tongue as the participants. She is more familiar with Chinese accents than the panel and could therefore more easily recognise individual sounds. There might be also a third reason. The evaluation of data done by the present author was based on the overall analysis of 1,600 test samples in both Test 1 and Test 2, which provides a more reasonable and reliable framework to facilitate the relatively accurate evaluation. Lack of sufficient data and differences in individual judgement could also result in the discrepancies in evaluation between the panel and the present author.

Another point which needs to be considered is that there were some discrepancies in the scores assigned by naïve language users and L1 language teachers. The reason might be, in contrast to Step 1 (holistic evaluation), the evaluation in Step 2 (word-by-word evaluation) was to judge the performance with an accurate score based on word-by-word/phoneme-by-phoneme recognition, acoustic blur and intact intonation pattern (for formulaic sequences). Given the different background of the two sub-groups in the

panel, different results inevitably emerged. Naïve language users, as L1 language speakers, pay more attention to overall comprehension, rather than perception of individual sounds. Besides, naïve language speakers have relatively less specialised knowledge of linguistics, i.e., formulaic language, intonation patterns; most of the time they evaluated by intuition. In contrast, for language teachers, having more knowledge of linguistics and more experience in dealing with non-L1 language learners' production, they tended to evaluate samples professionally with stricter criteria; therefore relatively low marks were awarded by the language teachers on the panel. Research carried out by Warren, et al. (2009, p.97) also confirms that in rating prosodic features, experienced English language teachers are no more reliable than naïve raters.

E. Conclusion of panel validation

As stated earlier, the panel consisted of eleven people; seven of them were naïve L1 English speakers and four were language teachers; all were L1 English speakers. Incorporating the data from both experienced L1 language teachers and naïve L1 language users is of methodological importance, in that evidence for different ratings in evaluation of non-native English speech by each sub-group is provided. As Warren, et al. (2009) note, very few studies – only theirs and that of Thompson (1991) – ‘compare the ratings of experienced and naïve raters’ (p.92).

Validation of methodology was carried out in two steps. Even though different results emerged in certain samples between the panel and the present author, and between the two sub-groups within the panel, however, the overall evaluation indicated that, the current evaluation methodology employed by the present author was acceptable and thus legitimate. Given the different background of the panel and the present author and

different evaluation methods adopted individually by them, it is not surprising that there is some difference between the two groups, and between the panel and the present author, especially in Step 2. However there is a clear similarity in the overall results. The evaluation method carried out by the present author, however, was a systematic evaluation framework based on 1,600 student samples, and thus could fairly be deemed more reliable.

8.2.3.5 DISCUSSION OF RESULTS OF TEST 2

After six months' exposure to authentic, natural English speech training materials, first of all, most of the participants registered an improvement over the results obtained in Test 1. Different samples (with the exception of Samples 2 and 8/4) were used in Test 1 and Test 2, and therefore the overall improvement cannot be directly measured in absolute terms. However, the most telling contrast was reflected in the results of the two samples tested in both Test 1 and Test 2, in that a significant increase in average scores was evident, especially in Sample 2, with a 10% improvement.

In addition to the overall improvement by the participants, the Test Group was also shown to achieve a higher level of intelligibility than the Control Group. Compared with the approximately same level of proficiency in English language between Test Group and Control Group shown in Test 1, a 7% gap improvement emerged in Test 2. Another finding which also showed a higher improvement by the Test Group was that, there were more participants who merited a higher score band, in particular, more students achieved from a partial or complete receptive mastery (i.e., Score '2' or '3') to a more proficient productive mastery in the native-like re-production of formulaic sequences (i.e., Score '4'). The absolute higher level of improvement achieved by the

Test Group clearly indicated the effect of the slow-down technology, which was the only significant variable between the two groups.

It would seem, therefore, that exposure to real, informal, connected English speech afforded the participants a chance to listen to and compare the discrepancies between dynamic, flowing NS and transcribed, citation-form, read-out performances. Access to the slow-down facility allowed the participants to perceive segmental and supra-segmental levels of linguistic features so as to improve intelligibility and cope with NS from an English speaking community.

8.2.4 Discussion of Tests

In this section, issues arising from the tests are considered. The discussion is structured as follows:

- A. justification of the aim of the tests
- B. issues of methods and procedures for evaluation
- C. factors influencing the interpretation of test data
- D. mfactors affecting the performance of individual participants

- A. Justification of the aim of the tests

Chapter 3 considered the issues facing Chinese EFL learners when involved in English native-speaking environment. Therefore the aim of the tests was to evaluate Chinese language learners' linguistic ability in imitating and re-producing native-like English speech, rather than training and teaching the pronunciation of isolated individual sounds.

There are two levels of intelligibility evaluated in the tests. One is the receptive level,

and the other is the productive level. Given that English is widely used by non-L1 English speakers, as long as sufficient segmentals can be extracted from the signal and correct semantic labels can be assigned to the speech signal and pragmatic inferences effected, then intelligibility is achieved and the speaker's communicative intention is realised. Non-L1 learners of English do not need to be taught or learn the production of native-like blur. Therefore, Score '3' is given to the participant's performance when a complete receptive mastery is demonstrated, i.e., correct citation form is produced. Citation imitations of NS blur are clear proof that the student understood the native-like blur, but either chose not to or could not imitate the principled blur of the NS.

Compared with Score '3', Score '4' emphasises a higher level of language communicative competence – the productive mastery of the target language. If the participant heard the native-like blur, and could also produce an intelligent blur, this demonstrates that they could tackle with the casual L1-L1 English speech, since blur is not due to oral laziness on the part of the speaker, but rather to a principled 'differential focussing' which allows listeners to concentrate on the parts of the production to which the speaker meant them to attend. This is not saying that native-like blur is necessary for production, but rather that there can be no doubt about language learners' ability to 'decode' blur if they can obviously 'encode' it. Therefore, to increase the learner's ability to recognise the principled nature of NS blur by hearing it, contrasting it with the citation form, slowing it down to follow the prosody and finally imitating it is the ambition proposed by the present author in the current study.

Thus, in order to evaluate the better application of the slow-down facility and the Dynamic Speech Corpus assets, in the facilitation of Chinese EFL learners in learning and acquisition of natural authentic English speech so as to cope with the target

language speaking community, the current research was carried out by the present author. During the study, some issues were also considered.

B. Issues of methods and procedures for evaluation

Firstly, from the tests it is clear that, despite the considerable improvement achieved in Test 2, in general, the performances of participants are at a very poor level, in that sometimes it is difficult to even distinguish one participant's response from the others'. Most of the exemplar samples used in the tests are very difficult for the participants to perceive and produce. This once more supports the argument proposed by the present author that Chinese EFL learners are too removed from principled access to real, natural, casual English speech, and that is why they need language learning technologies to bridge the gap.

Secondly, it is very difficult to make a clear-cut judgement and confidently award a score in many cases. That is also an explanation for some discrepancies in evaluation between the panel and the present author. Sometimes, the same student recording has to be re-listened to and re-scored several times. A score of '0' is easy to give. Score '1' is easy to assign for initial consonants in particular. Sometimes the score is to be assigned to a middle or terminal phoneme, even though it is clear that the participant had no idea what the correct word is, if he heard only one of the correct phonemes. This, however clearly demonstrates how much of the master recording was intelligible. Score '2' is difficult to distinguish from '3', in that it is unfair to give a lower score for a NNS production of the word, rather than listening to the participant's response as evidence of having heard the correct word. Score '3' is easy to give, where it is clear that the participant heard the correct word and re-produced it. Evidence of intelligibility, rather

than L1 production quality determines the score; that is to say, a Chinese accent is acceptable. In contrast to Score '3', the criterion for Score '4' insists on a NS-like production (for formulaic sequences), including acoustic blur and an intact prosody pattern. Judging the native-like blur is easier than the assessment of the intonation envelope, which is also discussed later in this section.

C. Factors influencing the interpretation of test data

Apart from the methods and procedures for evaluation, in the process of elicitation and validation of test data, some other factors can be identified which might influence the interpretation of test data.

Firstly, in terms of evaluation of test data, given that the assigned rating is mainly based on listening and evaluating, there is a degree of uncertainty, especially in the assessment of melodic features, i.e., in the phonological realisations of formulaic language. Objectivity might be enhanced by the incorporation of an advanced technique such as the low-pass filtering used by Van Els and De Bot (1987), in which detailed segmental information is removed and prosodic features are kept intact.

Secondly, in terms of a legitimate interpretation of test data, as mentioned above, it is not an easy task. Even though the process of evaluation as applied by the present author is to a large extent objective and consistent, a 100% deviation-free metric can never be attained. The reliability of assessment and validation of test data are affected by many factors outside the control of the present author, i.e., the time of the day, the physical and mental state of the students at the time of assessment, and the context in which the assessment is carried out.

D. Factors affecting the performance of individual participants

In addition to the limitations relating to testing evaluation, there are also some other factors which might affect the validation of the current research, i.e., personal attributes, and random factors. Personal attributes refer to gender, age, and background knowledge, etc., of the participants. As shown in Questionnaire 1 (see Appendix 3), there are differences in participants' gender between the Test Group and the Control Group in the senior secondary school group, and also an age gap between the Test Group and the Control Group in the junior secondary school group. Random factors refer to temporary fluctuations in participants' physical condition, stress levels or mental alertness at the time of the test. In the process of data collection, even though all the participants within the same group (i.e., GpA, GpB and GpC) were tested on the same day and under the same conditions, however, it could not be guaranteed that all 100 participants were tested at the same time or under the same testing context, or even ensure that all the participants were tested under the same physical or mental conditions. These factors are also anticipated to influence the test performance of individual participants.

Restricted by some potential controlled or uncontrolled factors as considered above, therefore, the current study inevitably illustrates to a certain extent some limitations which require further research.

8.3 Conclusion

Based on the scope of the current research, in order to evaluate the language ability of Chinese learners of English in imitating and re-producing native-like informal English speech on formulaic sequences by using speech technologies, a case study was described in this chapter. The application and evaluation of the use of speech

technologies, including Test 1, Test 2 and the training session, are described. The overall conclusion shows a high suitability of the application of speech technologies to non-L1 language learning and teaching. Feedback from linguistic researchers at the BAAL conference 2008 also justifies initial expectations (see Appendix 32). Some considerations on testing and evaluation methods are also discussed in this chapter.

In the next chapter, a discussion of the present research and how it fits into the current literature is provided.

9. Discussion of Present Research

The research work undertaken during the preparation of this thesis is detailed in the previous chapters. By revisiting the current literature, the nature of the contributions in the present study is highlighted, as follows:

In respect of the linguistic research carried out by the present author leading to Contribution 1, several new concepts are introduced into the body of knowledge. The first is the introduction and use of the ‘flow unit’. As examined in Section 7.2.2, the flow unit is an ideal production unit for segmenting and tagging the speech sequence, due to its unique features of brevity, accurate indication of speed of delivery and its ability to facilitate the subtle perception of pitch contour. By way of contrast, other linguistic units, as reviewed in Section 2.2.2, tend to be longer and with relatively intact semantic meaning or syntactic structure. However, as reviewed in Chapter 2, natural, dialogic, informal L1-L1 English speech is mostly characterised by frequent pauses, imperfect ‘flaws’ or sudden changes in pitch direction. The flow unit, therefore, is a natural, interactive unit which better matches the production unit of casual English speech. As discussed in Section 7.2.2, the flow unit is better suited than conventional units to capture the flow features of natural, informal, dialogic English speech, especially speed of delivery and pitch range. However, there is no direct relationship between the flow unit and formulaic language, except that there is a tendency for formulaic sequences to occur within a single flow unit due to their phonological and prosodic integrity.

The second concept introduced is ‘communicative speed’. Communicative speed, as considered in Section 7.2.2, is a new metric for dialogic communication, which

incorporates the effect of flow features such as contractions and elisions. It measures the number of citation equivalent syllables of any word present – no matter how minimally – in the speech signal. This elevated speed allows the learner to more easily identify those iterations which contain more phonetic reduction, especially in formulaic sequences. The learning value consists in contrasting the form which the learner has internalised and the form he actually hears, and eventually guiding the listener's attention to speaker-determined prominences. By contrast, articulation rate and speaking rate, as employed by other researchers, can only indicate the speed of delivery (whether excluding or including pauses) by calculating the number of syllables actually spoken in the signal. The natural features of phonetic 'blur' in real, authentic native-to-native English speech therefore cannot be demonstrated, and cannot easily be noticed and appreciated by language learners.

The third new concept is the introduction of phonological characteristics of formulaic language into the framework of formulaicity. As reviewed in Sections 2.3.3 and 2.4, certain terms and categories are defined and identified by several researchers, and some phonological features as studied by several researchers are also investigated. These analyses are either based on written texts or spoken forms with phonological features determined by the setting of the recordings (e.g., TV studio recordings). In addition, there is no analysis in the literature aimed at establishing the correlation between the various categories of formulaic language and their phonological realisations. The research work undertaken by the present author is based on real, natural L1-L1 speech assets, and links specific prosodic features, such as speed of delivery and pitch range, to different categories. This fills the identified gap in the literature by enriching the understanding and use of the communicative values of formulaic language. The advantages of this approach were confirmed in a conversation with Alison Wray

(Personal communication, 11 September 2008) at the 2008 BAAL conference, where the present author presented this work.

The research work relating to Contribution 2 is investigated in the evaluation of the application of a slow-down technology, combined with deliberately devised training materials, in facilitating Chinese EFL learners in the imitation and re-production of native acoustic ‘blur’ and intonation patterns. As discussed in Chapters 5 and 6, a 40% speed was exclusively chosen in the training sessions. The two and a half time’s longer exposure to the natural flow of speech allowed the learners to better perceive the reduced production and to follow the native speech prosody. Compared with previous research, as reported in the literature, which mainly focused on the segmental level of word recognition and pronunciation, the current study concentrates on the perception and production of the supra-segmental level of native phonetic blur and L1-L1 intonation patterns exemplified in formulaic language by using the slow-down facility. The investigation of the effectiveness of the technologies applied in bridging the east-west production and prosodic gap extends the application of the slow-down algorithm in language learning and teaching, beyond the still natural-sounding 80% speed to exploit the perceptible advantages of the less natural but more informative 40% speed. This therefore contributes to the study of TELL technologies and informs further research in this area.

A pedagogical implication is elaborated in the current study, which relates to Contribution 3. As reviewed and discussed in Chapters 2, 3, 4 and 5, most L2/EFL learners learn their L2 in a non-English-speaking environment and mainly concentrate on the written form. This is especially the case with Chinese learners. On the other hand, given the constraints of current English learning and teaching in China, Chinese learners

and teachers have very limited access to language learning and teaching facilities. Even though there are some technologies currently employed to assist learners' self-study, the drawbacks of these tools impede efficient learning on the part of EFL students. The present study, to a large extent, exposes Chinese language learners to an authentic, English speaking community for a relatively long period of six months by providing learners with natural, dialogic speech samples. It is proposed that EFL teachers and learners should be exposed to real, everyday English and have access to time-scaling technology as a principled pedagogical approach. The research work carried out by the present author in this area makes new knowledge available and contributes to the body of work on EFL pedagogy.

An innovative metric for evaluating EFL speech production is developed in the current study, which leads to Contribution 4. The aim of the tests is to evaluate the correct imitation and re-production of native English speech, especially the intelligent 'blur' and L1 prosodic patterns present in formulaic language. By reviewing the existing evaluation methodologies, such as Jones and Haywood (2004) and Kim (2006), among others (as discussed in Section 8.2.1.3), it is clear that none of them would fit the aims of this study. While a 5-point Likert scale is employed in some linguistic research, but not designed specifically to evaluate the production of phonetic blur and intonation patterns. The innovative application of a 5-point Likert scale in the present study can not only evaluate native speech production, but also distinguish the higher productive level of principled blur and the lower receptive level of accidental imitation, which adds new knowledge to the study of EFL testing methodology. This evaluation system is also validated by an independent panel of eleven L1 speakers with specialist and non-specialist knowledge of spoken English, which also brings new insights to the validation methodology and informs further study in EFL evaluation.

The application of the assets from the Dynamic Speech Corpus is also investigated in the current study, which relates to Contribution 5. Compared with other existing spoken corpora, these natural speech assets have unique advantages. As discussed in Section 5.3.3, the first advantage is their naturalness, which can better demonstrate the real, interactive features in informal L1-L1 dialogic English speech. Other corpora use broadcast materials which, by their nature, do not exemplify relaxed, informal L1-L1 communication. Their second advantage is the high standard of audio quality, which better allows the application of the slow-down technology. Another advantage embedded in the natural speech assets is that genuine, spontaneous dialogues can be separated into two ‘semi-logues’ without crossed signals, which provides the basis for acoustic analysis of cross-talking, which occurs frequently in formulaic language. Developing and tagging the natural, native-to-native speech assets adds new knowledge to the study of corpus linguistics. Exposing and immersing language learners in this virtual target language speaking community (via training and testing materials taken from the natural speech assets) inevitably advances their EFL learning. This is justified by the improvement of the test subjects in perception and production of formulaic language with NS-like quality. This investigation also contributes to EFL learning and teaching.

In the next chapter, an overall conclusion of the current research is given. Reflections on the current study and recommendations for further research are also made.

10. Conclusion

The research work undertaken by the present author is reported in the previous chapters. This current chapter provides a general conclusion in 10.1, with recommendations for further research in 10.2.

10.1 Conclusion

This section provides an overall conclusion in 10.1.1. Following that, 10.1.2 discusses the research questions which formed the basis of the current research. 10.1.3 discusses the specific contributions made by the present author in the course of this research and relates each to the research questions posed. Limitations and implications of the current study then follow in 10.1.4. Finally, reflections on the current study are considered in 10.1.5.

10.1.1 Overall Conclusion

This thesis, based on the in-depth study of formulaic language taken from real, natural, dynamic L1-L1 assets in the DSC, demonstrates the correlation of various categories of formulaic language and their phonological realisations, e.g., speed of delivery and pitch range. By her work in this area, the present author has contributed to the body of knowledge.

In order to enhance language learning through the application of technology, the slow-down facility combined with deliberately devised training materials is applied and validated in the present research. The considerable improvement of Chinese EFL learners in perception and production of the natural flow of English speech demonstrates its effectiveness.

A pedagogical methodology is also suggested in the current research and contributes to EFL pedagogy, especially for Chinese learners and teachers, i.e., exposing language learners to natural, authentic, L1-L1 speech materials, assisted by language learning technologies, rather than concentrating on the written form.

In addition, an evaluation methodology for testing the communicative linguistic competence of Chinese EFL learners in imitating and re-producing native casual English speech, especially formulaic language, is investigated in this study and has thus contributed to the body of knowledge. Qualitative and quantitative evaluation of test data and the innovative application and validation of a 5-point Likert scale is discussed and justified in this thesis as well.

All the training and testing materials used in this study are taken from the assets of the Dynamic Speech Corpus. Therefore, using this real, natural, informal English speech to expose learners to a virtual language learning environment and thus improve their skills in dealing with L1 speech is also examined in the current research.

10.1.2 Summary of Work

This thesis documents an investigation into the perception and acquisition of natural, authentic English speech by Chinese language learners using DIT's speech technologies. A review of spoken English and formulaic language, and Chinese speech and English learning and teaching in China was undertaken first (Chapters 2 and 3). These chapters suggested that classroom-based Chinese learners of English needed to increase their exposure to real, informal, native English speech so as to tackle an interactive, dialogic English speaking environment, especially in dealing with formulaic language. In Chapter 2, there was also a gap identified by the present author that the phonological

characteristics of formulaic sequences themselves could be more important than their lexical realisations.

On the basis of a review of the literature, especially given the limited English learning and teaching situations in China, in Chapters 4 and 5, access to effective TELL technologies was proposed and discussed, in particular, the slow-down technology and assets from Dynamic Speech Corpus. As a result of the review and discussions, several areas of research were identified. The rationale for designing the study, and the procedures and methodology for its implementation were described in Chapter 6, leading to a statement of the research questions which defined the scope of work for this thesis:

RQ1: What influence do speed of delivery and pitch range have on the categorisation of formulaic language?

RQ2: Does the slow-down facility, coupled with suitable training materials, improve Chinese EFL learners' ability to perceive and produce formulaic language with NS-like quality?

RQ3: Can real natural English speech be incorporated into EFL pedagogy?

RQ4: How can EFL learning be evaluated in the context of natural speech?

RQ5: Can assets from the DIT Dynamic Speech Corpus improve EFL learners' facility with L1 speech?

The first research question was investigated in the analysis of the Dynamic Speech Corpus assets (Chapter 7), in which an inverse relationship between speed of delivery and pitch range was identified and a prototype formulaic language typology was recommended as a starting point for further research.

The second and fifth research questions were evaluated during a case study of the effectiveness of the slow-down technology with Chinese learners of English in bridging the intonational gap between Mandarin prosody and English prosody, specifically incorporating the use of the Dynamic Speech Corpus assets (Chapter 8).

The third research question was also demonstrated in the case study (Chapter 8), in which the considerable improvement in Chinese EFL learners' language competence illustrated the pedagogical effectiveness of using natural, informal, dialogic English speech with technological support.

The fourth research question of an innovative evaluation methodology based on qualitative and quantitative analysis of test data was also verified during the case study (Chapter 8), which proposed a new method of assessing the production of intelligent, native-like blur.

10.1.3 Contributions to the Body of Knowledge

This thesis has presented several original contributions to the field of formulaic language, and EFL learning and teaching by:

1. Definition of the relationship between formulaicity and prosody. A correlation of various categories of formulaic language and their phonological realisations, e.g., speed of delivery and pitch range, is proposed to fill the gap in the current literature. This is investigated in Chapter 7 and relates to RQ1.
2. Clear demonstration of the effectiveness and acceptability of the use of a Speech Slow-down facility (discussed in Chapter 8). The study demonstrates the effectiveness of the chosen assets in promoting phoneme recognition and

pronunciation. It also improves EFL learners' communicative linguistic ability in imitation and re-production of native-like acoustic 'blur' and intonation patterns of formulaic sequences. This relates to RQ2.

3. Implications of an EFL pedagogy, specifically for Chinese learners and teachers, by exposing students to real, interactive speech samples occurring in everyday life, aided by language learning technologies, rather than adherence to the internalised, idealised forms influenced by the written language. This is investigated in Chapter 8 and relates to RQ3.
4. Development of an evaluation methodology for testing EFL speech production, which incorporates an assessment of the messy 'blur' of rapid speech and communicative intonation patterns, rather than being based on judgements using citation phonemes. The test results obtained from 100 participants from different levels of linguistic competence and language learning background indicate a certain level of both qualitative and quantitative validation. This is discussed in Chapter 8 and relates to RQ4.
5. Investigation of the application of assets from DIT's language learning resource, i.e., Dynamic Speech Corpus. Using real, natural English dialogic speech can expose EFL learners to a natural, authentic, target language speaking environment, which therefore facilitates their acquisition of pragmatic, communicative language competence. This is discussed in Chapter 8 and relates to RQ5.

10.1.4 Limitations and Implications of the Current Study

Due to time scale limitations inherent in the scope of a PhD, even though several areas of contributions have been made in the current thesis, there are also some limitations which to some extent restrict its reliability and further development.

Firstly, as discussed earlier, given that the Dynamic Speech Corpus was still under development, this resulted in a relatively limited amount of data being available for use and analysis, which in turn consequently affects the reliability of the current research.

For example, the analysis of formulaic language and its phonological realisations of speed of delivery and pitch range are mainly based on four L1-L1 dialogues. The value setting for various bands of speed and pitch range is very limited, obtained from only five speakers, who are also from different age bands – one teenager, one in her 70s, and three in the 40-50 age bracket. There is also a gender imbalance between the five speakers. Therefore, for increased reliability of the analysis of the salient features of formulaic sequences, more dialogues are needed. The choice of speakers should also be increased, with more speakers, rather than concentrating on one main speaker. The age and gender of the speakers should also be kept in balance, since speakers at different ages and genders have different physical values of speed and pitch range. While more examples would help to confirm the hypothesis and be helpful in finding other possible significant features, within the scope of the thesis, it can be argued that the results are indeed valid.

Besides, given the very heavy learning and teaching syllabus imposed on the three Chinese English teachers and 100 participants, the time spent on the training session is not balanced between the three groups, which might influence the testing performance and validation. Therefore, a fully controlled training session is recommended for any further investigation in this area.

In addition, there are three Chinese teachers who represented the present author and implemented the training sessions in China. They are all qualified teachers of English, and all have teaching experience. However, their spoken competence might be different, which could also influence the training results and therefore the testing performance. If the training sessions could be controlled by one teacher, that would diminish this possibility.

There is also another factor which was not considered when designing the tests. The present author used a population of convenience. However, in GpB (senior cycle junior secondary school), there are twenty students in the test group and also twenty in the control group. However, the gender between two groups is not balanced. There are six males and fourteen females in the test group, while there are eight males and twelve females in the control group. In Test 2, the Test Group achieved a 19% higher score than the Control Group, in which the gender factor might influence the testing results. Therefore, gender balance needs to be considered in further research in this area.

10.1.5 Reflections on the Current Study

During the long journey of the current research, the present author has made several contributions to the body of knowledge. All five contributions are important to English language and EFL pedagogy. Among them, two are substantial. First is the analysis of formulaic language. The prototype analysis on the phonological realisations of various categorisations of formulaic language brings significant prosodic features into the perception and production of formulaic language, which fills in a major gap in the literature.

The investigation of the application of slow-down technology also adds new knowledge to the body of EFL technology-assisted language learning. Through the present research, the effectiveness and acceptability of the slow-down facility in helping Chinese EFL learners with imitation and re-production of a natural flow of L1-L1 English speech, rather than recognising and re-producing its citation form, is clearly demonstrated.

Based on the literature, and to the best of the present author's knowledge, the current research is valid. The methodology applied for undertaking this study is correct and suitable.

Tests and training sessions carried out in China were necessary and also successful. However, some suggestions can be considered for further study. First is on the balance between test participants and test samples. In Test 2, in order to obtain a wide range of testing data and also to fit into the PhD time scale, the same number of participants is used as in Test 1, while the number of testing samples is reduced. Further tests would be more informative if more testing samples are chosen and tested, while keeping the same number of participants.

Secondly, there are different levels of improvement in Test 2 between the three groups. Compared with GpA, which had the pressure of written College English Exams, and GpC which was younger and had less knowledge of English, GpB was preparing for their transfer exam (including listening and written exams) to senior secondary school and had more motivation to participate actively in the training activities, which may be why GpB makes significant improvement compared to the other two groups. In future research, more tests should be designed and carried out concentrating on GpB for more

accurate validation of the application of technology in EFL language learning and acquisition.

In addition, since the present author could only monitor the training sessions from a distance, detailed training processes were out of her control. Therefore, if possible, full control of the training sessions by the author is recommended for further study.

10.2 Recommendation for Further Research

Work carried out in this thesis has produced significant improvement in our understanding of the areas of formulaic language, and EFL learning and teaching. Several major areas of investigation were undertaken, and some are now considered for further research.

10.2.1 Further Analysis of the Relationship between Formulaicity and Prosody

Initial findings from the current research on the correlation between various formulaic language and their phonological realisations, i.e., speed of delivery and pitch range, are proving to be of interest to fellow researchers at linguistic conferences. More data are needed to determine whether the initial indications are valid and generalisable.

The investigation of the link between various locations within fluent speech of formulaic language and their speed of delivery is at present inconclusive. There might be some other factors which need to be considered, as discussed in Section 7.2.4. The current analysis is based on what has been defined as the flow unit, which is a relatively small unit for segmenting and tagging. Formulaic language is not only marked using pitch range, but also can be marked by being part of an ‘expressive envelope’, which

can override a unit of greater extent than the flow unit, i.e., phrase or sentence. Therefore, further analysis is needed based on different hierarchical units.

In addition to the speed of delivery and pitch range, more work needs to be undertaken, based on the analysis of natural interactive dialogue, to identify other phonological features related to the categorisation of formulaic language, such as pause, and stress.

10.2.2 Analysis Related to the Realisation of Formulaic Language

Apart from the analysis of phonological features which are related to the realisation of formulaic language, another area of interest is whether there are any other factors which promote the use of formulaic sequences, for example, the relationship between the speakers and the roles of each speaker in the conversation. In the analysis of the speech corpus assets, an unbalanced use of formulaic language is shown between two speakers which is likely to be linked to their different roles in the conversation. More data are needed to verify this hypothesis.

Another observation concerns the quick change of pitch and/or tack which sometimes accompanies the use of formulaic language. In one recording, the speaker suddenly changes her expression from ‘going down (and up)’ to ‘up and down’. Whether this results from visualisation in the brain, or whether it is somehow constrained by the syntactic structure, because ‘up and down’ is more acceptable than ‘down and up’, remains an open question and requires further investigation.

10.2.3 Fuller Investigation of Flow Unit and Other Units

The flow unit is defined as an ideal unit for use in the current study. The speed of delivery of each flow unit is measured for using in a tag-based search. However, the

relationship of the flow unit and its phonological realisations, e.g., prominence, tone, key, termination, requires further study for further refinement of the flow unit.

Another issue which needs further investigation is to compare and contrast the hierarchy of units related to the flow unit. Compared with the tone unit as reviewed in Section 2.2.2, the flow unit is characterised as being a relatively short unit at 3-4 words, thus whether grouping flow units together would form a tone unit and whether there is a correlation between them will be left open for further research.

10.2.4 Further Development of DIT's Speech Technologies for the Acquisition and Use of L1 Speech by EFL Learners

As discussed earlier, all training and testing materials used in the current study were taken from the Dynamic Speech Corpus assets, comprising real, interactive L1-L1 speech features. By exposing them to these authentic, dynamic language learning materials, combined with the use of slow-down technology, Chinese EFL learners advanced their linguistic competence in coping with the natural flow of English speech.

Given that the DSC is not yet available to language learners, some more work is needed to progress its development. When these functions are available, a dedicated language learning toolset will be of use to and benefit language learners and also contribute to the development of TELL.

10.2.5 Further Analysis of Construction of Dialogic Con-fluence

Based on the tagging of speech attributes occurring in real, natural, dynamic native-to-native conversation, some correlations have been initially analysed (as outlined in Appendix 1) on how speakers realise their speech intention and turn construction; for example, what the relationship is between the use of discourse markers, formulaic

language and non-linguistic vocalisations, and the expression of speaker intention; how phonetic, prosodic, discourse and disfluent features contribute to the realisation of turn interaction between speakers.

Limited by the available data, the analysis undertaken by the present author in this area is merely at a preliminary stage. Its aim is to highlight the natural, dynamic interaction in authentic, informal L1-L1 English speech, and how the realisation of speaker intention and turn construction affects prosody, which extends the research beyond formulaic sequences. A wider range of data is needed for further analysis of attributes which contribute towards dialogic con-fluence. This could be an area of extensive further research.

References

- Adolphs, S. & Durow, V. (2004) Social-cultural integration and the development of formulaic sequences. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.107-26.
- Aitchison, J. (1994) Understanding words. In: G. Brown K. Malmkjær A. Pollitt & J. Williams, eds. 1994. *Language and understanding*. Oxford: Oxford University Press.
- Altenberg, B. (1990) Speech as linear composition. In: G. Caie K. Haastrup A.L. Jakobsen J.E. Nielsen J. Sevaldsen H. Specht & A. Zettersten, eds. 1990. *Proceedings from the Fourth Nordic Conference for English studies, vol. 1*. Copenhagen: Dept of English, University of Copenhagen, pp.133-43.
- Altenberg, B. (1998) On the phraseology of spoken English: the evidence of recurrent word-combinations. In: A. Cowie, ed. 1998. *Phraseology: theory, analysis and applications*. Oxford: Clarendon Press, pp.101-22.
- Anderson, J.R. (1983) *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.
- Anonymous internet source, (2007) *Evaluation of English classroom teaching activity by countryside senior secondary school students*. [Online] Available at: www.xzedu.net.cn/ktoblog/UploadFiles/2007-11/1114349564.doc [Accessed 24 December 2008].
- Aston, G. (1988) *Learning Comity: an approach to the description and pedagogy of interactional speech*. Bologna: Editrice Club Bologna.
- Aston, G. (1995) Say 'thank you': some pragmatic constraints in conversational closings. *Applied Linguistics*, 16 (1), pp.57-86.
- Bacalu, C. & Delmonte, R. (1999) *Prosodic modeling for speech recognition*. [Online] Available at: <http://www.parlaritaliano.it/parlare.eng/index.html> [Accessed 30 March 2007].
- Bachman, L.F. (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bamgbose, A. (1998) Torn between the norms: innovations in world Englishes. *World Englishes*, 17 (1), pp.1-14.
- Bao, T.R. (2004) *Situations and thoughts on primary school English teaching in China*. [Online] Available at: <http://www.elishar.com/?action=text|admin|4|cn|123|75> [Accessed 03 February 2009].
- Beattie, G. (1983) *Talk: an analysis of speech and non-verbal behaviour in conversation*. Milton Keynes: Open University Press.

- Becker, J. (1975) The phrasal lexicon. In: B. Nash-Webber & R. Schank, eds. 1975. *Theoretical issues in natural language processing 1*. Cambridge, Mass.: Bolt, Beranek, and Newman.
- Bishop, H. (2004) The effect of typographic salience on the look up and comprehension of unknown formulaic sequences. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.227-48.
- Bolinger, D.L. (1975) *Aspects of Language*. 2nd ed. New York: Harcourt Brace Jovanovich.
- Bolinger, D.L. (1976) Meaning and memory. *Forum Linguisticum*, 1, pp.1-14.
- Boomer, D.S. & Laver, J.D.M. (1968) Slips of the tongue. *British Journal of Disorders of Communication*, 3, pp.2-12.
- Boulton, A. (2009) *Discussion on attribute tagging in the Dynamic Speech Corpus [Conversation]* (Personal communication, 12 September 2009).
- Bradley, P. (2008) *Discussion on testing methodology [Conversation]* (Personal communication, 5 March 2008).
- Brazil, D. (1994) *Pronunciation for Advanced Learners of English*. student's book. Cambridge: Cambridge University Press.
- Brazil, D. (1995) *Describing English Language*. Oxford: Oxford University Press.
- Brazil, D. (1996) Designing an integrated pronunciation course. *Speak Out!*, 17, pp.5-9.
- Brazil, D. (1997) *The Communicative Value of Intonation in English*. 2nd ed. Cambridge: Cambridge University Press.
- Brett, D. (2004) Computer generated feedback on vowel production by learners of English as a second language. *ReCALL*, 16 (1), pp.103-13.
- Brown, G. (1990) *Listening to Spoken English*. 2nd ed. London: Longman.
- Brown, G. (1995) *Speakers, Listeners and Communication: explorations in discourse analysis*. Cambridge: Cambridge University Press.
- Brown, G. & Yule, G. (1983a) *Teaching the Spoken Language: an approach based on the analysis of conversational English*. Cambridge: Cambridge University Press.
- Brown, G. & Yule, G. (1983b) *Discourse Analysis*. Cambridge: Cambridge University Press.
- Brown, P. & Levinson, S. (1987) *Politeness: some universals in language usage*. Cambridge: Cambridge University Press.
- Brown, R. (1973) *A First Language*. London: Allen & Unwin.

- Burnaby, B. & Sun, Y.L. (1989) Chinese teachers' views of western language teaching: context informs paradigms. *TESOL Quarterly*, 23 (2), pp.219-38.
- Bush, M.D. & Terry, R.M. (1997) *Technology-enhanced Language Learning*. Lincolnwood, IL, USA: National Textbook Company.
- Bybee, J. (1998) The emergent lexicon. *Chicago Linguistics Society*, 34, pp.421-35.
- Campbell, D. (2007) *Discussion on the usage of the slowed-down speed* [Conversation] (Personal communication, 23 April 2005).
- Campbell, D. (2008) *Discussion on measurement of speed of delivery* [Conversation] (Personal communication, 9 July 2008).
- Campbell, D. Meinardi, M. Richardson, B. Wang, Y. & McDonnell, C. (2009) DIT's Dynamic Speech Corpus and dialogic fluency. In: EUROCALL (European Association for Computer-Assisted Language Learning), *EuroCALL 2009*. Gandia, Spain 9-12 September 2009.
- Campbell, D. Wang, Y. Kelleher, J. Meinardi, M. & Richardson, B. (2006) DIT Speech Corpus. In: IVACS (Inter-Varietal Applied Corpus Studies), *3rd IVACS international conference: language at the interface*. Nottingham, UK 23-24 June 2006.
- Campbell, D. Wang, Y. & McDonnell, C. (2007) A prototype speech corpus. In: EUROCALL (European Association for Computer-Assisted Language Learning), *EuroCALL 2007: mastering multimedia: teaching languages through technology*. Coleraine, UK 5-8 September 2007.
- Campbell, D. Wang, Y. & McDonnell, C. (2008) FS \neq FS: formulaicity and prosody. In: BAAL (British Association for Applied Linguistics), *41st BAAL annual meeting*. Swansea, UK 11-13 September 2008.
- Campbell, R. & Wales, R. (1970) The study of language acquisition. In: J. Lyons, ed. 1970. *New horizons in linguistics*. Harmondsworth: Penguin, pp.242-60.
- Carter, R. & McCarthy, M. (1988) *Vocabulary and Language Teaching*. London: Longman.
- Carter, R. & McCarthy, M. (1995) Grammar and the spoken language. *Applied Linguistics*, 16 (2), pp.141-58.
- Carter, R. & McCarthy, M. (1997) *Exploring Spoken English*. Cambridge: Cambridge University Press.
- Carter, R. & McCarthy, M. (2006) *Cambridge Grammar of English: a comprehensive guide spoken and written English grammar and usage*. Cambridge: Cambridge University Press.
- Cauldwell, R. (2000) *Grasping the nettle: the importance of perception work in listening comprehension*. [Online] Available at: http://www.developingteachers.com/articles_tchtraining/perception1_richard.htm [Accessed 05 March 2006].

- Cauldwell, R. (2001) *The functional irrhythmicality of spontaneous speech: a discourse view of speech rhythms*. [Online] Available at: <http://www.solki.jyu.fi/apples/> [Accessed 07 September 2005].
- Cauldwell, R. (2002) *Phonology for listening: relishing the messy*. [Online] Available at: http://www.speechinaction.net/SPARC_Richards_publications.htm [Accessed 22 February 2006].
- Cauldwell, R. (2003) *The two-sides rule in teaching listening and pronunciation*. [Online] Available at: http://www.developingteachers.com/articles_tchtraining/two_sides1_richard.htm [Accessed 10 October 2005].
- Cauldwell, R. (2005) *Bricking up and streaming down: two approaches to naturalness in pronunciation materials*. [Online] (Updated 16 June 2005) Available at: <http://www.phon.ucl.ac.uk/home/johnm/ptlc2005/pdf/ptlcp36.pdf> [Accessed 09 April 2006].
- Chafe, W.L. (1979) The flow of thought and the flow of language. In: T. Givon, ed. 1979. *Syntax and semantics volume 12: discourse and syntax*. New York: Academic Press.
- Chan, A.Y.W. & Li, D.C.S. (2000) English and Cantonese phonology in contrast: explaining Cantonese ESL learners' English pronunciation problems. *Language, Culture and Curriculum*, 13 (1), pp.67-85.
- Chang, J. (2001) Chinese speakers. In: M. Swan & B. Smith, eds. 2001. *Learner English: a teacher's guide to interference and other problems*. 2nd ed. Cambridge: Cambridge University Press, pp.310-24.
- Chao, Y.R. (1933) Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology, Academia Sinica*, 4 (2), pp.121-34.
- Chao, Y.R. (1934) The non-uniqueness of phonemic solutions of phonetic systems. In: M. Joos, ed. 1957. *Readings in linguistics I*. Chicago: University of Chicago Press, pp.38-54.
- Chao, Y.R. (1968) *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chaudron, C. (1982) Vocabulary elaboration in teachers' speech to L2 learners. *Studies in Second Language Acquisition*, 4, pp.170-80.
- Chaudron, C. & Richards, J. (1986) The effect of discourse markers on the comprehension of lectures. *Applied Linguistics*, 7, pp.113-27.
- Chen, M.Y. (2000) *Tone Sandhi: patterns across Chinese dialects*. Cambridge: Cambridge University Press.
- Chen, P. (1999) *Modern Chinese: history and sociolinguistics*. Cambridge: Cambridge University Press.

- Chen, ZH.Q. (2008) *Discussion on English listening and speaking teaching in primary and secondary school in Anshan city* [Phone call] (Personal communication, 12 June 2008).
- Cheng, CH.CH. (1973) *A Synchronic Phonology of Mandarin Chinese*. The Hague: Mouton.
- China channel network, (2006) *Chinese IELTS students rank the last in the world in oral test*. [Online] Available at: <http://news.eastday.com/eastday/node81741/node81762/node136275/userobject1ai2052931.html> [Accessed 21 February 2007].
- China education and research network, (1994) *Hanyu Pinyin Syllabus*. [Online] Available at: <http://www.edu.cn/20011114/3009777.shtml> [Accessed 05 March 2007].
- China resources network, (2006) *A report of general teaching situations in primary school*. [Online] Available at: <http://www.ceozg.com/book/30882.html> [Accessed 19 November 2008].
- Cicourel, A.V. (1973) *Cognitive Society*. Harmondsworth: Penguin Books.
- Code, C. (1994) Speech automatism production in aphasia. *Journal of Neurolinguistics*, 8 (2), pp.135-48.
- Cook, G. (1989) *Discourse*. Oxford: Oxford University Press.
- Cortazzi, M. & Jin, L. (1996) Cultures of learning: language classrooms in China. In: H. Coleman, ed. 1996. *Society and the language classroom*. Cambridge: Cambridge University Press, pp.169-206.
- Coulmas, F. (1981) Introduction: conversational routine. In: F. Coulmas, ed. 1981. *Conversational routine*. The Hague: Mouton, pp.1-17.
- Cowie, A.P. (1988) Stable and creative aspects of vocabulary use. In: R. Carter & M. McCarthy, eds. 1988. *Vocabulary and language teaching*. London: Longman, pp.126-39.
- Cowie, A.P. (1992) Multiword lexical units and communicative language teaching. In: P.J.L. Arnaud & H. Bejoint, eds. 1992. *Vocabulary and applied linguistics*. Basingstoke: Macmillan, pp.1-12.
- Cowie, A.P. & Howarth, P. (1996) Phraseological competence and written proficiency. In: G. Blue & R. Mitchell, eds. 1996. *Language and education*. Clevedon: Multilingual Matters, pp.80-93.
- Crabbe, D. (2003) The quality of language learning opportunities. *TESOL Quarterly*, 37 (1), pp.9-34.
- Crick, F.H.C. (1979) Thinking about the brain. *Scientific American*, 241, pp.219-32.
- Cruttenden, A. (1981) Item-learning and system-learning. *Journal of Psycholinguistic Research*, 10 (1), pp.79-88.

- Cruttenden, A. (1997) *Intonation*. 2nd ed. Cambridge: Cambridge University Press.
- Crystal, D. (1996) The past, present and future of English rhythm. *Speak Out!*, 18, pp.8-13.
- Crystal, D. (2002) *The English Language: a guided tour of the language*. 2nd ed. London: Penguin Books.
- Crystal, D. (2003a) *The Cambridge Encyclopedia of the English Language*. 2nd ed. Cambridge: Cambridge University Press.
- Crystal, D. (2003b) *A Dictionary of Linguistics & Phonetics*. 5th ed. Oxford: Blackwell Publishing Ltd.
- Cullen, C. Vaughan, B. Kousidis, S. Wang, Y. MacDonnail, C. & Campbell, D. (2006) Generation of high quality audio natural emotional speech corpus using task based mood induction. In: *InSciT2006 Conference*, Merida, Spain 25-28 October 2006.
- Dahlmann, I. & Adolphs, S. (2007) *Pauses as an indicator of psycholinguistically valid multi-word expressions (MWEs)?* [Online] Available at: <http://www.acl.ldc.upenn.edu> [Assessed 01 September 2008].
- Dalton, C. & Seidlhofer, B. (1994) *Pronunciation*. Oxford: Oxford University Press.
- de Beaugrande, R. (1980) *Text, Discourse and Process*. London: Longman.
- DeCarrico, J.S. & Nattinger, J.R. (1988) Lexical phrases for the comprehension of academic lectures. *English for Specific Purposes*, 7, pp.91-102.
- DeCock, S. Granger, S. Leech, G. & McEnery, T. (1998) An automated approach to the phrasicon of EFL learners. In: S. Granger, ed. 1998. *Learner English on computer*. London, New York: Addison Wesley Longman, pp.67-79.
- Deese, J. (1984) *Thought into Speech: the psychology of a language*. Englewood Cliffs, NJ: Prentice-Hall.
- DeFrancis, J. (1986) *The Chinese Language: fact and fantasy*. Honolulu: University of Hawaii Press.
- Dennett, D.C. (1990) Making sense of ourselves. In: W.C. Lycan, ed. 1990. *Mind and cognition*. Oxford: Basil Blackwell, pp.184-99.
- Derwing, T.M. (1990) Speech rate is no simple matter: rate adjustment and NS-NNS communicative success. *Studies in Second Language Acquisition*, 12, pp.303-13.
- Derwing, T.M. & Munro, M.J. (1997) Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, 19, pp.1-16.
- Dörnyei, Z. (2003) *Questionnaires in Second Language Research*. Mahwah NJ: Lawrence Erlbaum.
- Dörnyei, Z. Durow, V. & Zahran, K. (2004) Individual differences and their effects on formulaic sequence acquisition. In: N. Schmitt, ed. 2004. *Formulaic sequences*:

acquisition, processing and use. Amsterdam: John Benjamins Publishing Company, pp.87-106.

Doyle, D. (2006) Connecting theory and practice to enhance teaching and learning encounters with Chinese students. In: ACELS (The Advisory Council for English Language Schools), *MATSDA/ACELS conference: what we know and what we do: connecting theory and practice in materials development for language teaching*. Dublin, Ireland 21-22 January 2006.

Duanmu, S. (2002) *The Phonology of Standard Chinese*. New York: Oxford University Press.

Ellis, N.C. (1996) Sequencing in SLA: phonological memory, chunking and points of order. *Studies in Second Language Acquisition*, 18, pp.91-126.

Ellis, R. (1994) *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

English education network, (2007a) *New mind, new idea, new pedagogy: English course-book for nine-year compulsory education*. [Online] Available at: <http://wz.yingyu.com.cn/2007/2007-01-25/20070125134854.shtml> [Accessed 24 November 2008].

English education network, (2007b) *Several points to be considered in English listening teaching*. [Online] Available at: <http://wz.yingyu.com.cn/2007/2007-04-23/20070423205820.shtml> [Accessed 22 November 2008].

Erman, B. & Warren, B. (2000) The idiom principle and the open-choice principle. *Text*, 20, pp.29-62.

Feng, S. (2003) *Investigation and strategy for spoken English in secondary school*. [Online] Available at: <http://www.mt56.com/Explore/ArticleShow.asp?ArticleID=5846> [Accessed 23 February 2007].

Fernando, C. & Flavell, R. (1981) *On Idiom: critical views and perspectives* (Exeter Linguistic Studies 5). Exeter: University of Exeter.

Field, J. (1999) 'Bottom-up' and 'top-down'. *ELT Journal*, 53 (4), pp.338-39.

Field, J. (2000) 'Not waving but drowning': a reply to Tony Ridgway. *ELT Journal*, 54 (2), pp.186-95.

Field, J. (2003a) *Psycholinguistics: a resource book for students*. London: Routledge.

Field, J. (2003b) Promoting perception: lexical segmentation in L2 listening. *ELT Journal*, 57 (4), pp.325-34.

Field, J. (2004) *Psycholinguistics: the key concepts*. London: Routledge.

Field, J. (2008) *Discussion on testing methodology* [Conversation] (Personal communication, 11 September 2008).

Firth, J.R. (1964) *The Tongues of Men and Speech*. London: Oxford University Press.

- Flavell, L. & Flavell, R. (1992) *Dictionary of Idioms and their Origins*. London: Kyle Cathie.
- FLUENT network, (2008) [Online] Available at: http://fluent.dmc.dit.ie/acl_users/credentials_cookie_auth/login_form?came_from=http%3A//fluent.dmc.dit.ie/front-page/document_view [Accessed 12 March 2009].
- Free dictionary network, (2009) [Online] Available at: <http://www.thefreedictionary.com/attribute> [Accessed 08 June 2009].
- Fu, L.Y. (2008) *On factors of influencing English listening teaching and learning*. [Online] Available at: <http://www.studa.net/yingyu/080528/17491148.html> [Accessed 10 December 2008].
- Furnham, A. & Bochner, S. (1989) *Culture Shock: psychological reactions to unfamiliar environments*. London: Routledge.
- Garvey, C. (1977) Play with language & speech. In: S. Ervin-Tripp & C. Mitchell-Kernan, eds. 1977. *Child discourse*. New York: Academic Press, pp.27-47.
- Gass, S. & Varonis, E.M. (1984) The effect of familiarity on the comprehensibility of nonnative speech. *Language Learning*, 34, pp.65-89.
- Gimson, A.C. (2001) *Gimson's Pronunciation of English*. 6th ed. Revised by A. Cruttenden. London: Edward Arnold.
- Givon, T. (1989) *Mind, Code and Context*. Hillsdale, NJ: Lawrence Erlbaum.
- Goh, C. (1997) Metacognitive awareness and second language listeners. *ELT Journal*, 51 (4), pp.361-69.
- Goldman-Eisler, F. (1968) *Psycholinguistics: experiments in spontaneous speech*. London: Academic Press.
- Golinkoff, R.M. & Hirsh-Pasek, K. (1995) Reinterpreting children's sentence comprehension: toward a new framework. In: P. Fletcher & B. MacWhinney, eds. 1995. *The handbook of child language*. Oxford: Blackwell, pp.430-61.
- Gorokhova, S. (2008) Decomposition of idioms during sentence production: evidence from speech errors. In: BAAL (British Association for Applied Linguistics), 41st BAAL annual meeting. Swansea, UK 11-13 September 2008.
- Graddol, D. (1997) *The Future of English?: a guide to forecasting the popularity of the English language in the 21 century*. The British Council.
- Graddol, D. (2006) *English Next: why global English may mean the end of 'English as a Foreign Language'*. The British Council.
- Gu, Q. (2003) *Bridging belief gaps in ELT teacher education in cross-cultural settings*. [Online] Available at: http://www.developingteachers.com/articles_tchtraining/bridgingpf_qing.htm [Accessed 27 October 2005].

- Hakuta, K. (1976) A case study of a Japanese child learning English as a second language. *Language Learning*, 26 (2), pp.321-51.
- Halliday, M.A.K. (1967) Notes on transitivity and theme in English: part 2. *Journal of Linguistics*, 3, pp.199-244.
- Halliday, M.A.K. (1970) *A Course in Spoken English: introduction*. Oxford: Oxford University Press.
- Han, J.F. (2007) *Drawing attention to pronunciation teaching*. [Online] (Updated 01 March 2007) Available at: <http://jcjykc.cersp.com/Post/ShowArticle.asp?ArticleID=6623> [Accessed 07 March 2007].
- Hasselgreen, A. (2004) *Testing the Spoken English of Young Norwegians: a study of test validity and the role of 'smallwords' in contributing to pupils' fluency*. Cambridge: Cambridge University Press.
- Hawkins, P. (1984) *Introducing Phonology*. London: Hutchinson.
- He, Y. & Jing, S. (1992) Intonations of the Beijing dialect: an experimental exploration. *Yuyan Jiaoxue Yu Yanjiu*, 2, pp.71-96.
- Heffernan, N. & Wang, S.H.D. (2007) Network-based English pronunciation self-training. In: EUROCALL (European Association for Computer-Assisted Language Learning), *EuroCALL 2007: mastering multimedia: teaching languages through technology*. Coleraine, UK 5-8 September 2007.
- Hewings, M. (2004) *Pronunciation Practice Activities: a resource book for teaching English pronunciation*. Cambridge: Cambridge University Press.
- Hickey, T. (1993) Identifying formulas in first language acquisition. *Journal of Child Language*, 20, pp.27-41.
- Hong, H.L. (2008) *Problems and strategies in senior vocational school English listening teaching*. [Online] Available at: <http://www.govyi.com/lunwen/2008/200810/266083.shtml> [Accessed 17 December 2008].
- House, J. (1996) Developing pragmatic fluency in English as a foreign language. *Studies in Second Language Acquisition*, 18, pp.225-52.
- Howarth, P. (1996) *Phraseology in English Academic Writing: some implications for language learning and dictionary making* [lexicographica, Series maior 75]. Tübingen: Max Niemeyer.
- Howarth, P. (1998) Phraseology and second language proficiency. *Applied Linguistics*, 19 (1), pp.24-44.
- Howie, J. (1976) *An Acoustic Study of Mandarin Tones and Vowels*. Cambridge: Cambridge University Press.

- Hu, Y.W. (2005) *Current problems and measures on secondary school English phonetics teaching*. [Online] (Updated 03 October 2006) Available at: <http://english.cersp.com/ketang/200603/468.html> [Accessed 03 February 2007].
- Hu, ZH.Q. (2008) *A report of primary school English teaching in Jianwei County*. [Online] Available at: <http://www.qwedu.net/user1/bluesky/archives/2008/1948.html> [Accessed 29 December 2008].
- Huang, J. & Hatch, E.M. (1978) A Chinese child's acquisition of English. In: E.M. Hatch, ed. 1978. *Second language acquisition: a book of readings*. Rowley, MA: Newbury House, pp.118-31.
- Huang, L.SH. (2010) The potential influence of L1 (Chinese) on L2 (English) communication. *ELT Journal*, 64 (2), pp.155-64.
- Huang, X.L. (2007) *Application of multi-media in English language learning and teaching*. [Online] Available at: <http://studa.net/jiaoxue/070127/15250610.html> [Accessed 12 February 2007].
- Hudson, J. (1998) *Perspectives on Fixedness: applied and theoretical*. Lund: Lund University Press.
- Hulstijn, J.H. (2003) Incidental and intentional learning. In: C. Doughty & M. Long, eds. 2003. *The handbook of second language acquisition*. Oxford: Blackwell, pp.349-81.
- Humaine emotion-research network, (2006) [Online] Available at: <http://emotion-research.net/toolbox/toolboxdatabase.2006-09-26.3936825048?searchterm=Belfast> [Accessed 18 October 2006].
- Hymes, D.H. (1968) The ethnography of speaking. In: J.A. Fishman, ed. 1968. *Readings in the society of language*. The Hague/Paris: Mouton, pp.99-138.
- Irujo, S. (1993) Steering clear: avoidance in the production of idioms. *International Review of Applied Linguistics in Language Teaching*, 31, (3), pp.205-19.
- Itoh, H. & Hatch, E.M. (1978) Second language acquisition: a case study. In: E.M. Hatch, ed. 1978. *Second language acquisition: a book of readings*. Rowley, MA: Newbury House, pp.76-88.
- Jenkins, J. (1998a) Rethinking phonology in teacher education. *Vienna English Working Papers*, 7 (1), pp.40-46.
- Jenkins, J. (1998b) Which pronunciation norms and models for English as an International Language? *ELT Journal*, 52 (2), pp.119-26.
- Jenkins, J. (2000) *The Phonology of English as an International Language: new models, new norms, new goals*. Oxford: Oxford University Press.
- Jenkins, J. (2005) Teaching pronunciation for English as a lingua franca: a sociopolitical perspective. In: C. Gnutzmann & F. Intemann, eds. 2005. *The globalisation of English and the English language classroom*. Tübingen: Gunter Narr Verlag, pp.145-58.

- Jespersen, O. (1922) *Language: its nature, development and origin*. New York: Macmillan.
- Jespersen, O. (1976) Living grammar. In: D.D. Bornstein, ed. 1976. *Readings in the theory of grammar*. Cambridge, MA: Winthrop Publisher, pp.82-93.
- Jiang, SH.Y. (2000) *Survey of English teaching in senior secondary school in China*. [Online] (Updated 24 November 2004) Available at: http://www.360doc.com/content/060228/00/6187_73407.html [Accessed 28 September 2008].
- Johnstone, R. (2007) Languages in the global age: opportunities and threats. In: RIA (Royal Irish Academy), Dublin, Ireland 15 October 2007.
- Jones, M. & Haywood, S. (2004) Facilitating the acquisition of formulaic sequences: an exploratory study in an EAP context. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.269-300.
- Juffs, A. (1990) Tone, syllable structure and interlanguage phonology: Chinese learners' stress errors. *IRAL*, XXVIII (2), pp.99-115.
- Kellerman, S. (1992) 'I see what you mean': the role of kinesic behaviour in listening and implications for foreign and second language learning. *Applied Linguistics*, 13 (3), pp.239-58.
- Kenworthy, J. (1987) *Teaching English Pronunciation*. London: Longman.
- Kim, I.S. (2006) Automatic speech recognition: reliability and pedagogical implications for teaching pronunciation. *Educational Technology & Society*, 9 (1), pp.322-34.
- Knowles, G. (1987) *Patterns of Spoken English*. London: Longman.
- Kousidis, S. Dorran, D. Wang, Y. Vaughan, B. Cullen, C. Campbell, D. McDonnel, C. & Coyle, E. (2008) Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues. In: ASSTA (Australian Speech Science and Technology Association) and ISCA (International Speech Communication Association), *INTERSPEECH 2008 incorporating SST 2008*. Brisbane, Australia 22-26 September 2008.
- Krashen, S. & Scarcella, R. (1978) On routines and patterns in language acquisition and performance. *Language Learning*, 28 (2), pp.283-300.
- Kuiper, K. (1996) *Smooth Talkers. Vol. 1: studies in everyday communication*. Hillsdale NJ: Lawrence Erlbaum Associates.
- Kuiper, K. (2004) Formulaic performance in conventionalized varieties of speech. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.37-54.
- Ladefoged, P. (1993) *A Course in Phonetics*. 3rd ed. Orlando: Harcourt Brace Jovanovich.

- Ladefoged, P. (2001) *A Course in Phonetics*. 4th ed. Orlando: Harcourt College Publishers.
- Lambacher, S. (2001) The taming of English vowels. In: *Proceedings of the PAC3 at JALT 2001 Convention*, Kitakyushu, Japan 22-25 November 2001.
- Lass, R. (1984) *Phonology: an introduction to basic concepts*. Cambridge: Cambridge University Press.
- Laver, J. (1970) The production of speech. In: J. Lyons, ed. 1970. *New horizons in linguistics*. Harmondsworth: Penguin, pp.53-75.
- Laver, J. (1994) *Principles of Phonetics*. Cambridge: Cambridge University Press.
- Leech, G. (1983) *Principles of Pragmatics*. London: Longman.
- Lennon, P. (2000) The lexical element in spoken second language fluency. In: H. Riggensbach, ed. 2000. *Perspectives on fluency*. Ann Arbor: University of Michigan Press, pp.25-42.
- Levelt, W.J.M. (1989) *Speaking: from intention to articulation*. Cambridge, MA: MIT.
- Li, C. & Thompson, S.A. (1981) *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.
- Li, F.K. (1966) The zero initial and the zero syllabic. *Language*, 42, pp.300-02.
- Li, J. (2007) Questions to be reconsidered in teaching pronunciation. *Sino-US English Teaching*, 4 (9), pp.35-39.
- Lin, M. (1995) A perceptual study on the domain of tones in Beijing Mandarin. *Acta Acustica*, 20 (6), pp.437-45.
- Lin, P.M.S. (2006) *The role of formulaic language in the spoken performances of native and nonnative speakers of English*. Unpublished MPhil dissertation: University of Hong Kong, Hong Kong.
- Lin, P.M.S. & Bahlmann, I. (2008) The role of prosodic features in the identification of formulaic sequences. In: BAAL (British Association for Applied Linguistics), 41st BAAL annual meeting. Swansea, UK 11-13 September 2008.
- Lin, Y.H. (2007) *The Sounds of Chinese*. Cambridge: Cambridge University Press.
- Littlewood, W. (2000) Do Asian students really want to listen and obey? *ELT Journal*, 54 (1), pp.31-36.
- Liu, D. (1998) Ethnocentrism in TESOL: teacher education and the neglected needs of international TESOL students. *ELT Journal*, 52 (1), pp.3-10.
- Locke, J. (1995) Development of the capacity for spoken language. In: P. Fletcher & B. MacWhinney, eds. 1995. *The handbook of child language*. Oxford: Blackwell, pp.278-302.

- Locke, J. (1997) A theory of neurolinguistic development. *Brain and Language*, 58, pp.265-326.
- Lord, A.B. (1960) *The Singer of Tales*. Cambridge MA: Harvard University Press.
- Lu, K. (2007) *Characteristics and strategies for large-class English teaching in primary school*. [Online] Available at: <http://www.szls.com.cn/printpage.asp?ArticleID=1081> [Accessed 07 February 2007].
- Luo, CH.P. & Wang, J. (1981) *Outline of General Phonetics*. Beijing: Shangwu Yinshuguan.
- Ma, H.Q. (2001) *Common mistakes on English phonetics teaching and how to correct them*. [Online] Available at: <http://vip.ilib.cn/A-QCode~cjxyxb200102018.html> [Accessed 15 December 2006].
- Makkai, A. (1972) *Idiom Structure in English*. The Hague: Mouton.
- Marks, J. (1999) Is stress-timing real? *ELT Journal*, 53 (3), pp.191-99.
- McCarthy, M. (1988) Some vocabulary patterns in conversation. In: R. Carter & M. McCarthy, eds. 1988. *Vocabulary and language teaching*. London: Longman, pp.181-200.
- McCarthy, M. (1990) *Vocabulary*. Oxford: Oxford University Press.
- McCarthy, M. (1991) *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.
- McCarthy, M. (1998) *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press.
- McCarthy, M. (2001) *Issues in Applied Linguistics*. Cambridge: Cambridge University Press.
- McCarthy, M. (2004) *Touchstone: from corpus to course book*. Cambridge: Cambridge University Press.
- McCarthy, M. (2006) Fluency and confluence: what fluent speakers do. In: M. McCarthy, ed. 2006. *Explorations in corpus linguistics*. Cambridge: Cambridge University Press. Ch.1.
- McCarthy, M. (2008) Profiling spoken fluency. *The Language Teacher*, 32 (7), pp.32-34.
- McCarthy, M. (in press) Spoken fluency revisited. *English Profile Journal*, 1 (1).
- McCarthy, M. & Carter, R. (2006a) Ten criteria for a spoken grammar. In: M. McCarthy, ed. 2006. *Explorations in corpus linguistics*. Cambridge: Cambridge University Press. Ch.3.

- McCarthy, M. & Carter, R. (2006b) This that and the other: multi-word clusters in spoken English as visible patterns of interaction. In: M. McCarthy, ed. 2006. *Explorations in corpus linguistics*. Cambridge: Cambridge University Press. Ch.2.
- McCarthy, M. & O’Keeffe, A. (2006) Spoken grammars, written grammars. In: ACELS (The Advisory Council for English Language Schools), *ACELS publishers’ workshop*. Dublin, Ireland 16 June 2006.
- Mehrabian, A. (1981) *Silent Messages: implicit communication of emotions and attitudes*. 2nd ed. Belmont, CA: Wadsworth.
- Meinardi, M. (2006) *The use of ‘real’ English in language learning: making authentic NS speech accessible through a novel digital Slow-down tool*. Unpublished PhD thesis: Dublin Institute of Technology, Dublin.
- Meyer, B.J.F. Brandt, D.M. & Bluth, G.J. (1980) Use of the top-level structure in text: key to reading comprehension of ninth-grade students. *Reading Research Quarterly*, 16, pp.72-103.
- Miller, G.A. (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *The Psychological Review*, 63 (2), pp.81-97.
- Ministry of Education of the People’s Republic of China, (2001) *Guidance of Ministry of Education on establishing English teaching in primary school January 18, 2001*. [Online] Available at: <http://www.moe.edu.cn/edoas/website18/level3.jsp?tablename=1162&infoid=727> [Accessed 17 October 2005].
- Ministry of Education of the People’s Republic of China, (2003a) *Basic education in China*. [Online] Available at: http://www.moe.edu.cn/english/basic_b.htm [Accessed 21 February 2007].
- Ministry of Education of the People’s Republic of China, (2003b) *Report of education statistics volumn 1 No. 26: Department of Planning Ministry of Education Feb. 27, 2003*. [Online] Available at: http://www.moe.edu.cn/english/planning_s.htm [Accessed 21 February 2007].
- Mitchell, R. & Martin, C. (1997) Rote learning, creativity and ‘understanding’ in classroom foreign language teaching. *Language Teaching Research*, 1 (1), pp.1-27.
- Moon, R. (1997) Vocabulary connections: multi-word items in English. In: N. Schmitt & M. McCarthy, eds. 1997. *Vocabulary: description, acquisition and pedagogy*. Cambridge: Cambridge University Press, pp.40-63.
- Moon, R. (1998) *Fixed Expressions and Idioms in English*. Oxford: Clarendon Press.
- Moore, R.K. (2005) Cognitive informatics: the future of spoken language processing? In: *10th International conference on speech and computer SPECOM*. Patras, Greece 17-19 October 2005.
- Mortimer, C. (1985) *Elements of Pronunciation*. Cambridge: Cambridge University Press.

- Nagy, W. (1978) Some non-idiom larger-than-word units in the lexicon. *Papers from the Parasession on the Lexicon*, Chicago Linguistic Society, pp.289-300.
- Nation, P. (2001) *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nation, P. & Coady, J. (1988) Vocabulary and reading. In: R. Carter & M. McCarthy, eds. 1988. *Vocabulary and language teaching*. London: Longman, pp.97-110.
- Nattinger, J.R. (1988) Some current trends in vocabulary teaching. In: R. Carter & M. McCarthy, eds. 1988. *Vocabulary and language teaching*. London: Longman, pp.62-82.
- Nattinger, J.R. & DeCarrico, J.S. (1992) *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Nelson, C.L. (1995) Intelligibility and world Englishes in the classroom. *World Englishes*, 14 (2), pp.273-79.
- Nelson, G. Al-Batal, M. & Echols, E. (1996) Arabic and English compliment responses: potential for pragmatic failure. *Applied Linguistics*, 17 (4), pp.411-32.
- Nelson, K. (1981) Individual differences in language development: implications for development and language. *Developmental Psychology*, 17 (2), pp.170-87.
- Ninth software network, (2006) *An approach of 'Three Transition' in classroom teaching to improve senior secondary school English teaching*. [Online] Available at: http://d9soft.com/article/274/Article11161_1.htm [Accessed 26 November 2008].
- O'Keeffe, A. McCarthy, M.J. & Carter, R.A. (2007) *From Corpus to Classroom*. Cambridge: Cambridge University Press.
- Olsen, J. (1972) *The Girls in the Office*. New York: Simon & Schuster.
- Oppenheim, N. (2000) The importance of recurrent sequences for nonnative speaker fluency and cognition. In: H. Riggensbach, ed. 2000. *Perspectives on fluency*. Ann Arbor: University of Michigan Press, pp.220-40.
- Pawley, A. (1991) How to talk cricket: on linguistic competence in a subject matter. In: R. Blust, ed. 1991. *Currents in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*. Canberra: Pacific Linguistics C-117, pp.339-68.
- Pawley, A. & Syder, F.H. (1983) Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In: J.C. Richards & R.W. Schmidt, eds. 1983. *Language and communication*. New York: Longman, pp.191-226.
- Pawley, A. & Syder, F.H. (2000) The one-clause-at-a-time hypothesis. In: H. Riggensbach, ed. 2000. *Perspectives on fluency*. Ann Arbor: University of Michigan Press, pp.163-99.
- Perkins, M.R. (1999) Productivity and formulaicity in language development. In: M. Garman C. Letts B. Richards C. Schellletter & S. Edwards, eds. 1999. *Issues in normal*

- & *disordered child language: from phonology to narrative*. Reading: University of Reading, pp.51-67.
- Peters, A.M. (1977) Language learning strategies: does the whole equal the sum of the parts? *Language*, 53 (3), pp.560-73.
- Peters, A.M. (1983) *The Units of Language Acquisition*. Cambridge: Cambridge University Press.
- Plunkett, K. (1993) Lexical segmentation and vocabulary growth in early language acquisition. *Journal of Child Language*, 20, pp.43-60.
- Primary English network, (2008) *A report on the situations of primary school English teaching*. [Online] Available at: http://www.pep.com.cn/xe/jszx/jxyj/jxdl/200806/t20080606_471667.htm [Accessed 23 October 2008].
- Purcell, E. & Suter, R. (1980) Predictors of pronunciation accuracy: a reexamination. *Language Learning*, 30, pp.271-87.
- Qin, B. & Lian, W. (2008) *Problems and measures on secondary school English phonetics teaching*. [Online] (Updated 24 September 2008) Available at: <http://www.lwxz8.com/jylw/yyjx/200809/11806.html> [Accessed 03 February 2009].
- Raupach, M. (1984) Formulae in second language speech production. In: H.W. Dechert D. Mohle & M. Raupach, eds. 1984. *Second language production*. Tübingen: Gunter Narr, pp.114-37.
- Richardson, B. (2009) *The potential use of Slow-down technology to improve pronunciation of English for international communication*. Unpublished PhD thesis: Dublin Institute of Technology, Dublin.
- Ridgway, T. (2000) Listening strategies – I beg your pardon? *ELT Journal*, 54 (2), pp.179-85.
- Roach, P. (1982) On the distinction between ‘stress-timed’ and ‘syllable-timed’ languages. In: D. Crystal, ed. 1982. *Linguistic controversies*. London: Edward Arnold.
- Roach, P. (2000) *English Phonetics and Phonology: a practical course*. 3rd ed. Cambridge: Cambridge University Press.
- Sacks, H. Schegloff, E.A. & Jefferson, G. (1974) A simplest systematics for the organisation of turn-taking for conversation. *Language*, 50 (4), pp.696-735.
- SALERO network, (2006) [Online] Available at: <http://www.salero.eu> [Accessed 12 January 2006].
- Saussure, F.De. (1966) *Course in General Linguistics*. New York: McGraw-Hill.
- Schiffrin, D. (1987) *Discourse Markers*. Cambridge: Cambridge University Press.

- Schmidt, R.W. (1983) Interaction, acculturation, and the acquisition of communicative competence: a case study of an adult. In: N. Wolfson & E. Judd, eds. 1983. *Sociolinguistics and language acquisition*. Rowley, MA: Newbury House, pp.137-74.
- Schmitt, N. & Carter, R. (2004) Formulaic sequences in action: an introduction. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.1-22.
- Schmitt, N. Dörnyei, Z. Adolphs, S. & Durow, V. (2004) Knowledge and acquisition of formulaic sequences: a longitudinal study. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.55-86.
- Schmitt, N. Grandage, S. & Adolphs, S. (2004) Are corpus-derived recurrent clusters psycholinguistically valid? In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.127-51.
- Schober, M.F. & Clark, H.H. (1989) Understanding by addressees and overhearers. *Cognitive Psychology*, 21, pp.211-32.
- Schumann, J.H. (1986) Research on the acculturation model for second language acquisition. *Journal of Multilingual and Multicultural Development*, 7, pp.379-92.
- Sebeok, T.A. (1964) The structure and content of Cheremis charms. In: D. Hymes, ed. 1964. *Language in culture and society: a reader in linguistics and anthropology*. New York: Harper & Row, pp.356-71.
- Secondary school English teaching resources network, (2006) *Change of English teachers under the new teaching content and new notion*. [Online] Available at: <http://en.ruiwen.com/news/7680.htm> [Accessed 24 November 2008].
- Selkirk, E. (1982) The syllable. In: H. van der Hulst & N. Smith, eds. 1982. *The structure of phonological representation*. Dordrecht: Foris, pp.337-83.
- Selkirk, E. & Shen, T. (1990) Prosodic domains in Shanghai Chinese. In: S. Inkelas & D. Zec, eds. 1990. *The phonology-syntax connection*. Stanford, California: CSLI, Stanford University; distributed by University of Chicago Press, pp.313-37.
- Selman, A. (2009) *Happy New Year*. [Online] Available at: http://uk.groups.yahoo.com/group/iatefl_pronsig/message/170 [Accessed 05 January 2009].
- Shen, X.N.S. (1989) *The Prosody of Mandarin Chinese*. Berkeley, CA: University of California Press.
- Shin, D. & Nation, P. (2008) Beyond single words: the most frequent collocations in spoken English. *ELT Journal*, 62 (4), pp.339-48.
- Sinclair, J.McH. (1987) Collocation: a progress report. In: R. Steele & T. Treadgold, eds. 1987. *Essays in honour of Michael Halliday*. Amsterdam: John Benjamins, pp.319-31.

- Sinclair, J.McH. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J.McH. & Renouf, A. (1988) A lexical syllabus for language learning. In: R. Carter & M. McCarthy, eds. 1988. *Vocabulary and language teaching*. London: Longman, pp.140-60.
- Smith, J. (1991) *Salutations, Felicitations, and Terminations: a study in communicative performance*. University of Canterbury: M.A.
- Smith, L.E. (1992) Spread of English and issues of intelligibility. In: B.B. Kachru, ed. 1992. *The other tongue: English across cultures*. 2nd ed. Urbana: University of Illinois Press, pp.75-90.
- Smith, L.E. & Nelson, C.L. (1985) International intelligibility of English: directions and resources. *World Englishes*, 4 (3), pp.333-42.
- Speechinaction network, (2002) [Online] Available at: <http://www.speechinaction.com> [Accessed 08 September 2007].
- Spöttl, C. & McCarthy, M. (2004) Comparing knowledge of formulaic sequences across L1, L2, L3, and L4. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.191-225.
- Steele, V. (2005) *Connected speech*. [Online] Available at: http://www.teachingenglish.org.uk/think/pron/con_speech.shtml [Accessed 06 December 2005].
- StepbyStep network, (2005) [Online] Available at: <http://www.eebbk.com/fdj.htm> [Accessed 07 July 2008].
- Stevick, E. (1976) *Memory, Meaning and Method*. Newbury House: Rowley, Massachusetts.
- Stone, D. (2008) Ratty & the thief: phonemic & phonological self-assessment as self-awareness. In: BAAL (British Association for Applied Linguistics), 41st BAAL annual meeting. Swansea, UK 11-13 September 2008.
- Strassler, J. (1982) *Idioms in English: a pragmatic analysis*. Tübingen: Gunter Narr Verlag.
- Stubbs, M. (1993) British traditions in text analysis: from Firth to Sinclair. In: M. Baker G. Francis & E. Tognini-Bonelli, eds. 1993. *Text and technology: in honour of John Sinclair*. Philadelphia/Amsterdam: John Benjamins, pp.1-33.
- Sun, CH.F. (2006) *Chinese: a linguistic introduction*. Cambridge: Cambridge University Press.
- Sun, SH. & Zhan, J.H. (2001) Application of decoding theory in English listening teaching. *Teaching and Training*, 8, pp.25-26.
- Tajima, K. Port, R. & Dalby, J. (1997) Effects of temporal correction on intelligibility of foreign-accented English. *Journal of Phonetics*, 25, pp.1-24.

- Tannen, D. (1989) *Talking Voices, Repetition, Dialogue and Imagery in Conversational Discourse*. Cambridge: Cambridge University Press.
- Tatham, M. & Morton, K. (2002) Computational modelling of speech production: English rhythm. In: A. Braun & H.R. Masthoff, eds. 2002. *Phonetics and its applications*. Stuttgart: Steiner, pp.383-405.
- Tatham, M. & Morton, K. (2004) *Expression in Speech: analysis and synthesis*. New York: Oxford University Press.
- Tatham, M. & Morton, K. (2005) *Developments in Speech Synthesis*. Chichester: John Wiley & Sons.
- Tatham, M. & Morton, K. (2006) *Speech Production and Perception*. New York: Palgrave Macmillan.
- TELLMEMORE network, (2000) [Online] Available at: <http://www.tellmemore.com/content/view/full/128> [Accessed 06 August 2008].
- Teubert, W. & Cermakova, A. (2007) *Corpus Linguistics: a short introduction*. London: Continuum.
- Thompson, I. (1991) Foreign accents revisited: the English pronunciation of Russian immigrants. *Language Learning*, 41 (2), pp.177-204.
- Towell, R. Hawkins, R. & Bazergui, N. (1996) The development of fluency in advanced learners of French. *Applied Linguistics*, 17 (1), pp.84-119.
- Tsui, A. & Fullilove, J. (1998) Bottom-up or top-down processing as a discriminator of L2 listening performance. *Applied Linguistics*, 19 (4), pp.432-51.
- Turner, Y. (2001) When an unstoppable force meets an immovable object: Chinese students in the UK university system. *Chinese Students in Ireland: new opportunities, new needs, new challenges*. Papers from the ICOS Seminar 26 January 2001.
- Tyler, A.E. Jeffries, A.A. & Davies, C. E. (1988) The effect of discourse structuring devices on listener perceptions of coherence in non-native university teachers' spoken discourse. *World Englishes*, 7 (2), pp.101-10.
- Underhill, A. (1994) *Sound Foundations: living phonology*. Oxford: Heinemann.
- Underwood, G. Schmitt, N. & Galpin, A. (2004) The eyes have it: an eye-movement study into the processing of formulaic sequences. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.153-72.
- University of Cambridge ESOL Examinations network, (2009) *The Common European Framework of Reference*. [Online] Available at: <http://www.cambridgeesol.org/exams/exams-info/cefr.html> [Accessed 09 January 2009].
- Van Els, T. & De Bot, K. (1987) The role of intonation in foreign accent. *The Modern Language Journal*, 71, pp.147-55.

- Van Lancker, D. (1987) Nonpropositional speech: neurolinguistic studies. In: A.W. Ellis, ed. 1987. *Progress in the psychology of language*, vol. 3. Hillsdale, NJ: Lawrence Erlbaum, pp.49-118.
- Van Lancker, D. & Canter, G.J. (1981) Idiomatic versus literal interpretations of ditropically ambiguous sentences. *Journal of Speech and Hearing Research*, 24, pp.64-69.
- Van Lancker, D. Canter, G.J. & Terbeek, D. (1981) Disambiguation of ditropic sentences: acoustic and phonetic cues. *Journal of Speech and Hearing Research*, 24, pp.330-35.
- Varonis, E.M. & Gass, S.M. (1985) Miscommunication in native/nonnative conversation. *Language in Society*, 14 (3), pp.327-43.
- Walsh, S. (2009) CA and CL: lifelong partners or just good friends? In: IVACS (Inter-Varietal Applied Corpus Studies), *IVACS annual symposium: corpora and interaction*. Edinburgh, UK 22 January 2009.
- Walker, R. (2009) *Discussion on positive role of 'blur'* [Email] (Personal communication, 15 October 2008).
- Walter, C. (2008) Phonology in L2 reading: not an optional extra. In: BAAL (British Association for Applied Linguistics), *41st BAAL annual meeting*. Swansea, UK 11-13 September 2008.
- Wang, Y. (2005) A research on English pronunciation teaching in Chinese context. *Journal of Kunming University of Science and Technology*, 5 (4), pp.75-79.
- Warren, P. Elgort, I. & Crabbe, D. (2009) Comprehensibility and prosody ratings for pronunciation software development. *Language Learning & Technology*, 13 (3), pp.87-102.
- Watkins, C. (1992) The comparison of formulaic sequences. In: E.C. Polome & W. Winter, eds. 1992. *Reconstructing languages & cultures*. The Hague: Mouton, pp.391-418.
- Watts, R.J. (1989) Taking the pitcher to the 'well': native speakers' perception of their use of discourse markers in conversation. *Journal of Pragmatics*, 13, pp.203-37.
- Weinert, R. (1995) The role of formulaic language in second language acquisition: a review. *Applied Linguistics*, 16 (2), pp.180-205.
- Widdowson, H.G. (2007) *Discourse Analysis*. Oxford: Oxford University Press.
- Willett, J. (1995) Becoming first graders in an L2: an ethnographic study of L2 socialization. *TESOL Quarterly*, 29 (3), pp.473-503.
- Williams, M. (1988) Language taught for meetings and language used for meetings: is there anything in common? *Applied Linguistics*, 9, pp.45-58.
- Willis, D. (1990) *The Lexical Syllabus*. London: Harper Collins.

- Wilson, M. (2003) Discovery listening – improving perceptual processing. *ELT Journal*, 57 (4), pp.335-44.
- Wingstedt, M. & Schulman, R. (1984) Comprehension of foreign accents. In: W. Dressler, ed. 1984. *Phonologica 1984: proceedings of the 5th International Phonology Meeting*. Cambridge: Cambridge University Press, pp.339-44.
- Witt, S. & Young, S. (1998) Computer-assisted pronunciation teaching based on automatic speech recognition. In: S. Jager J. Nerbonne & A. van Essen, eds. 1998. *Language teaching & language technology*. Lisse, The Netherlands: Swets & Zeitlinger, pp.25-35.
- Wong Fillmore, L. (1979) Individual differences in second language acquisition. In: C.J. Fillmore D. Kempler & S.Y.W. Wang, eds. 1979. *Individual differences in language ability and language behaviour*. New York: Academic Press, pp.203-28.
- Wood, D. (2002) Formulaic language in acquisition and production: implications for teaching. *TESL Canada Journal*, 20, pp.1-15.
- Wood, M.M. (1986) *A Definition of Idiom*. Bloomington, IN: Indiana University Linguistics Club.
- Wray, A. (1992) *The Focusing Hypothesis: the theory of left hemisphere lateralized language re-examined*. Amsterdam: John Benjamins.
- Wray, A. (2002) *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Wray, A. (2004) ‘Here’s one I prepared earlier’. In: N. Schmitt, ed. 2004. *Formulaic sequences: acquisition, processing and use*. Amsterdam: John Benjamins Publishing Company, pp.249-68.
- Wray, A. (2008) *Discussion on formulaic language [Conversation]* (Personal communication, 11 September 2008).
- Wray, A. & Perkins, M.R. (2000) The functions of formulaic language: an integrated model. *Language & Communication*, 20, pp.1-28.
- Xiao, J. (2004) *An investigation on secondary school English teaching in Nanchang City*. [Online] Available at: <http://www.pep.com.cn/200503/ca652210.htm> [Accessed 07 February 2007].
- Xu, G. & Ferguson, G. (2008) Investigating pedagogic tasks: the impact of participant structure and task type on oral language performance. In: BAAL (British Association for Applied Linguistics), 41st BAAL annual meeting. Swansea, UK 11-13 September 2008.
- Xu, K. (2004) *Large class or small class: what is English learning for?* [Online] Available at: <http://www.nanfangdaily.com.cn/southnews/xy/rmht/200403030009.asp> [Accessed 07 February 2007].
- Xu, M.Y. (2007) *Investigation of strategies in secondary school English listening teaching*. [Online] (Updated 15 August 2007) Available at:

<http://english.cersp.com/ketang/lilunshijian/200708/2687.html> [Accessed 21 December 2008].

Yang, W. (2008) *Analysis of problems and measures on senior secondary school English phonetics teaching*. [Online] Available at: <http://studa.net/jiaoxue/080529/14254524.html> [Accessed 29 November 2008].

Yip, M. (1980) *The Tonal Phonology of Chinese*. MIT PhD dissertation. Published 1990, New York: Garland Publishing.

Yorio, C.A. (1980) Conventionalized language forms and the development of communicative competence. *TESOL Quarterly*, 14 (4), pp.433-42.

Yule, G. (1985) *The Study of Language*. Cambridge: Cambridge University Press.

Zielinski, B.W. (2008) The listener: no longer the silent partner in reduced intelligibility. *System*, 36, pp.69-84.

Zou, Q. (2008) *Application of communicative English teaching method in English teaching*. [Online] Available at: http://www.gztzy.edu.cn/teach.asp?page=1&F_1stClassid=29&F_ID=974 [Accessed 25 November 2008].

Perception and Acquisition of Natural Authentic English Speech for Chinese Learners Using DIT's Speech Technologies

Yi Wang, BA

PhD Thesis

Dublin Institute of Technology

College of Arts & Tourism

Supervisors:

Mr. Dermot F. Campbell

Prof. Ciaran McDonnell

Prof. Michael McCarthy

September 2010

Volume 2/2

Appendix 1: Further Potential of Authentic English Speech for Teaching EFL Learners

1. Introduction

The main research work carried out by the present author within the scope of her PhD framework is reported in Volume 1. The first issue was looking at EFL for Chinese students and what they should learn. The work on formulaic sequences was part of this. This led to the research discussed in Chapters 7 and 8 on the benefit of access to natural, authentic speech for Chinese learners, and a methodology for teaching them English using new assets and technology, as well as assessing them. It is in this context that further interactive L1-L1 dialogue materials are suggested here based on research into dialogic con-fluence (McCarthy uses the term ‘confluence’, as mentioned in Section 2.3.3, Volume 1), as the work on the Dynamic Speech Corpus develops. It is a natural extension of the previous work. The methodology described in Chapter 8 could be used to teach and evaluate the benefits of such materials as described in this section.

In this section, further analysis is undertaken based on the tagging of real, dynamic English speech, i.e., speech attributes. It first examines some significant interactive features, such as the use of formulaic language, discourse functions, phonetic features, prosodic patterns, disfluent features, non-linguistic vocalisations, and their contributions to the realisation of speaker intention and turn construction. A discussion on how this analysis might be helpful for the investigation of achieving dialogic con-fluence for non-L1 language learners of English now follows.

The extended research outlined in this section adds new insights to the body of knowledge in the area of dialogic speech. The natural authentic speech assets, the slow-down technology and the tagging 'Attribute Tree' (detailed in Section 2.1 below) all combine to provide the present author with an opportunity that has not been available before.

2. Analysis of Dialogic Con-fluence

2.1 Tagging for Speech Attributes

As reported in Section 7.2.2, Volume 1, the initial work in tagging of DSC audio assets is mainly carried out by the present author in the areas of speed of delivery and pitch range, which correlate with the phonological realisations of formulaic language. Apart from these two aspects, the key work for tagging the speech is to tag for speech attributes. A simple definition of attribute is 'a quality or characteristic inherent in or ascribed to someone or something'.

In contrast to tagging for speed of delivery and pitch range, which mainly occurs at the flow unit level, tagging for speech attributes is based on three hierarchical levels: turn, flow sequence (which is also called communicative sequence) and flow unit (or prosodic sequence). A definition given by Campbell, et al. (2009) is that 'turn' means a coherent, interactive sequence spoken by a speaker indicating possession of the 'floor'. A 'flow sequence' refers to the content of the message at the semantic level. A 'flow unit' is more closely related to physical speech production, bordered/terminated either by a pause or a sudden change in pitch direction (as discussed in Section 7.2.2,

Volume 1). Based on this working definition, orthographic text may be segmented into these three different levels, as shown in Table 36.

Orthographic Text	Flow Unit	Flow Sequence	Turn
That is	FUA3	FSA3	TA2
[inhales]	FUA4		
That is right	FUA5		
yeah	FUA6		
She was out there for	FUA7	FSA4	
[inhales]	FUA8		
for a while	FUA9		
She is not in there any more	FUA10	FSA5	
cause (because) she has moved on	FUA11		
into a	FUA12		
into a playschool	FUA13		
[ahm]	FUA14	FSA6	
in	FUA15		
closer	FUA16		
to us	FUA17		
but yeah	FUA18	FSA7	
the Wee Care was so ni..	FUA19		
[inhales]	FUA20		

Table 36: Segmentation of orthographic text into three hierarchical levels

Table 36 demonstrates part of the transcription of Speaker A – Harry – in the dialogue between Harry and Sue. This is Harry’s second turn, marked as TA2. Within this turn, there are five flow sequences, FSA3 – FSA7. In each of these flow sequences, there are several flow units, such as FUA3, FUA4, FUA5 and FUA6 within flow sequence FSA3. The whole dialogue runs 6 minutes and 33 seconds. As shown in Table 37, 81 turns are tagged. 35 of them are in Harry’s ‘semi-logue’ (half of an interactive dialogue), and 46 in Sue’s. Below the turn level, there are 128 flow sequences in Harry’s semi-logue, and

91 in Sue's. These flow sequences are then further segmented into flow units. Within the total of 656 flow units, Harry delivers 410, while Sue delivers 246. For full version of tagging, see below.

	Turn	Flow Sequence	Flow Unit
Harry	35	128	410
Sue	46	91	246
Total	81	219	656

Table 37: Segmentation of Harry-Sue dialogue based on three hierarchical levels

In order to tag for speech attributes, an Attribute Tree was first established by DSC researchers based on research undertaken by Sacks, et al. (1974), Cook (1989) and O'Keeffe, et al. (2007), among others. There are more than 160 attributes altogether across the three hierarchical levels, and each level includes several speech values. As shown in Figure 46, 'topic' is tagged at the turn level; 'speaker intention' and 'turn construction' are mainly labelled at the flow sequence level; while 'phonetic features', 'formulaic sequences', 'discourse function', etc., are only tagged at the flow unit level. Each of these attributes can have one or several values realised in the DSC as drop-down values, as shown in Figure 47. For example, speaker intention can be expressed in various ways, and is currently tagged with 34 different values, such as 'clarification', 'establish consensus', 'express opinion', 'inform/make statement', 'revise opinion/recast'.

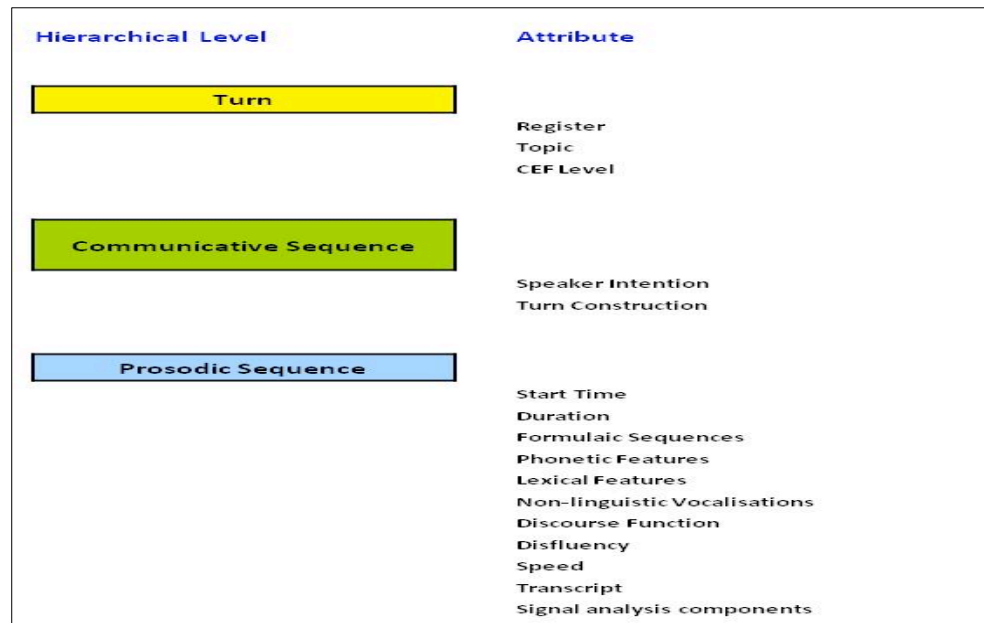


Figure 46: Attribute Tree

Hierarchical Level	Attribute	Drop-down value
Communicative Sequence	Speaker Intention	advice - give/receive apologise appreciate claim clarification command compare compliment congratulate deny establish consensus exclaim exhort express dislikes express likes express opinion face saving face threatening forbid inform / make statement invite invite a response offer permit predict promise question request revise opinion / recast suggest thank warn welcome UNSPECIFIED

Figure 47: Attribute values under the tag 'speaker intention'

The main principle for tagging is to tag those speech features which are significant for non-L1 language learners of English. That is to say, only those linguistic characteristics ‘noticed’ by an L1 listener or proficient language learner are tagged (similar to ‘noticing’ in Nation, 2001, as mentioned in Section 2.4.4, Volume 1). As Alex Boulton points out, L1 language users may tag different features, compared with non-L1 language learners (Personal communication, 12 September 2009). Since the present author is a non-L1 English user, she therefore has a relative advantage in tagging the speech attributes from the language learners’ point of view. Among all the attribute values applied to the main characteristics of speech in the dialogue Harry-Sue, 1,230 speech acts under the headings of ‘speaker intention’, ‘turn construction’, ‘formulaic sequences’, ‘phonetic features’, ‘discourse function’, and other linguistic and non-linguistic features are tagged as laid out in Table 38 below. For further details, see below. Potential links between these attribute values are analysed in detail in the next section.

	Harry	Sue
Speaker Intention	121	80
Turn Construction	31	33
Formulaic Sequences	135	85
Phonetic Features	122	130
Non-linguistic Vocalisations	80	42
Discourse Function	185	102
Disfluency	59	25
Total	733	497

Table 38: Attribute values in the dialogue Harry-Sue

2.2 Significant Features in Natural Authentic Interactive English Speech

By tagging for the attribute values of speech, some significant linguistic characteristics are further examined in this section. For example, 2.2.1 looks at the link between the realisation of speaker intention, and such features as disfluency, discourse function, the use of formulaic sequences and non-linguistic vocalisations. 2.2.2 discusses aspects of how a speaker implements turn constructions (i.e., keep a turn or ‘grab’ a turn), how the phonetic features, prosodic patterns, disfluent features and the use of discourse markers function in the realisation of turn behaviour. The detailed analysis of these features is only made possible by the application of DIT’s recording techniques to the uniquely natural, dialogic interchanges contained in the DSC.

2.2.1 Speaker Intention and its Realisations

This sub-section discusses how a speaker expresses his intention by using discourse markers, formulaic language, phonetic changes, disfluent features and non-linguistic vocalisations in real, informal L1-L1 English speech.

As shown in Table 38, there are 121 attribute values tagged under the heading of ‘speaker intention’ in Harry’s semi-logue. 34 of them, representing 28% of his tagged features, express the value ‘inform/make statement’, as shown in Table 39. While in Sue’s semi-logue, the majority of values expressing speaker intention is to ‘establish consensus’, which equals 43% of 80 tags. As for the tagging for discourse function, 45 out of 185 tags are used by Harry for buying time, which is the majority (24%) heading under discourse function; while the majority for Sue is backchannelling, representing 22% of the expressions of discourse features. In the analysis of formulaic sequences, 81 out of 135 formulaic sequences in Harry’s semi-logue (60%) are ‘integrated chunks’

(the parts are inseparable, not just a high probability that they are found together). While in Sue's part of the dialogue, the majority 43 tags out of 85 (51%) are under the heading 'structural chunk', which includes both grammatical and syntactical formulations.

	Speaker Intention	Discourse Function	Formulaic Sequences	Non-linguistic Vocalisations	Disfluency
Harry (1267 words)	inform/make statement 34 (28%)	buy time 45 (24%)	integrated chunk 81 (60%)	80	59
Sue (766 words)	establish consensus 34 (43%)	backchannelling 22 (22%)	structural chunk 43 (51%)	42	25

Table 39: Potential correlation between speaker intention and its realisations

Based on the analysis in Table 39, there is initial evidence in this unscripted, natural dialogue, of a link between the expression of speaker intention, the realisation of discourse function and the use of formulaic sequences. Harry is the main speaker in the Harry-Sue dialogue. Most of time, he holds the floor and delivers information. In order to keep his speech flow smooth and fluent, he employs many ways of buying time, such as by using formulaic sequences (e.g., 'you know', 'kind of', 'I guess', 'it was just'), and lexical fillers, for example, 'yeah', 'like', 'well'. While, Sue in contrast, being informed by Harry, mainly expresses consensus. Therefore, 22% of the discourse function tags in her semi-logue consists of backchannelling, such as 'right, right, right', 'yeah, yeah, yeah', 'yeah, OK', or repetition of Harry's expressions.

As for the use of formulaic sequences, as discussed above, due to the fact that the main function in Harry's discourse is to buy time, he uses more integrated chunks, which have a dialogic function, in order to maintain his speech flow and achieve dynamic fluency. This is different from Sue in that the majority of her formulaic sequences are structural chunks. Integrated chunks in Sue's formulaic sequences category rank third, with 13% of 85 tags, because she does not need them to help her buy time.

In order to buy time in making statements, Harry also uses more non-linguistic vocalisations, such as inhalations and filled pauses (i.e., 'ahm', 'erm', 'mm'). There are in total 80 non-linguistic vocalisations tagged in Harry's semi-logue. Compared with Harry, Sue only uses half of this number to help her make comments, i.e., backchannelling, and establish consensus. Apart from more non-linguistic vocalisations employed by Harry, there are also 59 disfluent acts tagged in his semi-logue (1267 words), which is more than twice that of Sue's total (see Table 39 above). He repeats his utterances or delivers his utterances disjointedly in order to buy time, or under time pressure, he does not finish sentences and even stammers.

The analysis above based on the study of informal, authentic, interactive native-to-native speech indicates a potential correlation between speaker intention, discourse function, the use of formulaic sequences, non-linguistic vocalisations and disfluency. The expression of discourse features and the use of formulaic language and non-linguistic vocalisations, and disfluent delivery, are determined by the speaker's intention. Speaker intention changes from flow sequence to flow sequence, and is realised by the use of various discourse functions, formulaic language and some non-

linguistic vocalisations. In the dialogue Harry-Sue, Harry seems to be more ‘disfluent’ than Sue. This results from his speaking intention and reflects his attempt to achieve dialogic fluency.

The tagging and analysis were based on and facilitated by the breaking down of the conversation into flow units and flow sequences. As discussed earlier in Section 2.1, real, interactive features in natural L1-L1 speech, such as turn construction, formulaic language, discourse function and disfluency, are better accessed and understood on the three hierarchical levels used in the DSC. This is the basic principle in tagging for speech attributes, which is also applied in the following analysis of the relationship between turn construction and its realisations in phonetic features, prosodic patterns, discourse function and disfluency.

2.2.2 Turn Construction and its Realisations

This sub-section deals with the analysis of turn construction, and the correlation between different turn behaviours and their realisations in phonetic features, prosodic patterns, discourse functions and disfluency characteristics. Five main turn behaviours are detailed, such as latching, turn attempt, turn grabbing, turn losing and turn keeping.

As with speaker intention, turn construction is tagged at the level of flow sequence. It is a dynamic process, which demonstrates the interactive flow of turns between speakers. Normal (i.e., un-contended) turn hand-over is not marked. In the tagging for marked turn construction in the dialogue Harry-Sue, there are in total 64 marked turn

behaviours which are identified and analysed, as shown in Table 40. The means of realisation of these turn constructions are also laid out below.

	No. of Tags	Means of Realisation
Latching	1	1. intonation pattern characterised by drawing out of phonetic item (100%)
Turn Attempt	8	1. speed increase only (by syllable reduction, elision, contraction or co-articulation); or additionally, discourse function (i.e., opener or lexical filler); or additionally, non-linguistic vocalisations (i.e., filled pause); or additionally, disfluency (i.e., incomplete) (37.5%) 2. volume increase only; or additionally, weakened vowel; or additionally, discourse function (i.e., opener) (37.5%) 3. intonation pattern (12.5%) 4. non-linguistic vocalisations (e.g., inhales) (12.5%)
Turn Grabbing	25	1. intonation pattern only (by drawing out of phonetic item, weakened vowel, contraction, co-articulation or hyper-articulated consonant); or additionally, disfluency (i.e., stammer or repetition) (44%) 2. volume increase only (by drawing out of phonetic item); or additionally, speed increase (by contraction, co-articulation, weakened vowel or IPA-n); or additionally, disfluency (i.e., incomplete or repetition) (40%)

		3. speed increase only (by weakened vowel, drawing out of phonetic item, elision, contraction or co-articulation); or additionally, disfluency (i.e., incomplete) (16%)
Turn Losing	22	<p>1. intonation pattern only (by drawing out of phonetic item or hyper-articulated consonant); or additionally, non-linguistic vocalisations (i.e., inhales, filled pause or laughs); or additionally, disfluency (i.e., incomplete, repetition, stammer or disjointed) (95%)</p> <p>2. disfluency only (i.e., incomplete) (5%)</p>
Turn Keeping	8	<p>1. volume increase only; or additionally, drawing out of phonetic item; or additionally, disfluency (i.e., stammer) (50%)</p> <p>2. volume increase plus speed increase (by syllable reduction, contraction, co-articulation or weakened vowel); or additionally, discourse function (i.e., lexical filler) (25%)</p> <p>3. speed increase only by weakened vowel and contraction (12.5%)</p> <p>4. intonation pattern plus discourse function (i.e., lexical filler) (12.5%)</p>
Total	64	

Table 40: Turn constructions and their realisations

a) latching

As shown in Table 40, there is only one turn behaviour marked as ‘latching’ (as mentioned in Section 2.2.4, Volume 1), which occurs when Sue is finishing Harry’s utterance by drawing out one of the phonemes and following Harry’s intonation pattern. This is one of the features of L1-L1 English speech, in which a listener is not a passive receiver of information, but rather he actively listens to and decodes the speaker’s utterance, therefore he can interpret the speaker’s intention in advance and actually finish his utterance.

b) turn attempt

There are eight turn behaviours which indicate that the speaker tries to cut in to take the floor, but does not succeed, which is marked as ‘turn attempt’. The main means to realise this linguistic feature is to increase speed of delivery and increase volume, in order to get the speaker’s attention, ‘jump the queue’, and to force a turn. Speed and volume increase are mostly realised by some phonetic features, such as syllable reduction, elision, contraction, co-articulation or weakened vowel. At the same time as increasing speed and volume, the speaker also applies other linguistic or non-linguistic features. For example, turn opener (i.e., ‘well’, ‘yeah’, ‘so’, ‘oh’) and lexical filler, such as ‘yeah’, ‘ah’, ‘so’, ‘see’, ‘like’, ‘well yeah’. Speakers use these small lexical words as ‘starters’ to help them tune into the conversation and prepare to grab a turn. Apart from linguistic discourse markers, speakers also use some non-linguistic vocalisations, i.e., filled pause, such as ‘erm’, ‘ahm’ to help them to get ready to start their utterances. Some disfluent characteristics also emerge in tagging for attribute values, e.g., the value ‘incomplete’. Since speakers speed up their utterances so as to gain a new turn, sometimes incomplete sentences are inevitable. 6 out of 8 ‘turn attempt’ tags in this

conversation are mainly realised by the increase of speed of delivery and volume. Apart from these, intonation and other non-linguistic features, such as ‘inhales’, can also be marked as an attempt at turn grabbing. A clear example given by Sue shows that she breathes in deeply and wants to take a turn, yet Harry does not allow her to cut in. Sue therefore gives up her attempt.

c) turn grabbing

In contrast to ‘turn attempt’, when the speaker actually gets his way and successfully grabs his turn, it is tagged as ‘turn grabbing’. There are 25 turn attributes marked in the dialogue Harry-Sue, 9 of them (29% of his turn constructions) are realised by Harry, and 16 by Sue which constitutes 48% of her turn behaviours. 44% of the total 25 attributes are realised mainly by intonation pattern via drawing out of phonetic item, weakened vowel, contraction, co-articulation or hyper-articulated consonant. 40% of them are realised by volume increase or, additionally, increased speed involved with phonetic changes. The remaining 16% of turn grabbing behaviours are mainly realised by speeding up the utterance. Apart from the factors of volume and speed of delivery, some disfluent features are common characteristics in most of turn grabbing acts, i.e., repetition, incomplete phrases and stammering. When the speaker is trying to grab the turn, because the volume is increased and the utterance may have to be speeded up, he might not yet be ready for a well-prepared utterance, therefore, he is likely to repeat the utterance, or deliver an incomplete fragment. Some stammered utterances are also likely to occur due to the time pressure to ‘jump the queue’.

d) turn losing

In addition to the contribution of disfluent features in the construction of ‘turn grabbing’, disfluent features are also common characteristics in the realisation of ‘turn losing’. Turn losing refers to the turn behaviour when the interlocutor succeeds in grabbing the turn, and therefore the speaker has to surrender his turn reluctantly, which is distinguished from normal turn hand-over. In the dialogue Harry-Sue, 22 turn constructions are tagged as turn losing. 15 of them are used by Harry, which constitutes 48% of his total turn behaviours. While 7 of 25 are marked in Sue’s semi-logue, which is 21% of her turn constructions. Intonation patterns, or, additionally, non-linguistic vocalisations are important factors which often accompany the loss of speaker’s turn, and which makes up of 95% of the total 22 tags. When the speaker concentrates on his intonation pattern, he is likely to draw out some segmentals or produce hyper-articulated consonants, or when he applies non-linguistic vocalisations, such as ‘inhalers’, ‘laughs’, or uses lexical filler, it is very likely that there might be a gap perceived by the listener. The listener seizes the opportunity, cuts into the conversation, and therefore the speaker loses his turn. As with turn grabbing, disfluent features are also noticed in the tagging of turn losing. When the speaker repeats his utterance, stammers, delivers incomplete or disjointed utterances, it is likely that he will lose control of the turn.

e) turn keeping

In contrast to ‘turn losing’, when the speaker does not want to give up his turn and succeeds in managing to hold his ground, this turn construction is tagged as ‘turn keeping’. The realisation of turn keeping is similar to turn attempt. Since the listener speeds up his utterance and increases his volume when he attempts to grab the turn, the

speaker therefore, in order to keep the turn, also needs to increase his own volume and his delivery speeds in a way that he can overwhelm the listener. There are eight turn keeping attributes in the dialogue Harry-Sue. 50% of them are realised by volume increase only, or, additionally, by drawing out some segmentals which therefore further increases the volume. 25% of them are realised by increasing both volume and speed of delivery. Some phonetic features are involved in the realisation of speed increase, e.g., syllable reduction, contraction, co-articulation and weakened vowel. 1 out of 8 is realised by speed increase only. 1 of them is realised by intonation pattern and lexical filler, which is shown in Figure 48.

In this part of dynamic flow of speech, Sue started to cut in when Harry was delivering his Flow Unit 69 (tagged as FUA69, as shown in Figure 49). However, Harry wanted to continue his utterance and he ignored Sue's attempt. After two attempts of trying to search for a proper word – 'they are not', 'they are not', tagged as Flow Unit 72 and Flow Unit 73 – accompanied by a non-linguistic vocalisation (i.e., inhales), he still could not find the word he wanted. However, Harry did not want to give up his turn, so he used a lexical filler 'yeah' (tagged as FUA75) to hold the floor and keep Sue away so that he could buy time and then finally finished his turn at his third attempt, tagged as FUA76-FUA78.



Figure 48: Intonation pattern and lexical filler used for the realisation of turn keeping

[illegible]

Figure 49: Interactive turn behaviour between speakers

The sample above shows the dynamic ‘tension’ between two speakers in natural speech, when they are contending the turn. The flat intonation pattern of using the lexical filler ‘yeah’ (Flow Unit 75 in Harry’s speech) indicates that the speaker does not actively engage in the comments made by the ‘intruder’, he uses it as a means to distance the ‘intruder’, so that he can finish what he wants to say. This is different from the backchannelling ‘yeah, yeah, yeah’ produced by Harry later in his Flow Unit 79, and also different with the turn opener ‘yeah’ Sue used (FUB 50 in her speech as shown in Figure 49) when she started to grab the turn. The communicative value of the same lexical filler – ‘yeah’, is decided by its intonation pattern.

In the investigation of tagging for turn construction, as analysed above, there is some initial evidence for a correlation between various kinds of turn behaviours, and their different realisations of phonetic features, prosody patterns, discourse function and some disfluent features. In natural L1-L1 English speech, turn construction dynamically flows between speakers, in that the interlocutor does not passively listen to the conversation, he frequently tries to find an opportunity to grab a turn; while the speaker is always in a position to protect his floor. This is one of the significant features of real, natural native-to-native English speech as opposed to the tidied-up versions used in course-books for non-L1 language learners of English.

3. Discussion

An initial analysis of tagging for attributes in natural, authentic English speech is reported in the previous section. Some potential relationships are demonstrated: for example, speaker intention, discourse function, disfluent features, and the use of

formulaic language and non-linguistic vocalisations. Depending on different intentions the speaker wishes to utter, different discourse markers are applied to facilitate the realisation of speaker intention. This link is also supported by the use of formulaic language. Formulaic language is one of the significant characteristics in natural L1-L1 speech. However, oriented by different intentions, different categories of formulaic language are employed by speakers. Apart from the function of discourse markers and formulaic language, the use of non-linguistic vocalisations and disfluent features also contribute to the realisation of speaker intention.

Another link is between the construction of turn behaviours and their different realisations of phonetic features, prosody patterns, discourse function and disfluency. Five main, non-neutral turn constructions are tagged and analysed. Latching is mainly linked to the use of intonation patterns. Turn attempt and turn keeping are mainly realised by speed and volume increase, and intonation patterns accompanied by discourse markers. Turn grabbing is more likely to be correlated with intonation patterns, and volume and speed increase; while turn losing is mainly realised by intonation patterns. The common feature of turn grabbing and turn losing is that disfluent characteristics function importantly in the process when both the listener is grabbing the turn and the speaker is losing his turn.

Some limits in the analysis need to be pointed out. One is that, tagging for speech attributes is only at an early stage in the linguistic work undertaken for the FLUENT project. Due to the time-scale of the PhD research, there is no time for the present author to tag every flow unit in the dialogue Harry-Sue, for intensity, speed of delivery

and pitch range. Therefore, the analysis is based on listening to the audio recording and assessing the changes of speed of delivery, pitch range and intensity, rather than analysing them instrumentally. Another limitation is that the data are based only on one dialogue Harry-Sue and both speakers are American. Some taggable features might be linked more to American culture rather than have general validity; some phonetic features might occur only in American accents. Therefore more data are needed to arrive at a more accurate and reasonable analysis.

The aim of the analysis above is to demonstrate the natural, dynamic flow of speech between two speakers. This initial analysis, although limited, is informative, and hopes to throw some light on the research of dialogic con-fluence in natural, authentic L1-L1 English speech.

4. Conclusion

The research work demonstrated in this section is in addition to the present author's initial research questions and is, in effect, an exploration of how this work might be expanded. By tagging for speech attributes, some relationships regarding the realisation of speaker intention, and the use of discourse markers and formulaic language, disfluent features and also some non-linguistic vocalisations are investigated. The correlation between the construction of turn behaviours, and their different realisations in phonetic, prosodic, discourse and disfluent features is also investigated in this section.

The aim of the analysis above is to highlight some natural, interactive characteristics in real, informal native-to-native English speech, which promise to be of interest for

building up the research of dialogic con-fluence and to be informative for EFL pedagogy.

Tagging for Speech Attributes in the dialogue Harry-Sue

HARRY

Hi Martina	FUA 1	FSA 1	TA1
How are you doing?	FUA 2	FSA 2	

TA1	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSA1	Intention=Compliment; Turn=TurnGrabbing
FSA2	Intention=Compliment
FUA2	Formulaic=IntegratedChunk; Lexical=Colloquial

That is	FUA 3	FSA 3	TA2
[inhales]	FUA 4		
That is right	FUA 5		
yeah	FUA 6		
She was out there for	FUA 7	FSA 4	
[inhales]	FUA 8		
for a while	FUA 9		
She is not in there any more	FUA 10	FSA 5	
cause (because) she has moved on	FUA 11		
into a	FUA 12		
into a playschool	FUA 13		
[ahm]	FUA 14	FSA 6	
in	FUA 15		
closer	FUA		

TA2	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSA3	Intention=EstablishConsensus
FSA4	Intention=inform
FSA5	Intention=inform
FSA6	Intention=inform
FSA7	Intention=ExpressOpinion; Turn=TurnLosing
FUA4	Non-LingVoc=Inhales
FUA5	Disfluency=Repetition
FUA6	Discourse=Add-on
FUA8	Non-LingVoc=Inhales; Discourse=BuyTime
FUA9	Formulaic=Template; Disfluency=Repetition
FUA10	Formulaic=Collocation; Discourse=There'sMore
FUA11	Lexical=Colloquial
FUA12	Discourse=BuyTime

	16		
to us	FUA 17		
but yeah	FUA 18	FSA 7	
the Wee Care was so ni..	FUA 19		
[inhales]	FUA 20		

she had been there yeah	FUA 21	FSA 8	TA3
-------------------------	--------	-------	-----

FUA13	Disfluency=Repetition
FUA14	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA15	Disfluency=Disjointed
FUA16	Disfluency=Disjointed
FUA17	Disfluency=Disjointed
FUA18	Discourse=ChangeTack
FUA19	Disfluency=Incomplete
FUA20	Non-LingVoc=Inhales

TA3	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSA8	Intention=EstablishConsensus; Turn=TurnLosing
FUA21	Formulaic=StructuralChunk; Discourse=Add-on, Take-up

[inhales]	FUA 22	FSA 9	TA4
yeah	FUA 23		
[inhales]	FUA 24	FSA 10	
ye- you know	FUA 25		
she did	FUA 26		
she did like it	FUA 27		
but there were some times she	FUA 28	FSA 11	
n- now that we have her in this other place	FUA 29	FSA 12	
and we are looking back at it	FUA 30		
it seems like she was	FUA 31	FSA 13	
maybe	FUA 32		
[mmm]	FUA 33		
you know	FUA 34		
[inhales]	FUA 35		
not as happy there as she might have been	FUA 36	FSA 14	
yeah yeah sh-	FUA 37	FSA 15	TA5

TA4	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSA9	Intention=EstablishConsensus
FSA10	Intention=Inform
FSA11	Intention=ReviseOpinion
FSA12	Intention=ExpressOpinion
FSA13	Intention=ExpressOpinion
FSA14	Intention=ExpressOpinion
FUA22	Non-LingVoc=Inhales
FUA23	Phonetic=DrawnOut; Discourse=Opener
FUA24	Non-LingVoc=Inhales
FUA25	Formulaic=IntegratedChunk; Discourse=BuyTime; Disfluency=Incomplete
FUA27	Phonetic=DrawnOut; Disfluency=Repetition
FUA28	Disfluency=Incomplete
FUA29	Lexical=Colloquial, Deixis; Discourse=Aside; Disfluency=FalseStart
FUA30	Discourse=Aside
FUA31	Discourse=There'sMore
FUA32	Discourse=Hedging
FUA33	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA34	Formulaic=IntegratedChunk; Discourse=BuyTime, Repetition
FUA35	Non-LingVoc=Inhales
FUA36	Formulaic=StructuralChunk, Template

[inhales]	FUA 38	FSA 16	TA6
I think	FUA 39		
[pause]	FUA 40		
do you know what it is	FUA 41	FSA 17	

TA5	Register=Informal; Topic=Playschool, Creche; CEF=A2
-----	-----------------------------------------------------

kind of	FUA 42	FSA 18
[pause]	FUA 43	
dozy young ones	FUA 44	
[pause]	FUA 45	FSA 19
do you know	FUA 46	
[inhales]	FUA 47	
do you know what they are like	FUA 48	FSA 20
they are sort of like	FUA 49	
you you come in and they are	FUA 50	
just chatting among themselves	FUA 51	FSA 21
[em]	FUA 52	
I am not talking about the kids now	FUA 53	
[laughs]	FUA 54	FSA 22
you know	FUA 55	
if that was all	FUA 56	
that was going on that would be great	FUA 57	FSA 23
but	FUA 58	
[inhales]	FUA 59	
sometimes the creche staff are	FUA 60	FSA 24
[pause]	FUA 61	
ah you know what	FUA 62	
what do you expect	FUA 63	FSA 25
I guess	FUA 64	
they are	FUA 65	
they are	FUA 66	FSA 26
they are women in their twenties	FUA 67	
and they have other	FUA 68	
things to be doing in their lives	FUA 69	FSA 27
but	FUA 70	
[inhales]	FUA 71	
they are not	FUA 72	FSA 28
[inhales]	FUA 73	
they are not	FUA 74	
yeah	FUA 75	FSA 29
They do not seem to be that engaged with the kids	FUA 76	

FSA15	Intention=EstablishConsensus; Turn=TurnLosing
FUA37	Phonetic=DrawnOut; Discourse=Repetition; Disfluency=Incomplete

TA6	Register=Informal; Topic=Creche staff; CEF=B1
FSA16	Intention=ExpressOpinion
FSA17	Intention=ExpressOpinion
FSA18	Intention=ExpressDislikes
FSA19	Intention=ExpressOpinion
FSA20	Intention=ExpressOpinion
FSA21	Intention=ExpressDislikes
FSA22	Intention=ExpressOpinion
FSA23	Intention=ExpressOpinion
FSA24	Intention=ExpressOpinion
FSA25	Intention=ExpressOpinion
FSA26	Intention=ExpressDislikes
FSA27	Intention=EstablishConsensus; Turn=TurnKeeping
FSA28	Intention=ExpressDislikes; Turn=TurnLosing
FUA38	Non-LingVoc=Inhales
FUA39	Phonetic=Hyper-articulatedConsonant
FUA40	Non-LingVoc=Pause; Discourse=BuyTime
FUA41	Formulaic=IntegratedChunk; Phonetic=IPA-dZu, elision (phone)
FUA42	Formulaic=IntegratedChunk
FUA43	Non-LingVoc=Pause; Discourse=BuyTime
FUA44	Phonetic=DrawnOut; Lexical=Colloquial, Regional (Irish)
FUA45	Non-LingVoc=Pause
FUA46	Formulaic=IntegratedChunk; Phonetic=IPA-dZu; Discourse=Hedging, Repetition
FUA47	Non-LingVoc=Inhales
FUA48	Formulaic=IntegratedChunk; Phonetic=IPA-dZu; Discourse=Repetition (Style)
FUA49	Formulaic=IntegratedChunk
FUA50	Disfluency=Stammer
FUA51	Discourse=There'sMore
FUA52	Non-LingVoc=FilledPause
FUA53	Discourse=Aside
FUA54	Non-LingVoc=Laughs
FUA55	Formulaic=IntegratedChunk; Discourse=Hedging

to be perfectly honest with you	FUA 77		
you know that was a concern	FUA 78		

yeah yeah yeah	FUA 79	FSA 29	TA7
----------------	-----------	-------------------	------------

yeah we- we- we did not see them changing so much	FUA 80	FSA 30	TA8
but just th..tha-	FUA 81		
they did not seem to	FUA 82		
d-	FUA 83		
do much	FUA 84		
you know	FUA 85	FSA 31	
the	FUA 86		
[inhales]	FUA 87		
it was kind of	FUA 88		
[erm]	FUA 89		
the	FUA 90		
you know	FUA 91		
the kids would come in and the teachers would throw some crayons at them	FUA 92	FSA 32	
	FUA 93		
and	FUA 94		
that would be the first hour	FUA 95		
and then they might	FUA 96	FSA 33	
let them out	FUA 97		
side	FUA 98		
for	FUA 99		
[inhales]	FUA 100		
the next hour or	FUA 101		
yeah not usually for a whole hour outside	FUA 102	FSA 34	
and then they would	FUA 103	FSA 35	
[inhales]	FUA 104		
it was just	FUA 105		
trying to make their way to the mealtimes	FUA 106		
[laughs]	FUA 107		
[inhales]	FUA		

FUA57	Phonetic=Contraction, WeakenedVowel
FUA58	Discourse=There'sMore
FUA59	Non-LingVoc=Inhales
FUA60	Phonetic=Hyper-articulatedConsonant, WeakenedVowel
FUA61	Non-LingVoc=Pause; Discourse=BuyTime
FUA62	Formulaic=IntegratedChunk; Non-LingVoc=FilledPause; Discourse=BuyTime, Repetition
FUA63	Phonetic=Hyper-articulatedConsonant; Disfluency=Repetition
FUA64	Formulaic=IntegratedChunk; Discourse=BuyTime
FUA65	Discourse=BuyTime
FUA66	Discourse=BuyTime; Disfluency=Repetition
FUA67	Formulaic=Template; Discourse=BuyTime; Disfluency=Repetition
FUA69	Formulaic=StructuralChunk
FUA70	Discourse=There'sMore
FUA71	Non-LingVoc=Inhales
FUA72	Discourse=BuyTime, Take-up
FUA73	Non-LingVoc=Inhales
FUA74	Discourse=BuyTime, Take-up; Disfluency=Repetition
FUA75	Discourse=LexicalFiller, BuyTime
FUA76	Phonetic=Contraction; Discourse=SyntacticalRepair
FUA77	Formulaic=Collocation; Discourse=Aside
FUA78	Formulaic=IntegratedChunk; Discourse=Hedging, Take-up, Repetition

TA7	Register=Informal; Topic=Creche staff; CEF=A2
FSA29	Intention=EstablishConsensus
FUA79	Discourse=Backchannelling, Repetition (Emphasis)

TA8	Register=Informal; Topic=Creche staff; CEF=B1
FSA30	Intention=ExpressDislikes
FSA31	Intention=ExpressDislikes
FSA32	Intention=ExpressDislikes
FSA33	Intention=ExpressDislikes
FSA34	Intention=ReviseOpinion/Recast
FSA35	Intention=ExpressDislikes
FSA36	Intention=ReviseOpinion/Recast; Turn=TurnLosing
FUA80	Phonetic=Contraction; Disfluency=Stammer
FUA81	Formulaic=IntegratedChunk; Phonetic=Hyper-

	108		
I do not know	FUA 109	FSA 36	
maybe it is the same everywhere	FUA 110		
Martina	FUA 111		
I do not know	FUA 112		

well you know it has a good reputation	FUA 113	FSA 37	TA9
there is no doubt	FUA 114		
and the price is not bad	FUA 115	FSA 38	
you know	FUA 116		
[inhales]	FUA 117		
[erm]	FUA 118		
[inhales]	FUA 119		
but I would I would have {fthph}	FUA 120	FSA 39	
we only had Stella in part-time so	FUA 121	FSA 40	
[inhales]	FUA 122		
we sort of felt	FUA 123	FSA 41	
ah at least we are getting her out	FUA 124		
and she can go to the park	FUA 125		
after that	FUA 126		
and she can	FUA 127		
you know	FUA 128		

Oh yeah	FUA	FSA	TA10
---------	-----	------------	-------------

	articulatedConsonant; Disfluency=Stammer
FUA82	Phonetic=Contraction
FUA83	Disfluency=Incomplete
FUA84	Phonetic=DrawnOut
FUA85	Formulaic=IntegratedChunk; Discourse=BuyTime
FUA87	Non-LingVoc=Inhales
FUA88	Formulaic=IntegratedChunk; Discourse=BuyTime; Disfluency=Incomplete
FUA89	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA90	Disfluency=Repetition
FUA91	Formulaic=IntegratedChunk; Discourse=BuyTime, Repetition
FUA92	Phonetic=Contraction; Discourse=There'sMore; Disfluency=Repetition
FUA93	Phonetic=WeakenedVowel
FUA94	Phonetic=DrawnOut; Discourse=There'sMore
FUA95	Phonetic=Contraction
FUA96	Discourse=OrganisationalMarker
FUA97	Phonetic=DrawnOut; Disfluency=Disjointed
FUA98	Phonetic=DrawnOut; Disfluency=Disjointed
FUA100	Non-LingVoc=Inhales
FUA101	Disfluency=Incomplete
FUA102	Discourse=LexicalFiller
FUA103	Phonetic=Contraction; Discourse=OrganisationalMarker, Repetition; Disfluency=Incomplete
FUA104	Non-LingVoc=Inhales
FUA105	Formulaic=IntegratedChunk; Discourse=ChangeTack
FUA106	Formulaic=Collocation
FUA107	Non-LingVoc=Laughs
FUA108	Non-LingVoc=Inhales
FUA109	Formulaic=IntegratedChunk; Discourse=Softening
FUA110	Phonetic=Contraction
FUA111	Discourse=Add-on
FUA112	Discourse=Add-on, Repetition, Softening

TA9	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA37	Intention=Evaluation; Turn=TurnGrabbing
FSA38	Intention=Evaluation
FSA39	Intention=ExpressDislikes
FSA40	Intention=ExpressOpinion
FSA41	Intention=ExpressOpinion; Turn=TurnLosing
FUA113	Formulaic=IntegratedChunk; Discourse=Opener
FUA114	Formulaic=Collocation; Discourse=Add-on
FUA115	Phonetic=DrawnOut, Elision (phone)
FUA116	Formulaic=IntegratedChunk; Discourse=Repetition
FUA117	Non-LingVoc=Inhales

	129	42	
that is exact	FUA 130		
that is exactly	FUA 131		
what we had to do	FUA 132		
you know	FUA 133	FSA 43	
we just had to juggle it	FUA 134		
and I mean	FUA 135	FSA 44	
you know	FUA 136		
obviously not having as much money around the house as	FUA 137		
[inhales]	FUA 138		
as we might	FUA 139		
but it meant being home with Stella	FUA 140	FSA 45	
and you know	FUA 141	FSA 46	
she only is	FUA 142		
she is only	FUA 143		
two	FUA 144		
once you know	FUA 145		
[inhales]	FUA 146		

yeah	FUA 147	FSA 47	TA11
------	---------	---------------	-------------

[aw]	FUA 148	FSA 48	TA12
Are they going to let you	FUA 149	FSA 49	
are they going to let you do that?	FUA 150		

Ah I see	FUA 151	FSA 50	TA13
----------	---------	---------------	-------------

FUA118	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA119	Non-LingVoc=Inhales
FUA120	Formulaic=StructuralChunk; Phonetic=Contraction; Disfluency=Incomplete, Repetition
FUA121	Lexical=Colloquial; Discourse=SyntacticalRepair
FUA122	Non-LingVoc=Inhales
FUA123	Formulaic=IntegratedChunk
FUA124	Formulaic=Collocation; Non-LingVoc=FilledPause
FUA127	Disfluency=Incomplete
FUA128	Formulaic=IntegratedChunk; Discourse=Hedging, Repetition

TA10	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA42	Intention=EstablishConsensus; Turn=TurnGrabbing
FSA43	Intention=Inform/MakeStatement
FSA44	Intention=Inform/MakeStatement
FSA45	Intention=Inform/MakeStatement
FSA46	Intention=Inform/MakeStatement; Turn=TurnLosing
FUA129	Discourse=Opener
FUA130	Phonetic=Contraction
FUA131	Phonetic=Contraction; Discourse=SyntacticalRepair
FUA132	Formulaic=StructuralChunk
FUA133	Formulaic=IntegratedChunk; Discourse=Add-on, Repetition
FUA134	Formulaic=StructuralChunk; Phonetic=Elision (phone), WeakenedVowel; Discourse=Add-on
FUA135	Formulaic=IntegratedChunk; Phonetic=WeakenedVowel
FUA136	Formulaic=IntegratedChunk; Discourse=Repetition
FUA137	Formulaic=Template; Phonetic=Elision (phone), SyllableReduction
FUA138	Non-LingVoc=Inhales
FUA139	Formulaic=Template; Disfluency=Repetition
FUA141	Formulaic=IntegratedChunk; Discourse=BuyTime, Repetition
FUA143	Phonetic=Contraction; Discourse=SyntacticalRepair; Disfluency=Disjointed
FUA144	Disfluency=Disjointed
FUA145	Formulaic=IntegratedChunk; Discourse=Hedging, Repetition
FUA146	Non-LingVoc=Inhales

TA11	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA47	Intention=EstablishConsensus
FUA147	Discourse=Backchannelling

that is exactly the thing	FUA 152	FSA 51	
they really have you	FUA 153	FSA 52	
by the short and curlies	FUA 154		
do not they	FUA 155		
[inhales]	FUA 156		

yeah	FUA 157	FSA 53	TA14
------	------------	-----------	------

[inhales]	FUA 158	FSA 54	TA15
Well I tell you	FUA 159		
you know I I	FUA 160	FSA 55	
I could not say to you	FUA 161		
do not	FUA 162	FSA 56	
let her go there	FUA 163		
they are abusive	FUA 164	FSA 57	
or they are	FUA 165		
they are	FUA 166		
they are	FUA 167		
stup-	FUA 168		
I mean	FUA 169	FSA 58	
the food was good	FUA 170		
I have to say that	FUA 171		
the food was good	FUA 172		
she	FUA 173	FSA 59	
I mean	FUA 174		
she is a picky eater	FUA 175		
and she	FUA 176		
[inhales]	FUA 177		
would eat	FUA 178		
nearly every day	FUA 179		
you know	FUA 180		
or	FUA 181		
this is kind of funny	FUA 182	FSA 60	
you look in the little book	FUA 183	FSA 61	
and it would say that	FUA 184		

TA12	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA48	Turn=TurnAttempt
FSA49	Intention=Question
FUA148	Non-LingVoc=FilledPause
FUA149	Formulaic=StructuralChunk; Phonetic=Elision (phone), SyllableReduction
FUA150	Formulaic=StructuralChunk; Phonetic=Elision (phone), SyllableReduction; Discourse=Repetition (Emphasis)

TA13	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA50	Intention=EstablishConsensus; Turn=TurnAttempt
FSA51	Intention=EstablishConsensus
FSA52	Intention=EstablishConsensus
FUA151	Phonetic=Coarticulation; Discourse=LexicalFiller
FUA152	Phonetic=Coarticulation, Contraction
FUA153	Lexical=Colloquial
FUA154	Formulaic=Idiom; Phonetic=Elision (phone)
FUA155	Phonetic=Contraction; Discourse=Add-on
FUA156	Non-LingVoc=Inhales

TA14	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA53	Intention=EstablishConsensus

TA15	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA54	Intention=ExpressOpinion
FSA55	Intention=ExpressOpinion
FSA56	Intention=ExpressOpinion
FSA57	Intention=ExpressOpinion
FSA58	Intention=Evaluation
FSA59	Intention=Inform/MakeStatement
FSA60	Intention=Inform/MakeStatement
FSA61	Intention=Inform/MakeStatement
FSA62	Intention=Inform/MakeStatement; Turn=TurnKeeping
FSA63	Intention=Inform/MakeStatement
FSA64	Intention=Inform/MakeStatement
FSA65	Intention=ExpressOpinion
FSA66	Intention=ExpressOpinion

she ate her lunch	FUA 185	
and then	FUA 186	
[inhales]	FUA 187	
you would ask her	FUA 188	FSA 62
and	FUA 189	FSA 63
she would say	FUA 190	
yeah I had one	FUA 191	FSA 64
piece of pasta	FUA 192	
or something	FUA 193	
[inhales]	FUA 194	
But no I have to say	FUA 195	FSA 65
you could tell	FUA 196	FSA 66
usually	FUA 197	
like	FUA 198	
how hungry is she	FUA 199	
when you get her out of there	FUA 200	
[inhales]	FUA 201	
and	FUA 202	FSA 67
[erm]	FUA 203	
you know	FUA 204	
she ate well there	FUA 205	
[inhales]	FUA 206	
she would	FUA 207	FSA 68
love	FUA 208	
being around other kids	FUA 209	
and you know	FUA 210	FSA 69
we were worried	FUA 211	
Martina	FUA 212	
cause (because)	FUA 213	
[inhales]	FUA 214	
she had not really been around other kids	FUA 215	FSA 70
that much	FUA 216	
but wh-	FUA 217	FSA 71
she was really	FUA 218	

FSA67	Intention=Inform/MakeStatement
FSA68	Intention=Inform/MakeStatement
FSA69	Intention=ExpressOpinion
FSA70	Intention=ExpressOpinion
FSA71	Intention=Inform/MakeStatement; Turn=TurnLosing
FUA158	Non-LingVoc=Inhales
FUA159	Formulaic=IntegratedChunk; Phonetic=Elision (phone); Discourse=Opener
FUA160	Formulaic=IntegratedChunk; Discourse=BuyTime; Disfluency=Stammer
FUA161	Phonetic=Contraction; Disfluency=Repetition
FUA162	Phonetic=Contraction, DrawnOut
FUA166	Discourse=BuyTime; Disfluency=Repetition
FUA167	Discourse=BuyTime; Disfluency=Repetition
FUA168	Disfluency=Incomplete
FUA169	Formulaic=IntegratedChunk
FUA170	Phonetic=DrawnOut
FUA171	Formulaic=StructuralChunk; Discourse=Aside
FUA172	Discourse=Repetition
FUA174	Formulaic=IntegratedChunk; Discourse=Repetition
FUA175	Phonetic=Contraction; Disfluency=Repetition
FUA177	Non-LingVoc=Inhales
FUA178	Phonetic=Elision (phone)
FUA180	Formulaic=IntegratedChunk; Discourse=Repetition
FUA181	Disfluency=Incomplete
FUA182	Formulaic=IntegratedChunk; Phonetic=SyllableReduction; Discourse=ChangeTack
FUA187	Non-LingVoc=Inhales
FUA188	Phonetic=Contraction
FUA189	Phonetic=DrawnOut; Discourse=There'sMore
FUA190	Phonetic=Contraction
FUA192	Formulaic=StructuralChunk
FUA193	Formulaic=IntegratedChunk
FUA194	Non-LingVoc=Inhales
FUA195	Formulaic=StructuralChunk; Discourse=Aside, LexicalFiller
FUA198	Discourse=LexicalFiller
FUA199	Phonetic=Coarticulation
FUA200	Phonetic=WeakenedVowel

No	FUA 219	FSA 72	TA16
it was just	FUA 220		
[inhales]	FUA 221		
that	FUA 222		
she would	FUA 223		
she would get a little bored	FUA 224		
you know	FUA 225		
and	FUA 226		
[erm]	FUA 227		
[inhales]	FUA 228	FSA 73	
she went through a period	FUA 229		
where she was crying	FUA 230		
a lot of the time	FUA 231		
you know (I have)	FUA 232		
that she was there	FUA 233		
and	FUA 234		
[inhales]	FUA 235	FSA 74	
and it does not	FUA 236		
you know	FUA 237		
ther-	FUA 238		
these young ones did not like that	FUA 239	FSA 75	
that was	FUA 240		
[laughs]	FUA 241		
that was one thing	FUA 242		

it was not	FUA 243	FSA 76	TA17
no-	FUA 244		
well I mean	FUA 245	FSA 77	
I do not think	FUA 246		
teaching even enters into the equation	FUA 247		
to be perfectly honest with you	FUA 248		
I mean they have a few	FUA 249	FSA 78	
flashcards	FUA 250		
stuck up on the wall all right	FUA 251		
but it	FUA 252	FSA 79	
was not	FUA 253		

FUA201	Non-LingVoc=Inhales
FUA202	Phonetic=DrawnOut
FUA203	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA204	Formulaic=IntegratedChunk; Discourse=Repetition
FUA205	Discourse=There'sMore
FUA206	Non-LingVoc=Inhales
FUA207	Formulaic=StructuralChunk; Phonetic=Contraction
FUA208	Formulaic=StructuralChunk; Phonetic=DrawnOut
FUA209	Phonetic=SyllableReduction
FUA210	Formulaic=IntegratedChunk; Discourse=Repetition
FUA211	Discourse=Aside
FUA213	Lexical=Colloquial
FUA214	Non-LingVoc=Inhales
FUA215	Phonetic=Contraction;
FUA217	Discourse=BuyTime; Disfluency=Incomplete
FUA218	Disfluency=Incomplete

TA16	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA72	Intention=Inform/MakeStatement
FSA73	Intention=Inform/MakeStatement
FSA74	Intention=Inform/MakeStatement
FSA75	Intention=Inform/MakeStatement; Turn=TurnLosing
FUA219	Phonetic=DrawnOut; Discourse=There'sMore
FUA220	Formulaic=IntegratedChunk; Phonetic=Hyper-articulatedConsonant; Discourse=BuyTime
FUA221	Non-LingVoc=Inhales; Discourse=BuyTime
FUA223	Discourse=SyntacticalRepair
FUA224	Phonetic=Contraction; Disfluency=Repetition
FUA225	Formulaic=IntegratedChunk
FUA226	Phonetic=DrawnOut; Discourse=BuyTime
FUA227	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA228	Non-LingVoc=Inhales
FUA229	Formulaic=SyntacticalChunk; Phonetic=WeakenedVowel
FUA230	Phonetic=DrawnOut
FUA231	Formulaic=SyntacticalChunk
FUA232	Formulaic=IntegratedChunk; Phonetic=Contraction; Discourse=Repetition
FUA233	Discourse=SyntacticalRepair

[inhalas]	FUA 254	
but I do not really care about it	FUA 255	FSA 80
I do not really believe the little kids are about	FUA 256	FSA 81
learning to read and write	FUA 257	
or anything like that	FUA 258	
you know	FUA 259	
or if they are going to do that	FUA 260	FSA 82
they can do that with us at home	FUA 261	
[inhalas]	FUA 262	

yeah	FUA 263	FSA 83	TA18
but	FUA 264		
[inhalas]	FUA 265	FSA 84	
just in terms of	FUA 266		
[erm]	FUA 267		
feeling valued and engaged with	FUA 268		
by the	FUA 269		
staff	FUA 270		
I do not know	FUA 271	FSA 85	

[inhalas]	FUA 272	FSA 86	TA19
I thought that was a little bit lacking	FUA 273		
you know	FUA 274		
[inhalas]	FUA 275		
but you know	FUA 276	FSA 87	
those are	FUA 277		
precious little thing	FUA 278		
[laughs]	FUA 279	FSA 88	
you know	FUA 280		
with your own kids	FUA 281	FSA 89	
you are always going to look and say	FUA 282		
why does not everyone else	FUA 283		
treat them as precious as	FUA 284		
us	FUA 285		
yeah	FUA 286		

FUA235	Non-LingVoc=Inhalas
FUA236	Phonetic=Contraction, Elision(phone); Disfluency=Repetition
FUA237	Formulaic=IntegratedChunk; Discourse=BuyTime, Repetition
FUA238	Disfluency=Incompleted
FUA239	Phonetic=Contraction, WeakenedVowel; Lexical=Colloquial, Regional (Irish); Discourse=SyntacticalRepair
FUA241	Non-LingVoc=Laughs
FUA242	Disfluency=Repetition

TA17	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA76	Intention=Evaluation; Turn=TurnGrabbing
FSA77	Intention=ReviseOpinion/Recast
FSA78	Intention=ExpressOpinion
FSA79	Intention=ExpressOpinion
FSA80	Intention=ExpressOpinion; Turn=TurnKeeping
FSA81	Intention=ExpressOpinion
FSA82	Intention=ExpressOpinion; Turn=TurnLosing
FUA243	Phonetic=Coarticulation, Contraction; Discourse=Take-up
FUA244	Disfluency=Incompleted
FUA245	Formulaic=IntegratedChunk; Discourse=LexicalFiller
FUA246	Phonetic=Contraction
FUA247	Formulaic=StructuralChunk
FUA248	Formulaic=Collocation; Phonetic=Coarticulation; Discourse=Aside
FUA249	Formulaic=IntegratedChunk; Discourse=Repetition
FUA251	Formulaic=Collocation; Phonetic=Coarticulation, Elision (phone), WeakenedVowel
FUA253	Phonetic=Contraction
FUA254	Non-LingVoc=Inhalas
FUA255	Phonetic=Contraction, SyllableReduction, Plus VolumeIncreasing; Discourse=ChangeTack
FUA256	Phonetic=Contraction; Discourse=Repetition
FUA258	Formulaic=IntegratedChunk
FUA259	Formulaic=IntegratedChunk
FUA260	Formulaic=StructuralChunk; Phonetic=SyllableReduction, WeakenedVowel
FUA261	Phonetic=WeakenedVowel
FUA262	Non-LingVoc=Inhalas

[inhales]	FUA 287		
-----------	------------	--	--

yeah	FUA 288	FSA 90	TA20
you got another option you are saying	FUA 289	FSA 91	

[hmm]	FUA 290	FSA 92	TA21
-------	------------	-----------	------

an- aw	FUA 291	FSA 93	TA22
yeah yeah	FUA 292		

[inhales]	FUA 293	FSA 94	TA23
I would say	FUA 294		
that is a	FUA 295		
that is a tough one	FUA 296		
I mean	FUA 297	FSA 95	
ab-	FUA 298		
you know	FUA 299		
and ah	FUA 300		
[inhales]	FUA 301		
I got to say	FUA 302	FSA 96	
I know from	FUA 303		
friends of mine	FUA 304		
that the au-pair experience can be	FUA 305		
great	FUA 306		
I do not know	FUA 307	FSA 97	
I mean	FUA 308		

TA18	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA83	Intention=FaceSaving; Turn=TurnGrabbing
FSA84	Intention=ExpressOpinion
FSA85	Intention=ReviseOpinion/Recast; Turn=TurnLosing
FUA263	Phonetic=DrawnOut; Discourse=There'sMore
FUA265	Non-LingVoc=Inhales
FUA266	Formulaic=StructuralChunk; Phonetic=DrawnOut, Hyper-articulatedConsonant; Discourse=BuyTime
FUA267	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA269	Phonetic=DrawnOut; Discourse=BuyTime
FUA271	Formulaic=IntegratedChunk, Discourse=Softening

TA19	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSA86	Intention=EstablishConsensus; Turn=TurnGrabbing
FSA87	Intention=ExpressOpinion
FSA88	Intention=ExpressOpinion
FSA89	Intention=ExpressOpinion; Turn=TurnKeeping
FUA272	Non-LingVoc=Inhales
FUA273	Formulaic=StructuralChunk; Discourse=Take-up
FUA274	Formulaic=IntegratedChunk
FUA275	Non-LingVoc=Inhales
FUA276	Formulaic=IntegratedChunk; Discourse=Hedging, Repetition
FUA279	Non-LingVoc=Laughs
FUA280	Formulaic=IntegratedChunk; Discourse=Repetition
FUA281	Phonetic=VolumeIncreasing
FUA282	Formulaic=StructuralChunk; Phonetic=SyllableReduction
FUA283	Phonetic=Contraction
FUA284	Formulaic=StructuralChunk; Phonetic=Coarticulation, WeakenedVowel
FUA286	Discourse=LexicalFiller
FUA287	Non-LingVoc=Inhales

TA20	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA90	Intention=EstablishConsensus
FSA91	Intention=Question
FUA289	Phonetic=Coarticulation, WeakenedVowel; Discourse=ChangeTack

TA21	Register=Informal; Topic=Aupair; CEF=A2
FSA92	Intention=EstablishConsensus
FUA290	Non-LingVoc=FilledPause; Discourse=Backchannelling

TA22	Register=Informal; Topic=Aupair; CEF=A2
FSA93	Intention=EstablishConsensus

it is	FUA 309	FSA 98
[inhalés]	FUA 310	
and it is (or as?) such a big	FUA 311	
so enriching	FUA 312	
[inhalés]	FUA 313	

Oh I mean	FUA 314	FSA 99	TA24
it is just	FUA 315		
a	FUA 316		
flip (of) the coin	FUA 317		
flip (of) the coin	FUA 318		
but I would say	FUA 319	FSA 100	
[mm]	FUA 320		
you know	FUA 321		
on average	FUA 322		
most (of?) the ones I have heard about	FUA 323		
have been at least	FUA 324		
[inhalés]	FUA 325		
fine	FUA 326		
[inhalés]	FUA 327		
and some of them have been so	FUA 328		
you know	FUA 329		
they have made friends for life	FUA 330		
[inhalés]	FUA 331	FSA 102	
and	FUA 332		
you know	FUA 333		
they go off to Sweden	FUA 334		
to	FUA 335		
on their holidays	FUA 336		FSA 103
to visit with their au-pair	FUA 337		
their former au-pair in the family	FUA 338		
and the kids are teenagers now	FUA 339	FSA 104	
you know	FUA 340		
[inhalés]	FUA 341		
So it is funny	FUA 342	FSA 105	
[erm]	FUA 343		

FUA291	Disfluency=Incompleted
FUA292	Discourse=Backchannelling, Repetition

TA23	Register=Informal; Topic=Aupair; CEF=B1
FSA94	Intention=Inform/MakeStatement; Turn=TurnGrabbing
FSA95	Intention=Inform/MakeStatement
FSA96	Intention=Inform/MakeStatement
FSA97	Intention=Inform/MakeStatement
FSA98	Intention=Inform/MakeStatement; Turn=TurnLosing
FUA293	Non-LingVoc=Inhalés
FUA294	Phonetic=Contraction, VolumeIncreasing
FUA295	Phonetic=Contraction
FUA296	Phonetic=Contraction; Disfluency=Repetition
FUA297	Formulaic=IntegratedChunk
FUA298	Disfluency=Incompleted
FUA299	Formulaic=IntegratedChunk; Discourse=Hedging
FUA300	Phonetic=DrawnOut; Discourse=BuyTime, LexicalFiller
FUA301	Non-LingVoc=Inhalés
FUA302	Formulaic=StructuralChunk; Lexical=Colloquial
FUA303	Phonetic=DrawnOut; Disfluency=Disjointed
FUA304	Phonetic=SyllableReduction; Disfluency=Disjointed
FUA305	Phonetic=Coarticulation
FUA306	Phonetic=DrawnOut
FUA307	Formulaic=IntegratedChunk
FUA308	Formulaic=IntegratedChunk; Discourse=Repetition
FUA309	Phonetic=Contraction
FUA310	Non-LingVoc=Inhalés
FUA313	Non-LingVoc=Inhalés

TA24	Register=Informal; Topic=Aupair; CEF=B1
FSA99	Intention=EstablishConsensus
FSA100	Intention=Inform/MakeStatement
FSA101	Intention=Inform/MakeStatement
FSA102	Intention=Inform/MakeStatement
FSA103	Intention=Inform/MakeStatement
FSA104	Intention=Inform/MakeStatement

[inhales]	FUA 344	
but ah	FUA 345	
[inhales]	FUA 346	FSA 106
I do not know I mean	FUA 347	
in your study as well	FUA 348	FSA 107
I mean	FUA 349	
you	FUA 350	
you like to	FUA 351	
work at home sometimes	FUA 352	
do not you so	FUA 353	

yeah	FUA 354	FSA 108	TA25
------	---------	----------------	-------------

yeah yeah yeah	FUA 355	FSA 109	TA26
----------------	---------	----------------	-------------

yeah	FUA 356	FSA 110	TA27
------	---------	----------------	-------------

I know	FUA 357	FSA 111	TA28
yeah	FUA 358		
yeah	FUA 359		

FSA105	Intention=Inform/MakeStatement
FSA106	Intention=ExpressOpinion
FSA107	Intention=Question; Turn=TurnLosing
FUA314	Formulaic=IntegratedChunk; Discourse=Opener
FUA315	Formulaic=IntegratedChunk; Phonetic=Contraction, Hyper-articulatedConsonant
FUA317	Formulaic=Idiom; Phonetic=SyllableReduction
FUA318	Formulaic=Idiom; Phonetic=SyllableReduction; Discourse=Repetition
FUA320	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA321	Formulaic=IntegratedChunk
FUA322	Formulaic=Collocation; Phonetic=SyllableReduction
FUA323	Phonetic=Contraction, SyllableReduction
FUA324	Formulaic=Collocation
FUA325	Non-LingVoc=Inhales
FUA327	Non-LingVoc=Inhales
FUA328	Phonetic=DrawnOut
FUA329	Formulaic=IntegratedChunk; Discourse=Repetition
FUA330	Phonetic=Contraction; Discourse=SyntacticalRepair
FUA331	Non-LingVoc=Inhales
FUA333	Formulaic=IntegratedChunk; Discourse=Repetition
FUA334	Lexical=Deixis
FUA336	Lexical=Deixis; Discourse=Add-on
FUA337	Lexical=Deixis; Disfluency=Repetition
FUA338	Lexical=Deixis; Discourse=SyntacticalRepair
FUA339	Lexical=Deixis; Discourse=Aside
FUA340	Formulaic=IntegratedChunk
FUA341	Non-LingVoc=Inhales
FUA342	Phonetic=Contraction
FUA343	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA344	Non-LingVoc=Inhales
FUA345	Discourse=BuyTime, LexicalFiller
FUA346	Non-LingVoc=Inhales
FUA347	Formulaic=IntegratedChunk; Phonetic=Contraction; Discourse=ChangeTack, Repetition
FUA348	Formulaic=Collocation
FUA349	Formulaic=IntegratedChunk; Discourse=Repetition
FUA351	Disfluency=Repetition
FUA353	Phonetic=Contraction, Elision (phone); Disfluency=Incompleted

TA25	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA108	Intention=EstablishConsensus
FUA354	Discourse=Backchannelling

TA26	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA109	Intention=EstablishConsensus

yeah	FUA 360	FSA 112	TA29
------	------------	------------	------

[inhales]	FUA 361	FSA 113	TA30
I guess so	FUA 362		
I mean I would say that	FUA 363	FSA 114	
I do not like to	FUA 364		
to	FUA 365		
be judgmental about it	FUA 366		
obviously you know	FUA 367		
[inhales]	FUA 368		
I mean obviously I have been	FUA 369	FSA 115	
away from	FUA 370		
Stella a lot	FUA 371		
but	FUA 372		
[inhales]	FUA 373		
I	FUA 374	FSA 116	
you know	FUA 375		
[erm]	FUA 376		
[exhales]	FUA 377		
it	FUA 378	FSA 117	
it is	FUA 379		
as you say	FUA 380		
they are only li-	FUA 381	FSA 118	
they are only little	FUA 382		
once (ones?)	FUA 383		
are not they	FUA 384		
so	FUA		

FUA355	Discourse=Backchannelling, Repetition
--------	---------------------------------------

TA27	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA110	Intention=EstablishConsensus
FUA356	Discourse=Backchannelling

TA28	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA111	Intention=EstablishConsensus
FUA357	Discourse=Backchannelling
FUA358	Discourse=Add-on, Backchannelling
FUA359	Discourse=Add-on, Backchannelling, Repetition

TA29	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSA112	Intention=EstablishConsensus
FUA360	Discourse=Backchannelling

TA30	Register=Informal; Topic=Kids; CEF=B1
FSA113	Intention=EstablishConsensus
FSA114	Intention=ExpressDislikes
FSA115	Intention=Inform/MakeStatement
FSA116	Intention=Inform/MakeStatement
FSA117	Intention=Inform/MakeStatement
FSA118	Intention=Inform/MakeStatement
FSA119	Intention=Inform/MakeStatement
FSA120	Intention=Inform/MakeStatement
FSA121	Intention=ExpressLikes; Turn=TurnLosing
FUA361	Non-LingVoc=Inhales
FUA362	Formulaic=IntegratedChunk
FUA363	Formulaic=IntegratedChunk; Phonetic=Contraction
FUA364	Phonetic=Contraction
FUA365	Disfluency=Disjointed, Repetition
FUA366	Disfluency=Disjointed
FUA367	Formulaic=IntegratedChunk; Phonetic=WeakenedVowel
FUA368	Non-LingVoc=Inhales
FUA369	Formulaic=IntegratedChunk; Phonetic=Coarticulation, Contraction, WeakenedVowel; Discourse=Repetition
FUA371	Formulaic=StructuralChunk
FUA373	Non-LingVoc=Inhales
FUA375	Formulaic=IntegratedChunk; Discourse=BuyTime, Repetition
FUA376	Non-LingVoc=FilledPause; Discourse=BuyTime

	385		
[inhales]	FUA 386	FSA 119	
and [erm]	FUA 387		
I mean	FUA 388		
I do not know	FUA 389		
I mean Stella is a pain (in) the arse	FUA 390	FSA 120	
I am sure Maisie is a pain (in) the arse some times too	FUA 391		
but still	FUA 392		
[inhales]	FUA 393		
you would rather	FUA 394	FSA 121	
[mmm]	FUA 395		
have them be	FUA 396		
your pain in the arse	FUA 397		
than somebody else's pain the arse	FUA 398		
for twelve hours a day	FUA 399		
you know	FUA 400		

yeah	FUA 401	FSA 122	TA31
[laughs]	FUA 402		
yeah	FUA 403		
[inhales]	FUA 404		

ah it was great talking to you Martina	FUA 405	FSA 123	TA32
good luck with that	FUA 406	FSA 124	

OK	FUA 407	FSA 125	TA33
----	---------	----------------	-------------

all right	FUA 408	FSA 126	TA34
bye bye	FUA 409	FSA 127	

bye	FUA 410	FSA 128	TA35
-----	---------	----------------	-------------

FUA377	Non-LingVoc=Exhales; Discourse=BuyTime
FUA378	Discourse=Aside
FUA379	Phonetic=Contraction; Discourse=Aside; Disfluency=Repetition
FUA380	Phonetic=Elision (phone); Discourse=Aside
FUA381	Lexical=Deixis; Discourse=SyntacticalRepair; Disfluency=Incompleted
FUA382	Lexical=Deixis; Disfluency=Repetition
FUA384	Phonetic=Contraction; Lexical=Deixis
FUA386	Non-LingVoc=Inhales
FUA387	Phonetic=DrawnOut; Non-LingVoc=FilledPause; Discourse=BuyTime
FUA388	Formulaic=IntegratedChunk; Discourse=Repetition
FUA389	Formulaic=IntegratedChunk
FUA390	Formulaic=IntegratedChunk, Idiom; Phonetic=Contraction; Lexical=Profanity; Discourse=Repetition
FUA391	Formulaic=IntegratedChunk, Idiom; Phonetic=Contraction, Elision (phone); Lexical=Profanity; Discourse=Repetition
FUA393	Non-LingVoc=Inhales
FUA394	Formulaic=StructuralChunk; Phonetic=Contraction
FUA395	Non-LingVoc=FilledPause; Discourse=BuyTime
FUA397	Formulaic=Idiom; Lexical=Profanity; Discourse=Repetition
FUA398	Formulaic=StructuralChunk, Idiom; Phonetic=SyllableReduction; Lexical=Profanity; Discourse=Repetition
FUA400	Formulaic=IntegratedChunk; Discourse=Hedging, Repetition

TA31	Register=Informal; Topic=Kids; CEF=A2
FSA122	Intention=EstablishConsensus
FUA401	Discourse=Backchannelling
FUA402	Non-LingVoc=Laughs
FUA403	Discourse=Backchannelling, Repetition
FUA404	Non-LingVoc=Inhales

TA32	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSA123	Intention=Appreciate; Turn=TurnGrabbing
FSA124	Intention=Compliment
FUA405	Phonetic=WeakenedVowel; Discourse=Add-on, Opener, LexicalFiller
FUA406	Formulaic=Collocation

TA33	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSA125	Intention=EstablishConsensus; Turn=TurnAttempt

TA34	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSA126	Intention=EstablishConsensus; Turn=TurnGrabbing
FSA127	Intention=Compliment
FUA408	Formulaic=Collocation

TA35	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSA128	Intention=Compliment; Turn=TurnLosing

SUE

Dave	FUB 1	FSB 1	TB1
I have been	FUB 2		
trying	FUB 3		
to reach you	FUB 4		
[inhales]	FUB 5		

yeah hi	FUB 6	FSB 2	TB2
[laughs]	FUB 7		
Listen	FUB 8	FSB 3	
[inhales]	FUB 9		
[ahm]	FUB 10		
we	FUB 11		
are	FUB 12		
[ahm]	FUB 13		
really	FUB 14		
looking hard for a place	FUB 15		
for our little girl Maisie	FUB 16		
and I know	FUB 17	FSB 4	
that you had Stella	FUB 18		
in a really good creche	FUB 19		
[inhales]	FUB 20		
and	FUB 21	FSB 5	
I think	FUB 22		
that you said	FUB 23		
[inhales]	FUB 24		
that it was the	FUB 25		
Wee Care creche	FUB 26		
in Monkstown	FUB 27		
is that where she goes	FUB 28		

TB1	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSB1	Intention=Inform/MakeStatement; Turn=TurnLosing
FUB2	Formulaic=StructuralChunk; Phonetic=Contraction
FUB3	Phonetic=DrawnOut
FUB4	Lexical=Colloquial, Regional (American)
FUB5	Non-LingVoc=Inhales

TB2	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSB2	Intention=Compliment
FSB3	Intention=Inform
FSB4	Intention=Inform
FSB5	Intention=Inform
FSB6	Intention=EstablishConsensus
FUB6	Discourse=LexicalFiller, Take-up
FUB7	Non-LingVoc=Laughs
FUB8	Discourse=ChangeTack, SteerConversation
FUB9	Non-LingVoc=Inhales
FUB10	Non-LingVoc=FilledPause; Discourse=BuyTime
FUB11	Phonetic=DrawnOut
FUB13	Non-LingVoc=FilledPause; Discourse=BuyTime
FUB14	Phonetic=DrawnOut
FUB15	Discourse=SetScene
FUB16	Discourse=SetScene
FUB18	Phonetic=elision (phone), Weakened Vowel; Lexical=Colloquial
FUB20	Non-LingVoc=Inhales
FUB21	Phonetic=Hyper-articulatedConsonant
FUB22	Phonetic=DrawnOut
FUB24	Non-LingVoc=Inhales
FUB27	Discourse=There'sMore
FUB28	Discourse=Add-on

Oh	FUB 29	FSB 7	TB3
so she had	FUB 30		
been there	FUB 31		

We	FUB 32	FSB 8	TB4
may be	FUB 33		
being offered a place	FUB 34		
and I know	FUB 35	FSB 9	
that if we get the offer	FUB 36		
we are going to have to decide	FUB 37		
super quick	FUB 38		
it is going to be an overnight decision	FUB 39	FSB 10	
[inhales]	FUB 40		
so I just want to get	FUB 41	FSB 11	
as much information as possible	FUB 42		

TB3	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSB7	Intention=EstablishConsensus; Turn=TurnGrabbing
FUB29	Discourse=Opener
FUB30	Formulaic=StructuralChunk; Phonetic=DrawnOut; Discourse=LexicalFiller
FUB31	Formulaic=StructuralChunk

TB4	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSB8	Intention=Inform; Turn=TurnGrabbing
FSB9	Intention=Inform
FSB10	Intention=Inform
FSB11	Intention=Request
FUB32	Discourse=ChangeTack
FUB33	Discourse=SetScene; Phonetic=DrawnOut
FUB34	Discourse=SetScene, SyntacticalRepair
FUB37	Formulaic=StructuralChunk
FUB38	Formulaic=Collocation; Phonetic=DrawnOut
FUB39	Formulaic=Collocation, StructuralChunk
FUB40	Non-LingVoc=Inhales
FUB41	Formulaic=IntegratedChunk; Discourse=OrganisationalMarker
FUB42	Formulaic=Template

as she could have been	FUB 43	FSB 12	TB5
--------------------------------------	-----------	-----------	-----

Do you know wh-	FUB 44	FSB 13	TB6
what lead	FUB 45		
to that	FUB 46		
do you think	FUB 47	FSB 14	

TB5	Register=Informal; Topic=Playschool, Creche; CEF=B1
FSB12	Intention=EstablishConsensus; Turn=Latching
FUB43	Formulaic=StructuralChunk; Phonetic=DrawnOut

TB6	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSB13	Intention=Question; Turn=TurnGrabbing
FSB14	Intention=Question
FUB44	Phonetic=IPA-dZ; Disfluency=Incomplete
FUB45	Phonetic=DrawnOut; Disfluency=Repetition
FUB46	Phonetic=DrawnOut
FUB47	Phonetic=IPA-dZ; Discourse=Add-on

[laughs]	FUB 48	FSB 15	TB7
----------	-----------	-------------------	------------

TB7	Register=Informal; Topic=Creche staff
FSB15	Intention=EstablishConsensus
FUB48	Non-LingVoc=Laughs; Discourse=Backchannelling

yeah yeah yeah yeah	FUB 49	FSB 16	TB8
---------------------	-----------	-------------------	------------

TB8	Register=Informal; Topic=Creche staff; CEF=A2
FSB16	Intention=EstablishConsensus
FUB49	Discourse=Backchannelling, Repetition

yeah	FUB 50	FSB 17	TB9
and they are not necessarily well paid	FUB 51		
See these are the things we are concerned about	FUB 52	FSB 18	
they are	FUB 53		

TB9	Register=Informal; Topic=Creche staff; CEF=B1
FSB17	Intention=ExpressOpinion; Turn=TurnAttempt
FSB18	Intention=ExpressOpinion
FUB50	Discourse=Opener
FUB51	Phonetic=SyllableReduction
FUB52	Discourse=LexicalFiller
FUB53	Disfluency=Incomplete

yeah and what we are hearing	FUB 54	FSB 19	TB10
cause (because) obviously Miaisie is our first child	FUB 55	FSB 20	
we do not have (any) experience of creches	FUB 56	FSB 21	
but	FUB 57	FSB 22	
[pause]	FUB 58		
people are telling us	FUB 59		
that they have a high turnover of staff	FUB 60		
[inhales]	FUB 61		

TB10	Register=Informal; Topic=Creche staff; CEF=B1
FSB19	Intention=Inform; Turn=TurnGrabbing
FSB20	Intention=Inform
FSB21	Intention=Inform
FSB22	Intention=EstablishConsensus
FUB54	Discourse=Opener
FUB55	Phonetic=WeakenedVowel; Lexical=Colloquial; Discourse=Aside, SetScene
FUB56	Phonetic=Contraction; Discourse=Aside
FUB57	Discourse=There'sMore
FUB58	Non-LingVoc=Pause
FUB60	Phonetic=DrawnOut
FUB61	Non-LingVoc=Inhales

mmm yeah	FUB 62	FSB 23	TB11
yeah	FUB 63		

TB11	Register=Informal; Topic=Creche staff; CEF=A2
FSB23	Intention=EstablishConsensus
FUB62	Discourse=Backchannelling
FUB63	Discourse=Backchannelling, Repetition

yeah	FUB 64	FSB 24	TB12
------	-----------	-----------	------

TB12	Register=Informal; Topic=Creche staff; CEF=A2
FSB24	Intention=EstablishConsensus
FUB64	Discourse=Backchannelling

[inhalas]	FUB 65	FSB 25	TB13
-----------	-----------	-----------	------

TB13	
FSB25	Turn=TurnAttempt
FUB65	Non-LingVoc=Inhalas

is it	FUB 66	FSB 26	TB14
yeah	FUB 67	FSB 27	
I was just going to say	FUB 68		
do you	FUB 69	FSB 28	
is it your impression that	FUB 70		
that is like kind of	FUB 71		
typical behaviour	FUB 72		
because	FUB 73	FSB 29	
if that is it	FUB 74		
and if that is the	FUB 75		
best that we are going to be able to do	FUB 76		
[inhalas]	FUB 77		

TB14	Register=Informal; Topic=Creche staff; CEF=B1
FSB26	Intention=Clarification; Turn=TurnGrabbing
FSB27	Intention=Clarification
FSB28	Intention=Clarification
FSB29	Intention=Inform/MakeStatement; Turn=TurnLosing
FUB66	Disfluency=Incomplete
FUB67	Discourse=LexicalFiller
FUB68	Formulaic=StructuralChunk; Phonetic=WeakenedVowel; Discourse=Aside
FUB69	Disfluency=Incomplete
FUB70	Discourse=SyntacticalRepair
FUB71	Formulaic=IntegratedChunk; Phonetic=Contraction, WeakenedVowel
FUB74	Phonetic=Contraction
FUB75	Phonetic=Contraction; Discourse=Repetition
FUB76	Formulaic=StructuralChunk; Phonetic=DrawnOut, SyllableReduction; Discourse=There'sMore
FUB77	Non-LingVoc=Inhalas

yeah	FUB 78	FSB 30	TB15
------	-----------	-----------	------

TB15	Register=Informal; Topic=Playshool, Creche; CEF=A2
-------------	----------------------------------------------------

FSB30	Intention=EstablishConsensus
FUB78	Discourse=Backchannelling

wa- wa-s	FUB 79	FSB 31	TB16
was there some point when you	FUB 80		
considered that	FUB 81		
maybe the way to do this was	FUB 82		
f- maybe for one of you to work half-time and	FUB 83		
[inhales]	FUB 84		
the other one	FUB 85	FSB 32	
I do not know	FUB 86		
have a- a- a-	FUB 87		
opposing schedules	FUB 88		
so that she would not need to be	FUB 89		
cause (because) we are	FUB 90		
[inhales]	FUB 91		

TB16	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSB31	Intention=Clarification; Turn=TurnGrabbing
FSB32	Intention=Clarification; Turn=TurnLosing
FUB79	Disfluency=Stammer
FUB80	Discourse=There'sMore
FUB83	Phonetic=DrawnOut; Discourse=SyntacticalRepair
FUB84	Non-LingVoc=Inhales
FUB86	Formulaic=IntegratedChunk; Discourse=Softening
FUB87	Disfluency=Stammer
FUB89	Formulaic=StructuralChunk; Phonetic=Contraction
FUB90	Lexical=Colloquial; Discourse=ChangeTack; Disfluency=Incomplete
FUB91	Non-LingVoc=Inhales

yeah	FUB 92	FSB 33	TB17
------	--------	---------------	-------------

TB17	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB33	Intention=EstablishConsensus
FUB92	Discourse=Backchannelling

yeah exactly	FUB 93	FSB 34	TB18
yeah	FUB 94		
we are kind of	FUB 95	FSB 35	
contemplating that possibility	FUB 96		
as well	FUB 97		
[inhales]	FUB 98		
[erm]	FUB 99	FSB 36	
and I think where we are going	FUB 100		
is that we might	FUB 101		
try out	FUB 102		

TB18	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSB34	Intention=EstablishConsensus; Turn=TurnGrabbing
FSB35	Intention=Inform/MakeStatement
FSB36	Intention=Inform/MakeStatement
FSB37	Intention=Inform/MakeStatement
FSB38	Intention=ExpressDislikes; Turn=TurnKeeping
FSB39	Intention=ExpressDislikes
FSB40	Intention=ExpressDislikes; Turn=TurnKeeping
FUB93	Phonetic=Coarticulation
FUB94	Discourse=Repetition, There'sMore

the creche	FUB 103	
[inhales]	FUB 104	
if it is not working out	FUB 105	FSB 37
then kind of re-organise our work schedule	FUB 106	
(But)the problem is that	FUB 107	FSB 38
[inhales]	FUB 108	
this creche takes a	FUB 109	FSB 39
six	FUB 110	
month	FUB 111	
deposit	FUB 112	
[pause]	FUB 113	
a six month deposit	FUB 114	FSB 40
and we do not even know	FUB 115	
if it will work out	FUB 116	

FUB95	Formulaic=IntegratedChunk; Phonetic=WeakenedVowel
FUB97	Formulaic=Collocation; Phonetic=Coarticulation
FUB98	Non-LingVoc=Inhales
FUB99	Non-LingVoc=FilledPause; Discourse=BuyTime
FUB100	Phonetic=DrawnOut
FUB101	Phonetic=Coarticulation
FUB102	Phonetic=DrawnOut
FUB103	Discourse=There'sMore
FUB104	Non-LingVoc=Inhales
FUB105	Phonetic=Contraction
FUB106	Formulaic=IntegratedChunk; Phonetic=WeakenedVowel; Discourse=OrganisationalMarker, Repetition
FUB107	Phonetic=Coarticulation, VolumeIncreasing; Discourse=ChangeTack
FUB108	Non-LingVoc=Inhales
FUB109	Phonetic=Hyper-articulatedConsonant
FUB110	Phonetic=DrawnOut; Disfluency=Disjointed
FUB111	Phonetic=DrawnOut; Disfluency=Disjointed
FUB112	Phonetic=DrawnOut; Disfluency=Disjointed
FUB113	Non-LingVoc=Pause
FUB114	Discourse=Add-on, Repetition
FUB115	Phonetic=Contraction; Discourse=Add-on
FUB116	Phonetic=DrawnOut, WeakenedVowel; Discourse=Add-on, Repetition

Right right right	FUB 117	FSB 41	TB19
-------------------	---------	---------------	-------------

TB19	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB41	Intention=EstablishConsensus
FUB117	Discourse=Backchannelling, Repetition

I know they give it	FUB 118	FSB 42	TB20
yeah	FUB 119	FSB 43	

TB20	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB42	Intention=EstablishConsensus; Turn=TurnAttempt
FSB43	Intention=EstablishConsensus
FUB118	Phonetic=WeakenedVowel; Lexical=Deixis
FUB119	Discourse=Add-on

So you would not say that she was	FUB 120	FSB 44	TB21
unhappy there	FUB 121		

TB21	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSB44	Intention=Clarification; Turn=TurnGrabbing
FUB120	Phonetic=SyllableReduction; Discourse=Opener
FUB121	Phonetic=DrawnOut

Yeah OK	FUB 122	FSB 45	TB22
---------	------------	-----------	------

TB22	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB45	Intention=EstablishConsensus
FUB122	Discourse=Backchannelling

Yeah OK	FUB 123	FSB 46	TB23
---------	------------	-----------	------

TB23	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB46	Intention=EstablishConsensus
FUB123	Discourse=Backchannelling

so it was not really	FUB 124	FSB 47	TB24
an inspiring	FUB 125		
teaching sta-	FUB 126		
staff	FUB 127		
yeah	FUB 128	FSB 48	

TB24	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB47	Intention=Evaluation; Turn=TurnGrabbing
FSB48	Intention=Evaluation; Turn=TurnLosing
FUB124	Phonetic=Contraction, SyllableReduction; Discourse=Opener
FUB125	Phonetic=Coarticulation, DrawnOut
FUB126	Disfluency=Incompleted
FUB127	Disfluency=Repetition
FUB128	Discourse=Add-on

yeah the	FUB 129	FSB 49	TB25
----------	------------	-------------------	-------------

TB25	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB49	Intention=EstablishConsensus; Turn=TurnAttempt
FUB129	Discourse=opener

right right right	FUB 130	FSB 50	TB26
yeah	FUB 131		

TB26	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB50	Intention=EstablishConsensus
FUB130	Discourse=Backchannelling, Repetition
FUB131	Discourse=Backchannelling, Add-on

we do not need to put pressure on them	FUB 132	FSB 51	TB27
or anything	FUB 133		

TB27	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSB51	Intention=EstablishConsensus; Turn=TurnGrabbing
FUB132	Phonetic=Contraction
FUB133	Formulaic=IntegratedChunk

that was a little bit lacking	FUB 134	FSB 52	TB28
-------------------------------	------------	-------------------	-------------

TB28	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB52	Intention=EstablishConsensus; Turn=TurnGrabbing
FUB134	Formulaic=StructuralChunk; Phonetic=DrawnOut

a little bit lacking	FUB 135	FSB 53	TB29
yeah	FUB 136		

TB29	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB53	Intention=EstablishConsensus
FUB135	Formulaic=StructuralChunk; Discourse=Backchannelling, Take-up
FUB136	Discourse=Backchannelling, Add-on

[inhales]	FUB 137	FSB 54	TB30
our	FUB 138	FSB 55	
our other	FUB 139		
our other	FUB		

TB30	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB54	Turn=TurnAttemp
FSB55	Intention=Inform/MakeStatement
FUB137	Non-LingVoc=Inhales

	140	
option	FUB 141	

yeah that is true	FUB 142	FSB 56	TB31
I know	FUB 143	FSB 57	
you want the best	FUB 144		
for your own	FUB 145		
[inhales]	FUB 146		

Well yeah	FUB 147	FSB 58	TB32
we were thinking	FUB 148		
maybe	FUB 149		
the aupair route	FUB 150		
but the thing is	FUB 151	FSB 59	
if we take in an au-pair	FUB 152		
[inhales]	FUB 153		
we basically have to give up our study	FUB 154		
because the study would have to become	FUB 155	FSB 60	
the au-pair's room	FUB 156		
[inhales]	FUB 157		
and so	FUB 158	FSB 61	
we would always have	FUB 159		
another person	FUB 160		
in our little tiny house	FUB 161		
and we would	FUB 162	FSB 62	
of course have to be	FUB 163		
providing room and board	FUB 164		
[inhales]	FUB 165		
and [erm]	FUB 166		
[inhales]	FUB 167		
so that is one option	FUB 168	FSB 63	
and	FUB 169		

FUB138	Phonetic=VolumeIncreasing
FUB139	Disfluency=Repetition
FUB140	Phonetic=WeakenedVowel; Disfluency=Repetition

TB31	Register=Informal; Topic=Playshool, Creche; CEF=A2
FSB56	Intention=EstablishConsensus; Turn=TurnGrabbing
FSB57	Intention=EstablishConsensus
FUB142	Phonetic=Contraction
FUB146	Non-LingVoc=Inhales

TB32	Register=Informal; Topic=Aupair; CEF=B1
FSB58	Intention=ExpressOpinion
FSB59	Intention=ExpressDislikes
FSB60	Intention=ExpressDislikes
FSB61	Intention=ExpressDislikes
FSB62	Intention=ExpressDislikes
FSB63	Intention=ExpressOpinion; Turn=TurnLosing
FUB147	Discourse=LexicalFiller, Opener
FUB151	Phonetic=DrawnOut, Discourse=There'sMore
FUB152	Formulaic=StructuralChunk; Discourse=SetScene
FUB153	Non-LingVoc=Inhales
FUB154	Formulaic=StructuralChunk; Phonetic=SyllableReduction
FUB155	Formulaic=StructuralChunk
FUB157	Non-LingVoc=Inhales
FUB159	Formulaic=StructuralChunk
FUB160	Phonetic=DrawnOut
FUB161	Formulaic=Collocation
FUB162	Formulaic=StructuralChunk; Phonetic=Contraction
FUB163	Formulaic=Collocation, StructuralChunk
FUB164	Phonetic=DrawnOut
FUB165	Non-LingVoc=Inhales
FUB166	Non-LingVoc=FilledPause; Discourse=BuyTime
FUB167	Non-LingVoc=Inhales
FUB168	Phonetic=Contraction
FUB169	Disfluency=Incompleted

yeah	FUB 170	FSB 64	TB33
------	------------	-----------	------

i- depending entirely	FUB 171	FSB 65	TB34
on whether you	FUB 172		
get a good au-pair	FUB 173		

TB33	Register=Informal; Topic=Aupair; CEF=A2
FSB64	Intention=EstablishConsensus
FUB170	Discourse=Backchannelling

TB34	Register=Informal; Topic=Aupair; CEF=A2
FSB65	Intention=ExpressOpinion; Turn=TurnGrabbing
FUB171	Phonetic=VolumeIncreasing; Disfluency=Incompleted
FUB172	Phonetic=Coarticulation

yeah	FUB 174	FSB 66	TB35
------	------------	-----------	------

TB35	Register=Informal; Topic=Aupair; CEF=A2
FSB66	Intention=EstablishConsensus
FUB174	Discourse=Backchannelling

yeah yeah yeah yeah	FUB 175	FSB 67	TB36
[inhales]	FUB 176		
yeah	FUB 177		

yeah	FUB 178	FSB 68	TB37
------	------------	-----------	------

TB36	Register=Informal; Topic=Aupair; CEF=A2
FSB67	Intention=EstablishConsensus
FUB175	Discourse=Backchannelling, Repetition
FUB176	Non-LingVoc=Inhales
FUB177	Discourse=Backchannelling, Repetition

TB37	Register=Informal; Topic=Aupair; CEF=A2
FSB68	Intention=EstablishConsensus
FUB178	Discourse=Backchannelling

[inhalas]	FUB 179	FSB 69	TB38
yeah I do	FUB 180		
and	FUB 181	FSB 70	
and I want to have the	FUB 182		
and I want to have	FUB 183		
exactly	FUB 184		
I will be able to spend more time with Maisie	FUB 185		
if I am actually working	FUB 186	FSB 71	
from home more	FUB 187		
[inhalas]	FUB 188		
so that	FUB 189	FSB 72	
if she were in a creche	FUB 190		
at least I could get to her quicker	FUB 191		
than if I am coming	FUB 192		
out of town	FUB 193		
[inhalas]	FUB 194		
[ahm]	FUB 195	FSB 73	
so there is that flexibility	FUB 196		
[inhalas]	FUB 197	FSB 74	
a a and I think that	FUB 198		
[inhalas]	FUB 199		
you know	FUB 200		
it is interesting	FUB 201		
having this conversation with you	FUB 202		
perhaps what I needed most was	FUB 203	FSB 75	
kind of	FUB 204		
to hear	FUB 205		
myself	FUB 206		
talk	FUB 207		
about it	FUB 208		
in order to kind of arrive	FUB 209	FSB 76	
at some sort of a decision	FUB 210		
[inhalas]	FUB 211	FSB 77	
cause (because) what you are telling me	FUB 212		

TB38	Register=Informal; Topic=Playshool, Creche; CEF=B1
FSB69	Intention=EstablishConsensus; Turn=TurnGrabbing
FSB70	Intention=ExpressOpinion
FSB71	Intention=ExpressOpinion
FSB72	Intention=ExpressOpinion
FSB73	Intention=ExpressOpinion
FSB74	Intention=ExpressOpinion; Turn=TurnKeeping
FSB75	Intention=ExpressOpinion
FSB76	Intention=ExpressOpinion; Turn=TurnKeeping
FSB77	Intention=Evaluation
FSB78	Intention=ExpressOpinion
FSB79	Intention=ExpressOpinion
FSB80	Intention=ExpressOpinion
FUB179	Non-LingVoc=Inhalas
FUB180	Discourse=There'sMore
FUB182	Phonetic=WeakenedVowel; Disfluency=Repetition
FUB183	Phonetic=WeakenedVowel; Discourse=SyntacticalRepair
FUB184	Phonetic=Coarticulation
FUB185	Formulaic=StructuralChunk; Phonetic=Coarticulation, Contraction
FUB186	Phonetic=Contraction
FUB187	Phonetic=DrawnOut
FUB188	Non-LingVoc=Inhalas
FUB189	Phonetic=DrawnOut
FUB191	Formulaic=Collocation; Phonetic=Coarticulation; Lexical=Colloquial
FUB192	Phonetic=Contraction
FUB194	Non-LingVoc=Inhalas
FUB195	Non-LingVoc=FilledPause; Discourse=BuyTime
FUB196	Phonetic=Coarticulation, Contraction
FUB197	Non-LingVoc=Inhalas
FUB198	Phonetic=VolumeIncreasing; Disfluency=Stammer
FUB199	Non-LingVoc=Inhalas
FUB200	Formulaic=IntegratedChunk
FUB201	Phonetic=Contraction
FUB204	Formulaic=IntegratedChunk; Phonetic=WeakenedVowel

is really helpful	FUB 213	
and it is informative	FUB 214	
[inhales]	FUB 215	
but I am kind of	FUB 216	FSB 78
starting to work out	FUB 217	
what my own feelings are about it	FUB 218	
[inhales]	FUB 219	FSB 79
[ahm]	FUB 220	
and I think probably	FUB 221	
the way I feel is that	FUB 222	FSB 80
it is just not a great solu...	FUB 223	
solution regardless	FUB 224	
to have a really young child	FUB 225	
away from you so much	FUB 226	

FUB205	Phonetic=Coarticulation
FUB207	Disfluency=Disjointed
FUB208	Phonetic=DrawnOut, VolumeIncreasing; Disfluency=Disjointed
FUB209	Formulaic=IntegratedChunk, StructuralChunk; Phonetic=Coarticulation; Discourse=Repetition
FUB210	Formulaic=IntegratedChunk; Phonetic=WeakenedVowel, Elision (phone)
FUB211	Non-LingVoc=Inhales
FUB212	Lexical=Colloquial
FUB214	Phonetic=Contraction
FUB215	Non-LingVoc=Inhales
FUB216	Formulaic=IntegratedChunk; Phonetic=Contraction, WeakenedVowel; Discourse=Repetition
FUB219	Non-LingVoc=Inhales
FUB220	Non-LingVoc=FilledPause; Discourse=BuyTime
FUB223	Phonetic=Contraction; Disfluency=Incompleted
FUB224	Disfluency=Repetition
FUB226	Phonetic=DrawnOut

or to dwell on it	FUB 227	FSB 81	TB39
yeah	FUB 228	FSB 82	TB40

TB39	Register=Informal; Topic=Kids; CEF=A2
FSB81	Intention=EstablishConsensus; Turn=TurnAttempt

TB40	Register=Informal; Topic=Kids; CEF=A2
FSB82	Intention=EstablishConsensus
FUB228	Discourse=Backchannelling

yeah	FUB 229	FSB 83	TB41
------	---------	--------	------

TB41	Register=Informal; Topic=Kids; CEF=A2
FSB83	Intention=EstablishConsensus
FUB229	Discourse=Backchannelling

[laughs]	FUB 230	FSB 84	TB42
----------	---------	--------	------

TB42	Register=Informal; Topic=Kids
------	-------------------------------

FSB84	Intention=EstablishConsensus
FUB230	Non-LingVoc=Laughs; Discourse=Backchannelling

than somebody else's	FUB 231	FSB 85	TB43
most of the time anyway	FUB 232		
[clicks]	FUB 233	FSB 86	
listen Dave	FUB 234		
thanks so much	FUB 235		
I really appreciate that input	FUB 236		
[inhales]	FUB 237		
and	FUB 238	FSB 87	

yeah I will let you know	FUB 239	FSB 88	TB44
what we decide	FUB 240		
just just	FUB 241		
so you know	FUB 242		
OK	FUB 243	FSB 89	
[inhales]	FUB 244		

thanks Dave	FUB 245	FSB 90	TB45
-------------	---------	---------------	-------------

bye	FUB 246	FSB 91	TB46
-----	---------	---------------	-------------

TB43	Register=Informal; Topic=Kids; CEF=A2
FSB85	Intention=EstablishConsensus; Turn=TurnGrabbing
FSB86	Intention=Thank
FSB87	Intention=Inform/MakeStatement; Turn=TurnLosing
FUB231	Formulaic=StructuralChunk; Discourse=Take-up
FUB232	Lexical=Colloquial, Regional (Irish); Discourse=Add-on
FUB233	Non-LingVoc=Clicks
FUB234	Discourse=ChangeTack, SteerConversation
FUB235	Phonetic=DrawnOut
FUB236	Phonetic=SyllableReduction
FUB237	Non-LingVoc=Inhales
FUB238	Disfluency=Incompleted

TB44	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSB88	Intention=Inform/MakeStatement
FSB89	Intention=EstablishConsensus; Turn=TurnLosing
FUB239	Phonetic=Contraction; Discourse=Opener, LexicalFiller
FUB241	Discourse=BuyTime; Disfluency=Repetition
FUB243	Discourse=Add-on
FUB244	Non-LingVoc=Inhales

TB45	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSB90	Intention=Thank

TB46	Register=Informal; Topic=Playschool, Creche; CEF=A2
FSB91	Intention=Compliment; Turn=TurnGrabbing
FUB246	Discourse=Take-up

Statistics on main attribute values in both speakers

a) Speaker Intention

	Harry		Sue	
	no.	%	no.	%
appreciate	1	1%	0	0%
clarification	0	0%	6	8%
compliment	5	4%	2	3%
establish consensus	26	21%	34	43%
express dislikes	11	9%	7	9%
express likes	1	1%	0	0%
express opinion	31	26%	15	19%
evaluation	4	3%	3	4%
face saving	1	1%	0	0%
inform/make statement	34	28%	8	10%
question	3	2%	2	3%
request	0	0%	1	1%
revise opinion/recast	4	3%	0	0%
thank	0	0%	2	3%
total	121		80	

b) Turn Construction

	Harry		Sue	
	no.	%	no.	%
latching	0	0%	1	3%
turn attempt	3	10%	5	15%
turn keeping	4	13%	4	12%
turn grabbing	9	29%	16	48%
turn losing	15	48%	7	21%
total	31		33	

c) Discourse Function

	Harry		Sue	
	no.	%	no.	%
add-on	12	6%	11	11%
aside	12	6%	3	3%
backchannelling	13	7%	22	22%

buy time	45	24%	7	7%
change tack	6	3%	5	5%
set scene	0	0%	6	6%
hedging	9	5%	0	0%
lexical filler	10	5%	6	6%
opener	6	3%	8	8%
organisational marker	2	1%	2	2%
repetition	40	22%	13	13%
softening	3	2%	1	1%
steer conversation	0	0%	2	2%
syntactical repair	10	5%	4	4%
take-up	6	3%	4	4%
there's more ...	11	6%	8	8%
total	185		102	

d) Formulaic Sequences

	Harry		Sue	
	no.	%	no.	%
collocation	18	13%	18	21%
idiom	7	5%	7	8%
structural chunk	23	17%	43	51%
template	6	4%	6	7%
integrated chunk	81	60%	11	13%
total	135		85	

e) Phonetic Features

	Harry		Sue	
	no.	%	no.	%
assimilation	1	1%	2	2%
co-articulation	23	19%	24	18%
contraction	72	59%	73	56%
drawn out	1	1%	2	2%
elision (phone)	17	14%	17	13%
hyper-articulated consonant	1	1%	2	2%
IPA- <i>n</i>	1	1%	2	2%
syllable reduction	1	1%	2	2%
weakened vowel	2	2%	1	1%
volume increasing	3	2%	5	4%
total	122		130	

Appendix 2: Snippets for Testing and Training Samples

Z1	Particularly when a woman asks you.
Z2	It was enjoyable.
Z3	Which one are you thinking of.
Z4	People were generally quite nice.
Z5	But I have also had five hours a week at the DBS.
Z6	I have been teaching about four hours a week.
Z7	From pillar to post.
Z8	As I said earlier on.
Z9	You are very good cartoonist.
Z10	It is a pity you do not do a bit more.
Z11	That are particularly interesting.
Z12	Certainly made up for that.
Z13	I moved out of home when I was eighteen.
Z14	When did you move out of home?
Z15	Thanks for very much for taking the time to come in.
Z16	I do not know what you were doing in the eighties.
Z17	Do you know what I mean?
Z18	You know the mistakes that come up during the exam.
Z19	It is going to continue existing.
Z20	People just were not able to talk to him.
Z21	I do not want to go back.
Z22	Especially when you are working three nights a week.
Z23	Something to do with the fact that I am I am growing a bit older.
Z24	Fifty sixty percent of the people.
Z25	I cannot walk down the street without having to walk on the road.
Z26	They are always looking for people.
Z27	Yeah but got to remember.
Z28	He could not have known what it meant.
Z29	You have a list of questions?
Z30	Well that is exactly what the Italians would have been doing.
Z31	Are we on the see-food diet?
Z32	You heard about the award winning farmer?
Z33	And I got back to the hotel on Sunday evening
Z34	I know what you are going to say
Z35	Well I do have the occasional sleepless night.
Z36	Well do you know what was really shocking?
Z37	You are scraping the bottom of the barrel sir.
Z38	He had no idea what it meant
Z39	Would you believe it?
Z40	He was treated very badly
Z41	That is the only thing came into his head.
Z42	It would be to my favour
Z43	No I would have no question
Z44	It is fantastic is not it yeah
Z45	It is your busiest time of the year
Z46	Well do you have a back garden?

**Appendix 3: Questionnaire 1 for Both Control Group and Test Group
(English Version)**

Participant Number: _____ Date: _____

Gender: ☐ male ☐ female

Age: _____

Mother tongue: _____

Fluent in any other language? _____

No. years learning English: _____

No. hours English class per week: _____

No. minutes listening to English per week: _____

No. minutes speaking English per week: _____

Appendix 4: First Version of Data Assessment for Test 1

	which	one	are	you	thinking	of	SCORE (Max=24)	
Max-->	4	4	4	4	4	4		
1	3	0	0	3	2	3	11	46%
2	1	0	0	0	2	3	6	25%
3	1	0	0	1	2	3	7	29%
4	1	0	0	3	2	3	9	38%
5	3	1	3	3	1	1	12	50%
6	1	0	0	3	2	3	9	38%
7	1	0	0	0	2	1	4	17%
8	1	0	0	0	1	1	3	13%
9	3	0	0	1	2	1	7	29%
10	1	0	0	3	2	3	9	38%
11	1	0	0	1	3	3	8	33%
12	1	0	0	1	2	3	7	29%
13	1	0	0	3	2	3	9	38%
14	1	0	0	3	3	3	10	42%
15	1	0	0	1	2	3	7	29%
16	1	0	3	3	3	3	13	54%
17	1	0	0	3	2	3	9	38%
18	1	0	0	1	1	3	6	25%
19	1	0	0	3	2	3	9	38%
20	1	0	0	3	2	3	9	38%
21	1	0	0	1	3	1	6	25%
22	1	0	0	1	2	3	7	29%
23	1	0	0	2	2	3	8	33%
24	1	0	0	3	2	1	7	29%
25	1	0	0	3	2	3	9	38%
26	1	0	0	3	2	3	9	38%
27	1	0	0	3	3	3	10	42%
28	1	0	0	3	2	3	9	38%
29	1	0	0	3	2	3	9	38%
30	1	0	0	3	2	3	9	38%
31	1	0	0	0	0	3	4	17%
32	1	0	0	3	2	3	9	38%
33	1	0	0	3	0	3	7	29%
34	1	0	0	3	2	0	6	25%
35	1	0	0	3	2	3	9	38%
36	3	0	0	0	1	0	4	17%
37	1	0	0	3	2	3	9	38%
38	1	0	0	3	0	0	4	17%
39	1	0	0	3	2	3	9	38%
40	3	3	3	3	2	1	15	63%

Appendix 5: Refined Version of Data Assessment for Test 1

	which	one	are	you	thinking	of	SCORE (Max = 20)	
Max-->	3	3	4	4	3	3		
1	3	0	0	3	2	3	11	55%
2	1	0	0	0	2	3	6	30%
3	1	0	0	1	2	3	7	35%
4	1	0	0	3	2	3	9	45%
5	3	1	3	3	1	1	12	60%
6	1	0	0	3	2	3	9	45%
7	1	0	0	0	2	1	4	20%
8	1	0	0	0	1	1	3	15%
9	3	0	0	1	2	1	7	35%
10	1	0	0	3	2	3	9	45%
11	1	0	0	1	3	3	8	40%
12	1	0	0	1	2	3	7	35%
13	1	0	0	3	2	3	9	45%
14	1	0	0	3	3	3	10	50%
15	1	0	0	1	2	3	7	35%
16	1	0	3	3	3	3	13	65%
17	1	0	0	3	2	3	9	45%
18	1	0	0	1	1	3	6	30%
19	1	0	0	3	2	3	9	45%
20	1	0	0	3	2	3	9	45%
21	1	0	0	1	3	1	6	30%
22	1	0	0	1	2	3	7	35%
23	1	0	0	2	2	3	8	40%
24	1	0	0	3	2	1	7	35%
25	1	0	0	3	2	3	9	45%
26	1	0	0	3	2	3	9	45%
27	1	0	0	3	3	3	10	50%
28	1	0	0	3	2	3	9	45%
29	1	0	0	3	2	3	9	45%
30	1	0	0	3	2	3	9	45%
31	1	0	0	0	0	3	4	20%
32	1	0	0	3	2	3	9	45%
33	1	0	0	3	0	3	7	35%
34	1	0	0	3	2	0	6	30%
35	1	0	0	3	2	3	9	45%
36	3	0	0	0	1	0	4	20%
37	1	0	0	3	2	3	9	45%
38	1	0	0	3	0	0	4	20%
39	1	0	0	3	2	3	9	45%
40	3	3	3	3	2	1	15	75%

Appendix 6: Explanation of Training Purpose and Training Materials for Control Group

1. Problems for Chinese language learners when they involve in the native-to-native conversations

Many Chinese language learners, even advanced learners, inevitably encounter a dilemma with real English spoken communication. The pronunciation of words which they hear spoken by L1 English speakers is different from that which they learned in their language classes. Some words which are familiar to them become unrecognisable when spoken in an authentic English speaking community. The speech flow produced is often too rapid, vague and ‘messy’ to be understood in detail. So many complains from Chinese students are ‘listening is too difficult’.

2. What Chinese learners need is to expose to real authentic English speech

Since the difference between the ‘standard’ English pronunciations students learn in their class, that is, the citation forms of the words, and the co-articulation pronunciations in the stream of English connected speech, e.g. deletion, weak forms and contraction forms, etc., in order to understand ‘messy’, rapid, flow speech so as to facilitate mutual intelligibility, Chinese students need as much as possible to be exposed to natural, authentic spoken English.

Our traditional pronunciation teaching is to segment the sounds of language into discrete items which are what language teachers mostly demonstrate to students in class and are exclusively for non-L1 speakers. These ideal, isolated forms of words, however, are never spoken by L1-L1 speakers and are in no way similar to the spontaneous

English speech of everyday life. If language learners are constantly exposed to this idealised kind of spoken English, inevitably they will find it quite impossible to understand normal English speech. Since when L1 speakers talk with L1 speakers, they always tend to make the less efforts by producing co-articulations (running words together) in their speech. Of course, the way of making these co-articulations might be a little different in British, America, Australia, and the other English speaking countries. The main linguistic features, however, are similar. Thus, increasing exposure to real, informal English speech can give Chinese language learners more chances to capture these characteristics so that they can cope with the ‘untidy’ flow of natural English speech between L1 speakers.

Therefore, these real spontaneous native-to-native conversations made in CSAL Lab by DIT language and audio researchers are to demonstrate the salient linguistic features in authentic English communication. The main point among them, for Chinese language learners, is the reduced forms and weak forms of words, since the learners are used to producing every word clearly with its citation form. For example, Snippet 6 is transcribed as ‘well that is exactly what the Italians would have been doing’. Due to the fast speed of delivery, ‘would have been’ is reduced to /wudəbin/. So only when they are aware of and acquainted with the differences between the co-articulated forms and the isolated forms of words, then Chinese language learners can build up patterns of familiar expressions and achieve efficient understanding in a real, authentic stream-like English speaking community.

Appendix 7: Explanation of Training Purpose and Training Materials for Test Group

1. Problems for Chinese language learners when they involve in the native-to-native conversations

Many Chinese language learners, even advanced learners, inevitably encounter a dilemma with real English spoken communication. The pronunciation of words which they hear spoken by L1 English speakers is different from that which they learned in their language classes. Some words which are familiar to them become unrecognisable when spoken in an authentic English speaking community. The speech flow produced is often too rapid, vague and ‘messy’ to be understood in detail. So many complains from Chinese students are ‘listening is too difficult’.

2. What Chinese learners need is to expose to real authentic English speech

Since the difference between the ‘standard’ English pronunciations students learn in their class, that is, the citation forms of the words, and the co-articulation pronunciations in the stream of English connected speech, e.g. deletion, weak forms and contraction forms, etc., in order to understand ‘messy’, rapid, flow speech so as to facilitate mutual intelligibility, Chinese students need as much as possible to be exposed to natural, authentic spoken English.

Our traditional pronunciation teaching is to segment the sounds of language into discrete items which are what language teachers mostly demonstrate to students in class and are exclusively for non-L1 speakers. These ideal, isolated forms of words, however, are never spoken by L1-L1 speakers and are in no way similar to the spontaneous

English speech of everyday life. If language learners are constantly exposed to this idealised kind of spoken English, inevitably they will find it quite impossible to understand normal English speech. Since when L1 speakers talk with L1 speakers, they always tend to make the less efforts by producing co-articulations (running words together) in their speech. Of course, the way of making these co-articulations might be a little different in British, America, Australia, and the other English speaking countries. The main linguistic features, however, are similar. Thus, increasing exposure to real, informal English speech can give Chinese language learners more chances to capture these characteristics so that they can cope with the ‘untidy’ flow of natural English speech between L1 speakers.

Therefore, these real spontaneous native-to-native conversations made in CSAL Lab by DIT language and audio researchers are to demonstrate the salient linguistic features in authentic English communication. The main point among them, for Chinese language learners, is the reduced forms and weak forms of words, since the learners are used to producing every word clearly with its citation form. For example, Snippet 6 is transcribed as ‘well that is exactly what the Italians would have been doing’. Due to the fast speed of delivery, ‘would have been’ is reduced to /wudəbin/. So only when they are aware of and acquainted with the differences between the co-articulated forms and the isolated forms of words, then Chinese language learners can build up patterns of familiar expressions and achieve efficient understanding in a real, authentic stream-like English speaking community.

3. Slowed-down speed of delivery is helpful in getting language listeners to capture the important ‘missing’ phonetic characteristics of everyday listening

Passive exposure alone is insufficient. Even though language listeners can access the same listening materials as many times as possible, however, if these materials are played at the same speed, it is still difficult for them to become aware of the important segmental clues which they can rely on in understanding normal, informal English speech. Since most of the misleading information in the signal occurs in unstressed syllables, thus these unstressed segments are difficult to perceive for language learners when they are uttered at a fast speed. Therefore, language learners need a kind of comfortable speed which allows them to appreciate the way these significant elements are produced.

For example, Snippet 5 is actually uttered 'I can't walk down the street without having to walk on the road'. It is hard to hear the negative 't' in 'can't' that causes a problem for testing students, but it is made clear to the L1 listeners by extending the nasal '-n-' and following with 'without ...'. So only the slowed-down version can give language learners the chance to discover these differences and help them to be comfortable in an English speaking community.

Apart from helping language learners to recognise salient segmentals (word elements) in natural English speech, the slowed-down speed can also help learners follow the intonation and rhythm patterns in English. Chinese speakers, heavily influenced by their mother tongue, tend to produce their utterance with a lower tonal range (flatter) since they more rely on the tone which indicates the meaning of the words in Chinese rather than the intonation of the English phrases. Thus, with the slowed-down version, language learners can easily follow the patterns and tune themselves into informal, natural English conversations.

Appendix 8: Transcripts and Analysis of Training Materials 09/2007 for Control Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr01_100	Marc & Donal are talking about their teaching experience.	Marc: it would mean [inhales]	Donal: yeah well it is a good point well erm wh- which one are you thinking of erm which	1. reduced 'you' to /je/, 2. clearly pronounced /v/ in 'of', 3. weak form of 'are', produced as /ə/, 4. collocation 'think of', 5. grammatical paradigm FS 'are you'	Tr_Sq_01_100 Tr_Dlg_01_100
Tr02_100	Topic: Donal having worked in DIT and Maynooth (colleges in Dublin area).	Marc: well but you you yo you you certainly made up for that	Donal: did not do any work so I went went to maynooth	1. lively stress intonation pattern, 2. collocation 'make up for'	Tr_Sq_02_100 Tr_Dlg_02_100
Tr03_100	Topic: The first time they moved out of home.	Marc: [inhales] yeah which is I suppose a a big thing for erm i mean I moved out of home when I was	Donal: and he moved out of home for the first time	1. double word stress on 'eighteen', 2. weak form of 'I', produced as /ə/, 3. collocation 'move out of', 4. grammatical	Tr_Sq_03_100 Tr_Dlg_03_100

		eighteen		paradigm FS 'I was'	
Tr04_100	Topic: Newcomers to Marc's street.	Marc: but I I have I have found that er I looked around my street this morning and I I fifty sixty percent of the people on the street I I have never met		1. intonation indicates a range or a more exact figure, 2. big intonational change, 3. barely distinguishable 'of the' (= uh th), 4. collocation 'percent of'	Tr_Sq_04_100 Tr_Dlg_04_100
Tr05_100	Topic: Lack of consideration of some of Marc's neighbours.	Marc: I I just know that they park their cars in in inappropriate ways for example I cannot walk down the street without having to walk on the road I knock on the door and I complain and	Donal: yeah hmm	1. intonation and extension of 'road' indicate further phrase to follow, 2. huge intonational change, 3. it is hard to hear that 'can't' is negative (and not 'can'. Negativity is made clear to the L1 listener by extending the '-n-' and following with 'without ...',	Tr_Sq_05_100 Tr_Dlg_05_100

				NOT by articulating the ‘-t’ in ‘can’t’) 4. extreme reduction of ‘without’ to /wiə-/ 5. reduction of ‘having to’ to /havn to/, 6. grammatical paradigm FS ‘have to’	
Tr06_100	Topic: How other cultures deal with heavy snowfalls.	Marc: there is a tran-trans-humanance e isn’t that what they call it the movement from, yeah from from one from erm	Darragh: er up the hills that is exactly what well that is exactly what the Italians would have been doing except in in a modern-day format	1. really fast speed of delivery, 2. huge intonational change, 3. reduced ‘would have been’ to /wudəbin/, 4. weak form of ‘what’, produced as /wə/, 5. grammatical paradigm FS ‘would have been’	Tr_Sq_06_100 Tr_Dlg_06_100
Tr07_100	Topic: Marc groans at Darragh’s bad pun. Mock seriousness.	Marc: aw dear God no Darragh aw Lord Darragh aw Darragh	Darragh: or was it it was not a an electoral ward they were in was it Ah well that is what you have to put up with	1. idiom ‘scrap the bottom of the barrel’, 2. fast speed in unstressed elements 3. expressiveness of this idiom fits into the	Tr_Sq_07_100 Tr_Dlg_07_100

		[inhales] you are you are scraping the bottom of the barrel sir now	[laughs]	expressive envelope: 'Aw, Donal - How could you tell such a sick joke?!?' 4. weak form of 'of', produced as /ə/	
Tr08_100	Topic: Donal tells how he guessed a girl's age correctly when he meant to flatter her by politely subtracting a few years from her apparent age.		Donal: a girl on the course asked what what age what age we thought she was you know an an an an an I and I sort of ah as you do on a on a on a particularly when a woman asks you I I brought her age down	1. fast speed, 2. reduced 'particularly' to /pətikəlju/	Tr_Sq_08_100 Tr_Dlg_08_100
Tr09_100	Topic: Marc's timetable at DBS college.	Marc: but I have also had five hours a week at the DBS that is nine hours ah on top of about twenty three at DIT	Donal: yeah yeah yeah yeah	1. heavily reduced 'I have also had' to /ivalsəd/, 2. semi- fixed frame ' ... a week', 3. weak form of 'a' and 'at the', produced respectively as /ə/ and /əd/,	Tr_Sq_09_100 Tr_Dlg_09_100

				4. grammatical paradigm FS 'have had'	
Tr10_100	Topic: Beginning of the recording. Marc welcomes Donal to the studio.	Marc: welcome Donal thanks for very much for [inhales] taking the time to come in and ah help us out on this ah project		1. reduced 'for' to /fə/ and 'taking' to /takin/, 2. collocation 'come in' and 'thanks (very much) for ...'	Tr_Sq_10_100 Tr_Dlg_10_100

Notes:

There are 3 types of file:

- (1) The original **SNIPPET**. This is a short recording which is linguistically interesting because of the way it is spoken by the L1 speaker.
- (2) The same snippet recorded as part of the speaker's **SEQUENCE**. This is a phrase, a sentence or similar utterance.
- (3) The snippet within the context of the **DIALOGUE** between both speakers.

Tr01_100 – The SNIPPET at 100% speed

Tr_Sq_01_100 – The snippet in the context of the speaker's turn [i.e., a SEQUENCE] within the dialogue – at 100% speed

Tr_Dlg_01_100 – The snippet within the context of the DIALOGUE in which it was recorded – at 100% speed

Appendix 9: Transcripts and Analysis of Training Materials 09/2007 for Test Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr01_100 Tr01_40	Marc & Donal are talking about their teaching experience.	Marc: it would mean [inhalés]	Donal: yeah well it is a good point well erm wh- which one are you thinking of erm which	1. reduced 'you' to /je/, 2. clearly pronounced /v/ in 'of', 3. weak form of 'are', produced as /ə/, 4. collocation 'think of', 5. grammatical paradigm FS 'are you'	Tr_Sq_01_100 Tr_Sq_01_40 Tr_Dlg_01_100 Tr_Dlg_01_40
Tr02_100 Tr02_40	Topic: Donal having worked in DIT and Maynooth (colleges in Dublin area).	Marc: well but you you yo you you certainly made up for that	Donal: did not do any work so I went went to maynooth	1. lively stress intonation pattern, 2. collocation 'make up for'	Tr_Sq_02_100 Tr_Sq_02_40 Tr_Dlg_02_100 Tr_Dlg_02_40
Tr03_100 Tr03_40	Topic: The first time they moved out of home.	Marc: [inhalés] yeah which is I suppose a a big thing for erm i mean I moved out of home when I was	Donal: and he moved out of home for the first time	1. double word stress on 'eighteen', 2. weak form of 'I', produced as /ə/, 3. collocation 'move out of', 4. grammatical	Tr_Sq_03_100 Tr_Sq_03_40 Tr_Dlg_03_100 Tr_Dlg_03_40

		eighteen		paradigm FS 'I was'	
Tr04_100 Tr04_40	Topic: Newcomers to Marc's street.	Marc: but I I have I have found that er I looked around my street this morning and I I fifty sixty percent of the people on the street I I have never met		1. intonation indicates a range or a more exact figure, 2. big intonational change, 3. barely distinguishable 'of the' (= uh th), 4. collocation 'percent of'	Tr_Sq_04_100 Tr_Sq_04_40 Tr_Dlg_04_100 Tr_Dlg_04_40
Tr05_100 Tr05_40	Topic: Lack of consideration of some of Marc's neighbours.	Marc: I I just know that they park their cars in in in- appropriate ways for example I cannot walk down the street without having to walk on the road I knock on the door and I complain and	Donal: yeah hmm	1. intonation and extension of 'road' indicate further phrase to follow, 2. huge intonational change, 3. it is hard to hear that 'can't' is negative (and not 'can'. Negativity is made clear to the L1 listener by extending the '-n-' and following with 'without	Tr_Sq_05_100 Tr_Sq_05_40 Tr_Dlg_05_100 Tr_Dlg_05_40

				...’, NOT by articulating the ‘-t’ in ‘can’t’) 4. extreme reduction of ‘without’ to /wiə-/ 5. reduction of ‘having to’ to /havn to/, 6. grammatical paradigm FS ‘have to’	
Tr06_100 Tr06_40	Topic: How other cultures deal with heavy snowfalls.	Marc: there is a tran-trans-humanance e isn’t that what they call it the movement from, yeah from from one from erm	Darragh: er up the hills that is exactly what well that is exactly what the Italians would have been doing except in in a modern-day format	1. really fast speed of delivery, 2. huge intonational change, 3. reduced ‘would have been’ to /wudəbin/, 4. weak form of ‘what’, produced as /wə/, 5. grammatical paradigm FS ‘would have been’	Tr_Sq_06_100 Tr_Sq_06_40 Tr_Dlg_06_100 Tr_Dlg_06_40
Tr07_100 Tr07_40	Topic: Marc groans at Darragh’s bad pun. Mock seriousness.	Marc: aw dear God no Darragh aw Lord Darragh aw Darragh	Darragh: or was it it was not a an electoral ward they were in was it Ah well that is what you have to put up with	1. idiom ‘scrap the bottom of the barrel’, 2. fast speed in unstressed elements 3. expressiveness of this idiom fits into the	Tr_Sq_07_100 Tr_Sq_07_40 Tr_Dlg_07_100 Tr_Dlg_07_40

		[inhales] you are you are scraping the bottom of the barrel sir now	[laughs]	expressive envelope: 'Aw, Donal - How could you tell such a sick joke?!?' 4. weak form of 'of', produced as /ə/	
Tr08_100 Tr08_40	Topic: Donal tells how he guessed a girl's age correctly when he meant to flatter her by politely subtracting a few years from her apparent age.		Donal: a girl on the course asked what what age what age we thought she was you know an an an an an I and I sort of ah as you do on a on a on a particularly when a woman asks you I I brought her age down	1. fast speed, 2. reduced 'particularly' to /pətikəlju/	Tr_Sq_08_100 Tr_Sq_08_40 Tr_Dlg_08_100 Tr_Dlg_08_40
Tr09_100 Tr09_40	Topic: Marc's timetable at DBS college.	Marc: but I have also had five hours a week at the DBS that is nine hours ah on top of about twenty three at DIT	Donal: yeah yeah yeah yeah	1. heavily reduced 'I have also had' to /ivalsəd/, 2. semi- fixed frame ' ... a week', 3. weak form of 'a' and 'at the', produced respectively as /ə/ and /əd/,	Tr_Sq_09_100 Tr_Sq_09_40 Tr_Dlg_09_100 Tr_Dlg_09_40

				4. grammatical paradigm FS 'have had'	
Tr10_100 Tr10_40	Topic: Beginning of the recording. Marc welcomes Donal to the studio.	Marc: welcome Donal thanks for very much for [inhales] taking the time to come in and ah help us out on this ah project		1. reduced 'for' to /fə/ and 'taking' to /takin/, 2. collocation 'come in' and 'thanks (very much) for ...'	Tr_Sq_10_100 Tr_Sq_10_40 Tr_Dlg_10_100 Tr_Dlg_10_40

Notes:

There are 3 types of file:

- (1) The original **SNIPPET**. This is a short recording which is linguistically interesting because of the way it is spoken by the L1 speaker.
- (2) The same snippet recorded as part of the speaker's **SEQUENCE**. This is a phrase, a sentence or similar utterance.
- (3) The snippet within the context of the **DIALOGUE** between both speakers.

Tr01_100 – The SNIPPET at 100% speed

Tr01_40 – The same snippet at 40% speed

Tr_Sq_01_100 – The snippet in the context of the speaker's turn [i.e., a SEQUENCE] within the dialogue – at 100% speed

Tr_Sq_01_40 – The same sequence at 40% speed

- Tr_Dlg_01_100 – The snippet within the context of the DIALOGUE in which it was recorded – at 100% speed
- Tr_Dlg_01_40 – The same dialogue at 40% speed

Appendix 10: Transcripts and Analysis of Training Materials 10/2007 for Control Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr11_100	Donal is talking about the mistakes his students made in the exam.		Donal: if I told you have I sent you the erm you know the mistakes that come up during the exam the ah ah ah I must have sent you	1. reduced 'have', 2. reduced 'you know', 3. weak form of 'that', produced as /ðə/, 4. word stress on key words, 5. chunk 'you know', 6. collocation 'come up', 7. grammatical paradigm FS 'have sent'	Tr_Sq_11_100 Tr_Dlg_11_100
Tr12_100	Topic: Marc's timetable.	Marc: which makes it ah you know quite difficult ah especially when you are working three nights a week ahm the second semester is usually ah ahhhh less	Donal: yeah yeah	1. reduced 'especially' to 'specially', 2. huge reduction of 'when you are', 3. indistinguishable pronoun 'you', 4. weak form of 'a', 5. individual word stress, 6. semi-fixed frame '... a week', 7. grammatical paradigm FS 'are working'	Tr_Sq_12_100 Tr_Dlg_12_100

Tr13_100	Topic: Donal's plans for the summer.	Marc: have you plans to go to Spain or oh yeah [laughing]	Donal: ahm I do not know you see I might go to ah I might just try and keep keep s-go somewhere because I do not want to go back to that what I was doing last summer as as you saw in in	1. reduced 'want to' to 'wanna', 2. fast speed of delivery, 3. collocation 'go back'	Tr_Sq_13_100 Tr_Dlg_13_100
Tr14_100	Topic: What they were doing in the 1980s.	Marc: I mean I remem- I do not know th-what you were doing in the eighties but I was the manager of a [inhales] of a training company in the nineteen eighties and it was very very hard		1. reduced 'do not know' to 'dunno', 2. reduced 'what you were' to /wə tjuə/, 3. weak form of 'in', pronounced as /n/, 4. grammatical paradigm FS 'were doing', 5. collocation 'in the eighties'	Tr_Sq_14_100 Tr_Dlg_14_100

Tr15_100	Topic: Marc thinks he is getting more irritable (=grumpy) as he grows older.	Marc: I I suppose there is one thing that really b- has begun to annoy me and eh ah I suppose it has something something to do with the fact that I am I am growing a bit older and it is that	Donal: that is the grumpy old man thing	1. large reduction of 'i' in 'with', 2. reduction of 'that I am' to 'that am', 3. individual word stress, 4. semi-fixed frame 'something to do with', 5. collocation 'a bit', 6. grammatical paradigm FS 'am growing'	Tr_Sq_15_100 Tr_Dlg_15_100
Tr16_100	Topic: They are talking about a man people can't understand because he insists on speaking only Irish.	Marc: but you see you see [inhales]	Donal: and but yeah he he was he was sw- generally speaking people just were not able to talk to him so he was doing things like he was singing	1. reduction of 'were not able to' to 'wənt ebl tə', 2. weak form of 'to him' produced as /tə im/, 3. rejectionist tone in retelling a story, especially in 'talk', 4. grammatical paradigm FS 'were not able to', 5. collocation 'talk to'	Tr_Sq_16_100 Tr_Dlg_16_100
Tr17_100	Topic: One of Ciaran's neighbours.		Ciaran: ah one of my neighbors one day	1. weak form of 'or' produced as /ə/	Tr_Sq_17_100 Tr_Dlg_17_100

			about six or seven years ago I had not seen him in years and ah he went he was a s-sixteen or seventeen year-old		
Tr18_100	Topic: A program about China shown on TV.	Dermot: it was an hour long and I I it showed various scenes of him experienci ng China and of course he goes for the weird and wonderful and the way-out		1. reduction of 'and of course' to 'anə course', 2. slight reduction of 'and' to 'an' in the idiom 'weird and wonderful', 3. collocation 'of course' and 'go for', 4. collocation ' weird and wonderful'	Tr_Sq_18_100 Tr_Dlg_18_100
Tr19_100	Topic: The possible reason for why men and women prefer different colours.	Dermot: and if there were a biological reason why ah women should prefer pink would not you think that it would ah-translate into the Belgian communit y as well		1. reduction of 'would not you' to 'wouldn you', 2. vowel in second 'would' is reduced (despite pause) because it is a grammatical paradigm FS 'would + V.', 3. collocation 'translate ... into ...' and 'as well'	Tr_Sq_19_100 Tr_Dlg_19_100

Tr20_100	Topic: Discussing an author, an actor and the character he plays in a film	Dermot: ah no no	Ciaran: that was ah John Le Carré and it was what do you call him ah Alex Guinness Obieone Kanobe who who was in that the the tha- that so-	1. reduced 'do you' to 'dju', 2. weak form of 'him' produced as /im/	Tr_Sq_20_100 Tr_Dlg_20_100
----------	-------------------------------------------------------------------------------------------------	--------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------	-------------------------------

Appendix 11: Transcripts and Analysis of Training Materials 10/2007 for Test Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr11_100 Tr11_40	Donal is talking about the mistakes his students made in the exam.		Donal: if I told you have I sent you the erm you know the mistakes that come up during the exam the ah ah ah I must have sent you	1. reduced 'have', 2. reduced 'you know', 3. weak form of 'that', produced as /ðə/, 4. word stress on key words, 5. chunk 'you know', 6. collocation 'come up', 7. grammatical paradigm FS 'have sent'	Tr_Sq_11_100 Tr_Sq_11_40 Tr_Dlg_11_100 Tr_Dlg_11_40
Tr12_100 Tr12_40	Topic: Marc's timetable.	Marc: which makes it ah you know quite difficult ah especially when you are working three nights a week ahm the second semester is usually ah ahhhh less	Donal: yeah yeah	1. reduced 'especially' to 'specially', 2. huge reduction of 'when you are', 3. indistinguishable pronoun 'you', 4. weak form of 'a', 5. individual word stress, 6. semi-fixed frame '... a week', 7. grammatical paradigm FS 'are working'	Tr_Sq_12_100 Tr_Sq_12_40 Tr_Dlg_12_100 Tr_Dlg_12_40

Tr13_100 Tr13_40	Topic: Donal's plans for the summer.	Marc: have you plans to go to Spain or oh yeah [laughing]	Donal: ahm I do not know you see I might go to ah I might just try and keep keep s-go somewhere because I do not want to go back to that what I was doing last summer as as you saw in in	1. reduced 'want to' to 'wanna', 2. fast speed of delivery, 3. collocation 'go back'	Tr_Sq_13_100 Tr_Sq_13_40 Tr_Dlg_13_100 Tr_Dlg_13_40
Tr14_100 Tr14_40	Topic: What they were doing in the 1980s.	Marc: I mean I remem- I do not know th-what you were doing in the eighties but I was the manager of a [inhalés] of a training company in the nineteen eighties and it was very very hard		1. reduced 'do not know' to 'dunno', 2. reduced 'what you were' to /wə tjuə/, 3. weak form of 'in', pronounced as /n/, 4. grammatical paradigm FS 'were doing', 5. collocation 'in the eighties'	Tr_Sq_14_100 Tr_Sq_14_40 Tr_Dlg_14_100 Tr_Dlg_14_40

Tr15_100 Tr15_40	Topic: Marc thinks he is getting more irritable (=grumpy) as he grows older.	Marc: I I suppose there is one thing that really b- has begun to annoy me and eh ah I suppose it has something something to do with the fact that I am I am growing a bit older and it is that	Donal: that is the grumpy old man thing	1. large reduction of 'i' in 'with', 2. reduction of 'that I am' to 'that am', 3. individual word stress, 4. semi-fixed frame 'something to do with', 5. collocation 'a bit', 6. grammatical paradigm FS 'am growing'	Tr_Sq_15_100 Tr_Sq_15_40 Tr_Dlg_15_100 Tr_Dlg_15_40
Tr16_100 Tr16_40	Topic: They are talking about a man people can't understand because he insists on speaking only Irish.	Marc: but you see you see [inhales]	Donal: and but yeah he he was he was sw- generally speaking people just were not able to talk to him so he was doing things like he was singing	1. reduction of 'were not able to' to 'wəɾnt ebl tə', 2. weak form of 'to him' produced as /tə im/, 3. rejectionist tone in retelling a story, especially in 'talk', 4. grammatical paradigm FS 'were not able to', 5. collocation 'talk to'	Tr_Sq_16_100 Tr_Sq_16_40 Tr_Dlg_16_100 Tr_Dlg_16_40
Tr17_100 Tr17_40	Topic: One of Ciaran's neighbours.		Ciaran: ah one of my neighbors one day	1. weak form of 'or' produced as /ə/	Tr_Sq_17_100 Tr_Sq_17_40 Tr_Dlg_17_100 Tr_Dlg_17_40

			about six or seven years ago I had not seen him in years and ah he went he was a s- sixteen or seventeen year-old		
Tr18_100 Tr18_40	Topic: A program about China shown on TV.	Dermot: it was an hour long and I I it showed various scenes of him experiencing China and of course he goes for the weird and wonderful and the way-out		1. reduction of 'and of course' to 'anə course', 2. slight reduction of 'and' to 'an' in the idiom 'weird and wonderful', 3. collocation 'of course' and 'go for', 4. collocation 'weird and wonderful'	Tr_Sq_18_100 Tr_Sq_18_40 Tr_Dlg_18_100 Tr_Dlg_18_40
Tr19_100 Tr19_40	Topic: The possible reason for why men and women prefer different colours.	Dermot: and if there were a biological reason why ah women should prefer pink would not you think that it would ah-translate into the Belgian community as well		1. reduction of 'would not you' to 'wouldn you', 2. vowel in second 'would' is reduced (despite pause) because it is a grammatical paradigm FS 'would + V.', 3. collocation 'translate ... into ...' and 'as well'	Tr_Sq_19_100 Tr_Sq_19_40 Tr_Dlg_19_100 Tr_Dlg_19_40

Tr20_100 Tr20_40	Topic: Discussing an author, an actor and the character he plays in a film	Dermot: ah no no	Ciaran: that was ah John Le Carré and it was what do you call him ah Alex Guinness Obieone Kanobe who who was in that the the tha- that so-	1. reduced 'do you' to 'dju', 2. weak form of 'him' produced as /im/	Tr_Sq_20_100 Tr_Sq_20_40 Tr_Dlg_20_100 Tr_Dlg_20_40
---------------------	-------------------------------------------------------------------------------------------------	--------------------------------------	----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	-------------------------------------------------------------------------------------	--------------------------------------------------------------

Appendix 12: Transcripts and Analysis of Training Materials 11/2007 for Control Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr21_100	The difficulties of doing business in the 1980s.	Marc: [inhales] it was so hard and yeah [inhales] well hard because it was just too expensive I mean flying to Frankfurt in a- for example if eh w-	Donal: yeah yeah well hard to get it done because the i- even though the ideas were not there or something is that what is that what you mean or even though the ideas were there rather yeah	1. reduced 'what you mean' to /wə tju min/, 2. weak form of non-lexical words 'is' and 'that', produced as /zðe/	Tr_Sq_21_100 Tr_Dlg_21_100
Tr22_100	The future of the Irish language.		Donal: and therefore pessimistic about about the future of the language as as a cl- [inhales] it is going to continue existing i- i- in a in a funny hybrid	1. reduced 'going to' to 'gonna', 2. grammatical paradigm FS 'is going to', 3. terminal stress via slowing-down speed	Tr_Sq_22_100 Tr_Dlg_22_100

			as they say		
Tr23_100	Marc worries about things – even at night.	<p>Marc:</p> <p>that is right yeah</p> <p>well I do have the occasional sleepless night about [inhales] how I might approach a certain problem or particular issue or</p>	<p>Darragh:</p> <p>there is a lot of nervous energy expended</p> <p>you know looking after them</p>	<p>1. sentence stress on the word 'do',</p> <p>2. collocation: 'sleepless night'</p>	Tr_Sq_23_100 Tr_Dlg_23_100
Tr24_100	Marc's cousin making his first confession (religious ceremony)	<p>Marc:</p> <p>he is a sil- [sniffs]</p> <p>Ah no he is a real nice guy but ah ah</p> <p>[sniffs]</p> <p>he had no idea what it meant ah but he had no idea what anything meant un- until he was about forty</p>	<p>Darragh:</p> <p>but ah ah he had not a clue what it meant obviously</p> <p>but he was six for God's sake I mean</p> <p>he could not have known what it meant</p> <p>yeah</p>	<p>1. reduced the syntactic paradigm 'could not have known' to 'couldna',</p> <p>2. weak form of 'what it', produced to /ə i/,</p> <p>3. grammatical paradigm FS 'could not have known'</p>	Tr_Sq_24_100 Tr_Dlg_24_100
Tr25_100	Marc is talking about the s EA food he	<p>Marc:</p> <p>the lunch time I had had ah</p>	<p>Darragh:</p>	<p>1. reduced 'are we', produced as /ər wi/</p>	Tr_Sq_25_100 Tr_Dlg_25_100

	had that day which he thinks made him sick. Darragh jokes that he had a SEE -food diet (i.e. he cannot resist food).	seafood but I got violently ill [laughs] go on go on	are we on the see-food diet [laughs]	2. (joking) question intonation, 3. collocation 'on a diet'	
Tr26_100	The food in restaurants in Dublin and a joke there (outstanding in his field / out, standing in his field)	Marc: and I am an expert in the field you know ah nothing will faze me now he was outstanding in his field ah that is an old – ah	Darragh: Aw God yeah literally you heard about the award winning farmer did not you he was on he was yeah ya ya	1. question intonation, 2. reduction: 'did not you' to: 'didn't yə', 3. collocation 'hear about'	Tr_Sq_26_100 Tr_Dlg_26_100
Tr27_100	Darragh is talking about his trip to Machu picchu (Inca city in Peru)		Darragh: and what happened was that ah myself and the gang that had gone to the jungle got back on a Sunday evening and meant to get the train on the Monday and I got back to the hotel on Sunday	1. reduced 'and I' to 'anna', 2. collocation 'get back to'	Tr_Sq_27_100 Tr_Dlg_27_100

			evening and met one of the - Isabel one of my sort of co-leaders		
Tr28_100	They are playing a video game. The female character is shot, but still alive, and the male speaker jokes that he will kill her properly.	Dermot: what there you go there well we will s- we will put an end to that yeah [laughs]	Marty: that is terrible look and I lay aw oh no I am still my heart is still beating my heart is still beating no	1. weak form on the grammatical word 'will', 2. 'put' becomes 'pud' and there is a large reduction of 'put an end' to: 'pud an en', 3. intonation indicates a threatening intention, 4. collocation 'put an end to'	Tr_Sq_28_100 Tr_Dlg_28_100
Tr29_100	Dermot is trying to find where Marty's character is located in the game.	Dermot: what can you see something aw OK where are you oh there you are	Marty: I have just hit some vibrator oh yeah [laughs] I am going to hide	1. reduced 'I am going to' to 'ain gonna', 2. grammatical paradigm FS 'am going to'	Tr_Sq_29_100 Tr_Dlg_29_100
Tr30_100	They are talking about the positions of their characters in the game.	Dermot: there is a red stuff oozing aw I wonder	Marty: OK I will go over and finish you off	1. 'you' is reduced to 'je'	Tr_Sq_30_100 Tr_Dlg_30_100

		<p>now could I naw I am I am right behind you you have to lower your sights girl yeah</p>			
--	--	--------------------------------------------------------------------------------------------------------------------------------------	--	--	--

Appendix 13: Transcripts and Analysis of Training Materials 11/2007 for Test Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr21_100 Tr21_40	The difficulties of doing business in the 1980s.	Marc: [inhales] it was so hard and yeah [inhales] well hard because it was just too expensive I mean flying to Frankfurt in a- for example if eh w-	Donal: yeah yeah well hard to get it done because the i- even though the ideas were not there or something is that what is that what you mean or even though the ideas were there rather yeah	1. reduced 'what you mean' to /wə tju min/, 2. weak form of non-lexical words 'is' and 'that', produced as /zðe/	Tr_Sq_21_100 Tr_Sq_21_40 Tr_Dlg_21_100 Tr_Dlg_21_40
Tr22_100 Tr22_40	The future of the Irish language.		Donal: and therefore pessimistic about about the future of the language as as a cl- [inhales] it is going to continue existing i- i- in a in a funny hybrid	1. reduced 'going to' to 'gonna', 2. grammatical paradigm FS 'is going to', 3. terminal stress via slowing-down speed	Tr_Sq_22_100 Tr_Sq_22_40 Tr_Dlg_22_100 Tr_Dlg_22_40

			as they say		
Tr23_100 Tr23_40	Marc worries about things – even at night.	Marc: that is right yeah well I do have the occasional sleepless night about [inhales] how I might approach a certain problem or particular issue or	Darragh: there is a lot of nervous energy expended you know looking after them	1. sentence stress on the word 'do', 2. collocation: 'sleepless night'	Tr_Sq_23_100 Tr_Sq_23_40 Tr_Dlg_23_100 Tr_Dlg_23_40
Tr24_100 Tr24_40	Marc's cousin making his first confession (religious ceremony)	Marc: he is a sil- [sniffs] Ah no he is a real nice guy but ah ah [sniffs] he had no idea what it meant ah but he had no idea what anything meant un- until he was about forty	Darragh: but ah ah he had not a clue what it meant obviously but he was six for God's sake I mean he could not have known what it meant yeah	1. reduced the syntactic paradigm 'could not have known' to 'couldna', 2. weak form of 'what it', produced to /ə i/, 3. grammatical paradigm FS 'could not have known'	Tr_Sq_24_100 Tr_Sq_24_40 Tr_Dlg_24_100 Tr_Dlg_24_40
Tr25_100 Tr25_40	Marc is talking about the s EA food he	Marc: the lunch time I had had ah	Darragh:	1. reduced 'are we', produced as /ər wi/	Tr_Sq_25_100 Tr_Sq_25_40 Tr_Dlg_25_100 Tr_Dlg_25_40

	had that day which he thinks made him sick. Darragh jokes that he had a SEE -food diet (i.e. he cannot resist food).	seafood but I got violently ill [laughs] go on go on	are we on the see-food diet [laughs]	2. (joking) question intonation, 3. collocation 'on a diet'	
Tr26_100 Tr26_40	The food in restaurants in Dublin and a joke there (outstanding in his field / out, standing in his field)	Marc: and I am an expert in the field you know ah nothing will faze me now he was outstanding in his field ah that is an old – ah	Darragh: Aw God yeah literally you heard about the award winning farmer did not you he was on he was yeah ya ya	1. question intonation, 2. reduction: 'did not you' to: 'didn't yə', 3. collocation 'hear about'	Tr_Sq_26_100 Tr_Sq_26_40 Tr_Dlg_26_100 Tr_Dlg_26_40
Tr27_100 Tr27_40	Darragh is talking about his trip to Machu picchu (Inca city in Peru)		Darragh: and what happened was that ah myself and the gang that had gone to the jungle got back on a Sunday evening and meant to get the train on the Monday and I got back to the hotel on Sunday	1. reduced 'and I' to 'anna', 2. collocation 'get back to'	Tr_Sq_27_100 Tr_Sq_27_40 Tr_Dlg_27_100 Tr_Dlg_27_40

			evening and met one of the - Isabel one of my sort of co-leaders		
Tr28_100 Tr28_40	They are playing a video game. The female character is shot, but still alive, and the male speaker jokes that he will kill her properly.	Dermot: what there you go there well we will s- we will put an end to that yeah [laughs]	Marty: that is terrible look and I lay aw oh no I am still my heart is still beating my heart is still beating no	1. weak form on the grammatical word 'will', 2. 'put' becomes 'pud' and there is a large reduction of 'put an end' to: 'pud an en', 3. intonation indicates a threatening intention, 4. collocation 'put an end to'	Tr_Sq_28_100 Tr_Sq_28_40 Tr_Dlg_28_100 Tr_Dlg_28_40
Tr29_100 Tr29_40	Dermot is trying to find where Marty's character is located in the game.	Dermot: what can you see something aw OK where are you oh there you are	Marty: I have just hit some vibrator oh yeah [laughs] I am going to hide	1. reduced 'I am going to' to 'ain gonna', 2. grammatical paradigm FS 'am going to'	Tr_Sq_29_100 Tr_Sq_29_40 Tr_Dlg_29_100 Tr_Dlg_29_40
Tr30_100 Tr30_40	They are talking about the positions of their characters in the game.	Dermot: there is a red stuff oozing aw I wonder	Marty: OK I will go over and finish you off	1. 'you' is reduced to 'je'	Tr_Sq_30_100 Tr_Sq_30_40 Tr_Dlg_30_100 Tr_Dlg_30_40

		<p>now could I naw I am I am right behind you you have to lower your sights girl yeah</p>			
--	--	-----------------------------------------------------------------------------------------------------------------------------------------------	--	--	--

Appendix 14: Transcripts and Analysis of Training Materials 12/2007 for Control Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr31_100	Marc is talking about Donal's approach to teaching.	Marc: you wanted to to be more intricate you wanted to deal with many of the issues that concern you that are particularly interesting particularly fascinating which may not be of erm		1. reduced 'particularly' to 'particuly' – in both examples, 2. weak form of non-lexical words 'are', produced as /ə/	Tr_Sq_31_100 Tr_Dlg_31_100
Tr32_100	Marc complains about not having enough time.	Marc: you are a very good cartoonist or illustrator as I said earlier on and erm it is a pity you do not do a bit more but like myself you find yourself having so much to do and so little time		1. weak form of non-lexical word 'a', produced as /ə/, 2. reduced 'you', 3. reduced 'do not do' to 'dondo', 4. semi-fixed frame 'it is a pity ...', 5. collocation 'a bit'	Tr_Sq_32_100 Tr_Dlg_32_100

Tr33_100	Donal enjoyed his work last year.	Marc: last year I thought you di- had a fairly terrific set up I thought you did a very good job ahm yeah yeah yeah	Donal: it was enjoyable I did my best yeah well well we were wo- yeah so I mean it was it was enjoyable and it was certainly great to have the people well yourself and Conor were there and [sniffs]	1. stress on the non- lexical word 'was'	Tr_Sq_33_100 Tr_Dlg_33_100
Tr34_100	Topic: Newcomers to Marc's street.	Marc: I do not know them [inhales] and the(y) people have come in in the last say five six seven or eight years I do not know who they are I do not know what they ahhhw	Donal: yeah yeah yeah	1. reduction of the non- lexical word 'have', 2. 'seven or eight' reduced to /sevn ə e/ 3. five, six, seven and eight are parallel structure to try to find the correct time, 4. collocation 'come in', 5. grammatical paradigm	Tr_Sq_34_100 Tr_Dlg_34_100

				FS 'have come'	
Tr35_100	They are talking about people's attitude to a man who insists on speaking only Irish (only a minority speaks Irish well, even though it is the first official language).	Marc: and how how did he get on	Donal: he tried to traverse Ireland and not speak any English just speak Irish the whole time and basically people were generally quite nice but in Dublin he was kicked out of the pub	1. reduced 'basically', produced as 'basicly', 2. 'generally' is reduced to 'generly', 3. word and stress patterns for emphasis	Tr_Sq_35_100 Tr_Dlg_35_100
Tr36_100	Marc likes a quiet room for thinking.	Marc: well ah this is actually very comfortable I would love to have a meditation s- room like this where there is no sound coming from outside it is so peaceful in here my goodness me [inhales] I have a back garden yeah	Darragh: well do you have a back garden ah is there a yard	1. question intonation, 2. reduction of the non- lexical word 'do'	Tr_Sq_36_100 Tr_Dlg_36_100
Tr37_100	Topic: Marc's	Marc: ah he was a-	Darragh:	1. reduced 'would you'	Tr_Sq_37_100 Tr_Dlg_37_100

	cousin suffered at school because of a mistake he made at his first confession.	at school he was ah I saw him actually stand in the primary school classroom with a dunce's cap on would you believe it yeah	ah bru-the-that just reinforces the situation you know	to 'wouldje', 2. chunk 'would you believe it'	
Tr38_100	Topic: Buying a camera. (Irish and British L1 speakers reading out a dialogue exacted from the textbook of junior secondary school).	Dermot: good morning can I help you we have many cameras here some are made in China and some are made in other countries in Japan 5,000 <i>yuan</i> this is a digital camera sure what about those ones	Marty: oh yes please I would like to buy a camera oh this one looks very nice where is it made oh right and how much is it oh wow that is too expensive ahm I can not afford it do you have an ordinary one well this	1. weak form on the grammatical word 'are', produced as /ə/, then 'some are' is reduced to /sΛmə/, 2. weak form on the non-lexical word 'in', 'made in' produced as 'madn', especially in the second example, 3. collocation 'be made in'	Tr_Sq_38_100 Tr_Dlg_38_100

		yes it is made in Shanghai	one looks very good and the price is OK is it made in China OK I will take this one		
Tr39_100	The same topic as No. 38, but a more natural, unscripted, interactive dialogue is produced.	Dermot: yeah would you like a digital camera or a film camera oh g- we have got a full range here but wh- what sort of photographs would you like to take OK I would recommend a a moderate zoom lens this brand is pretty good most of our customers are very pleased with that they	Marty: hi good morning ahm I would like to buy a camera no digital please yeah well I I like taking portraits so what would be the best camera for something like that yeah OK and and how many pixels does that	1. reduced 'moderate' to 'modret'	Tr_Sq_39(40)_ 100 Tr_Dlg_39(40)_ 100

		<p>it has five megapixels</p> <p>well it is about 300 <i>euros</i></p> <p>yes it has a moderate wide angle</p> <p>very good</p>	<p>have oh yeah and is it very expensive</p> <p>hmmm OK and does it have a wide angle as well OK well that sounds great yeah I think I will I will take that one please</p>		
Tr40_100	The same dialogue as No. 39.	<p>Dermot: most of our customers are very pleased with that they</p> <p>it has five megapixels</p>	<p>Marty:</p> <p>OK and and how many pixels does that have oh yeah</p>	1. followed by the ending phoneme – ‘s’ in ‘pixels’, ‘does’ is produced as /sΛz/. It is the NON-lexical word which is reduced	Tr_Sq_39(40)_100 Tr_Dlg_39(40)_100

**Appendix 15: Transcripts and Analysis of Training Materials 12/2007 for
Test Group**

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr31_100 Tr31_40	Marc is talking about Donal's approach to teaching.	Marc: you wanted to to be more intricate you wanted to to deal with many of the issues that concern you that are particularly interesting particularly fascinating which may not be of erm		1. reduced 'particularly' to 'particuly' – in both examples, 2. weak form of non-lexical words 'are', produced as /ə/	Tr_Sq_31_100 Tr_Sq_31_40 Tr_Dlg_31_100 Tr_Dlg_31_40
Tr32_100 Tr32_40	Marc complains about not having enough time.	Marc: you are a very good cartoonist or illustrator as I said earlier on and erm it is a pity you do not do a bit more but like myself you find yourself having so much to do and so little time		1. weak form of non-lexical word 'a', produced as /ə/, 2. reduced 'you', 3. reduced 'do not do' to 'dondo', 4. semi-fixed frame 'it is a pity ...', 5. collocation 'a bit'	Tr_Sq_32_100 Tr_Sq_32_40 Tr_Dlg_32_100 Tr_Dlg_32_40

Tr33_100 Tr33_40	Donal enjoyed his work last year.	Marc: last year I thought you di- had a fairly terrific set up I thought you did a very good job ahm yeah yeah yeah	Donal: it was enjoyable I did my best yeah well well we were wo- yeah so I mean it was it was enjoyable and it was certainly great to have the people well yourself and Conor were there and [sniffs]	1. stress on the non- lexical word 'was'	Tr_Sq_33_100 Tr_Sq_33_40 Tr_Dlg_33_100 Tr_Dlg_33_40
Tr34_100 Tr34_40	Topic: Newcomers to Marc's street.	Marc: I do not know them [inhales] and the(y) people have come in in the last say five six seven or eight years I do not know who they are I do not know what they ahhhw	Donal: yeah yeah yeah	1. reduction of the non- lexical word 'have', 2. 'seven or eight' reduced to /sevn ə e/ 3. five, six, seven and eight are parallel structure to try to find the correct time, 4. collocation 'come in', 5. grammatical paradigm	Tr_Sq_34_100 Tr_Sq_34_40 Tr_Dlg_34_100 Tr_Dlg_34_40

				FS 'have come'	
Tr35_100 Tr35_40	They are talking about people's attitude to a man who insists on speaking only Irish (only a minority speaks Irish well, even though it is the first official language).	Marc: and how how did he get on	Donal: he tried to traverse Ireland and not speak any English just speak Irish the whole time and basically people were generally quite nice but in Dublin he was kicked out of the pub	1. reduced 'basically', produced as 'basicly', 2. 'generally' is reduced to 'generly', 3. word and stress patterns for emphasis	Tr_Sq_35_100 Tr_Sq_35_40 Tr_Dlg_35_100 Tr_Dlg_35_40
Tr36_100 Tr36_40	Marc likes a quiet room for thinking.	Marc: well ah this is actually very comfortable I would love to have a meditation s- room like this where there is no sound coming from outside it is so peaceful in here my goodness me [inhales] I have a back garden yeah	Darragh: well do you have a back garden ah is there a yard	1. question intonation, 2. reduction of the non- lexical word 'do'	Tr_Sq_36_100 Tr_Sq_36_40 Tr_Dlg_36_100 Tr_Dlg_36_40
Tr37_100 Tr37_40	Topic: Marc's	Marc: ah he was a-	Darragh:	1. reduced 'would you'	Tr_Sq_37_100 Tr_Sq_37_40

	cousin suffered at school because of a mistake he made at his first confession.	at school he was ah I saw him actually stand in the primary school classroom with a dunce's cap on would you believe it yeah	ah bru-the-that just reinforces the situation you know	to 'wouldje', 2. chunk 'would you believe it'	Tr_Dlg_37_100 Tr_Dlg_37_40
Tr38_100 Tr38_40	Topic: Buying a camera. (Irish and British L1 speakers reading out a dialogue exacted from the textbook of junior secondary school).	Dermot: good morning can I help you we have many cameras here some are made in China and some are made in other countries in Japan 5,000 <i>yuan</i> this is a digital camera sure what about those ones	Marty: oh yes please I would like to buy a camera oh this one looks very nice where is it made oh right and how much is it oh wow that is too expensive ahm I can not afford it do you have an ordinary one well this	1. weak form on the grammatical word 'are', produced as /ə/, then 'some are' is reduced to /sΛmə/, 2. weak form on the non-lexical word 'in', 'made in' produced as 'madn', especially in the second example, 3. collocation 'be made in'	Tr_Sq_38_100 Tr_Sq_38_40 Tr_Dlg_38_100 Tr_Dlg_38_40

		yes it is made in Shanghai	one looks very good and the price is OK is it made in China OK I will take this one		
Tr39_100 Tr39_40	The same topic as No. 38, but a more natural, unscripted, interactive dialogue is produced.	Dermot: yeah would you like a digital camera or a film camera oh g- we have got a full range here but wh- what sort of photographs would you like to take OK I would recommend a a moderate zoom lens this brand is pretty good most of our customers are very pleased with that they	Marty: hi good morning ahm I would like to buy a camera no digital please yeah well I I like taking portraits so what would be the best camera for something like that yeah OK and and how many pixels does that	1. reduced 'moderate' to 'modret'	Tr_Sq_39(40)_ 100 Tr_Sq_39(40)_ 40 Tr_Dlg_39(40)_ 100 Tr_Dlg_39(40)_ 40

		<p>it has five megapixels</p> <p>well it is about 300 <i>euros</i></p> <p>yes it has a moderate wide angle</p> <p>very good</p>	<p>have oh yeah and is it very expensive</p> <p>hmmm OK and does it have a wide angle as well OK well that sounds great yeah I think I will I will take that one please</p>		
Tr40_100 Tr40_40	The same dialogue as No. 39.	<p>Dermot: most of our customers are very pleased with that they</p> <p>it has five megapixels</p>	<p>Marty:</p> <p>OK and and how many pixels does that have oh yeah</p>	1. followed by the ending phoneme – ‘s’ in ‘pixels’, ‘does’ is produced as /sΛz/. It is the NON-lexical word which is reduced	<p>Tr_Sq_39(40)_100 Tr_Sq_39(40)_40 Tr_Dlg_39(40)_100 Tr_Dlg_39(40)_40</p>

Appendix 16: Transcripts and Analysis of Training Materials 03/2008 for Control Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr41_100	Darragh is talking about his experience of a recording for a radio show.	Marc: oh really yeah	Darragh: (a) jot down some ideas and what you like to do what sort of format for a show would be Irish music whatever and then just record it and and send it around to a couple of the places they are they are often looking for people ah	1. reduced 'they are', produced as 'the', 2. hyper-intensity on 'often', 3. collocation: 'look for', 4. reduced form of 'for' with schwa, produced as /fə/	Tr_Sq_41_100 Tr_Dlg_41_100
Tr42_100	They are talking about an election and joking about well-known families from the west of Ireland.	Marc: [laughs] no they are Jod Joyces the Joy... the Joyces yeah [laughs] well do you know what was really shocking	Darragh: [laughs] you know the you know the fighting tribe the- are they Limerick or Galway Wards [inhales] aw the Joyces yeah Joyces [laughs]	1. reduced 'do you', produced as 'dje', 2. weak form of 'what' and 'was', produced as /wəz/	Tr_Sq_42_100 Tr_Dlg_42_100

		ah			
Tr43_100	Topic: Marc's job.	Marc: for the first time in my erm sort of ah not very illustrious academic career it didn- true	Darragh: yeah but you got to remember that it ah A it is your busiest time of the year because you you I mean you have just come off a ah downtime in the summer and then the second thing is that	1.reduction of the non- lexical word 'of'	Tr_Sq_43_100 Tr_Dlg_43_100
Tr44_100	They are talking about the way of paying back a loan.	Marc: well I am actually happy in a way that it is is being has become a little bit drawn-out because financially it would be to my favour shall we say to to leave it for another couple of months	Darragh: em I think it also ge- gives you you know your time your head time to get	1. collocation: 'to one's favour', 2. rising intonation indicates further information to follow	Tr_Sq_44_100 Tr_Dlg_44_100

		so I can			
Tr45_100	Topic: Darragh jokes that Marc has prepared a list of questions to ask him.	Marc: [inhales] no I would have no questions it is just why no no no not at all [inhales] La- this is great ah	Darragh: you want to get through do you do you not I do not like you do not have any cause you would not see the bloody list in here anyway	1. reduced 'have', produced as /əv/	Tr_Sq_45_100 Tr_Dlg_45_100
Tr46_100	Topic: Marc tells a funny story about his cousin making his first confession (religious ceremony).	Marc: I have no idea I think he that is the only thing that came into his head it was probably one of those words ah he heard it yeah yeah [inhales] well we had a we had a priest in that school at the same time who was a a	Darragh: he had heard it probably he ma- he he may may have heard it	1. collocation: 'come into', 2. huge reduction of 'only', 3. weak form of 'his', produced as /iz/	Tr_Sq_46_100 Tr_Dlg_46_100
Tr47_100	Topic: Marc complains that some	Marc: somebody parked his car	Darragh: oh you hate	1. weak form of 'what', produced as	Tr_Sq_47_100 Tr_Dlg_47_100

	people park their cars in an inappropriate way.	[inhales] in such a way that I I had to walk out on onto the main car ferry road to get around this car so say I am not going to let it go this is just too too much too often and [inhales]	this I know what you are going to say yeah yeah you hate this yeah yeah emh	/ət/, 2. reduction of 'you', produced as /je/, weak form of 'are', then 'you are' produced as /je ə/, 3. 'going to' reduced to 'gonna', 4. grammatical paradigm FS 'are going to'	
Tr48_100	Topic: Making a call. (They are reading out a dialogue exacted from the textbook of junior secondary school, but in a more natural way).	Dermot: yes please I would like to book two rooms for the coming weekend oh just two Friday	Marty: hello the Rose Hotel can I help you next weekend you say well how many nights OK two nights from Friday to Saturday	1. weak form on the subject 'I', produced as /ə/, 2. weak form on the non-lexical word 'for', produced as /fə/, 3. collocation 'would like to'	Tr_Sq_48_100 Tr_Dlg_48_100

		<p>two double rooms please</p> <p>what's the price of a double room</p> <p>I see can I book the rooms now please</p> <p>Rick Smith</p>	<p>Friday the 20th and do you need single rooms or double</p> <p>right two doubles for Friday and Saturday July 20th and 21st</p> <p>well it is 400 <i>yuan</i> per night with breakfast</p> <p>certainly what is your full name please</p>		
Tr49_100	The same topic as No. 48, but a more natural interactive dialogue is produced.	<p>Dermot:</p> <p>yes ah I would like to book two rooms for the ah weekend please</p> <p>we would like one double room and one twin</p>	<p>Marty: hello the rose hotel can I help you</p> <p>right OK and what kind of a room would you like</p> <p>one- OK one double and one twin</p>	1. reduced 'breakfast' to /brekst/	Tr_Sq_49(50)_100 Tr_Dlg_49(50)_100

		<p>oh definitely non- smoking we are both non- smokers yeah</p> <p>both rooms yes yeah</p> <p>yeah Friday and Saturday</p> <p>[inhales] yes I would like a continenta l breakfast please</p> <p>for the other party as well</p> <p>yes</p> <p>thank you very much</p>	<p>do you have any preferences for smoking or non- smoking or a view non-smoking is that for both of the rooms</p> <p>OK and you said this weekend so that is the 20th and 21st</p> <p>right OK would you know whether you want breakfast or not included</p> <p>continental breakfast for ah ah also for both th- parties yeah great OK well let me just write that down so Friday Saturday July 20th and 21st one twin one double great thank you</p>		
--	--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--

		thank you very much			
		bye			
Tr50_100	The same dialogue as No. 49.	Dermot: I would like a continenta l breakfast please for the other party as well yes	Marty: continental breakfast for ah ah also for both th- parties yeah	1. reduced ‘as’ to /əz/, then ‘party as’ is produced as /pa:ti əz/, 2. collocation ‘as well’	Tr_Sq_49(50)_ 100 Tr_Dlg_49(50)_ 100

Appendix 17: Transcripts and Analysis of Training Materials 03/2008 for Test Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr41_100 Tr41_40	Darragh is talking about his experience of a recording for a radio show.	Marc: oh really yeah	Darragh: (a) jot down some ideas and what you like to do what sort of format for a show would be Irish music whatever and then just record it and and send it around to a couple of the places they are they are often looking for people ah	1. reduced 'they are', produced as 'the', 2. hyper-intensity on 'often', 3. collocation: 'look for', 4. reduced form of 'for' with schwa, produced as /fə/	Tr_Sq_41_100 Tr_Sq_41_40 Tr_Dlg_41_100 Tr_Dlg_41_40
Tr42_100 Tr42_40	They are talking about an election and joking about well-known families from the west of Ireland.	Marc: [laughs] no they are Jod Joyces the Joy... the Joyces yeah [laughs] well do you know what was really shocking	Darragh: [laughs] you know the you know the fighting tribe the- are they Limerick or Galway Wards [inhales] aw the Joyces yeah Joyces [laughs]	1. reduced 'do you', produced as 'dje', 2. weak form of 'what' and 'was', produced as /wəz/	Tr_Sq_42_100 Tr_Sq_42_40 Tr_Dlg_42_100 Tr_Dlg_42_40

		ah			
Tr43_100 Tr43_40	Topic: Marc's job.	Marc: for the first time in my erm sort of ah not very illustrious academic career it didn- true	Darragh: yeah but you got to remember that it ah A it is your busiest time of the year because you you I mean you have just come off a ah downtime in the summer and then the second thing is that	1.reduction of the non-lexical word 'of'	Tr_Sq_43_100 Tr_Sq_43_40 Tr_Dlg_43_100 Tr_Dlg_43_40
Tr44_100 Tr44_40	They are talking about the way of paying back a loan.	Marc: well I am actually happy in a way that it is is being has become a little bit drawn-out because financially it would be to my favour shall we say to to leave it for another couple of months	Darragh: em I think it also ge- gives you you know your time your head time to get	1. collocation: 'to one's favour', 2. rising intonation indicates further information to follow	Tr_Sq_44_100 Tr_Sq_44_40 Tr_Dlg_44_100 Tr_Dlg_44_40

		so I can			
Tr45_100 Tr45_40	Topic: Darragh jokes that Marc has prepared a list of questions to ask him.	Marc: [inhales] no I would have no questions it is just why no no no not at all [inhales] La- this is great ah	Darragh: you want to get through do you do you not I do not like you do not have any cause you would not see the bloody list in here anyway	1. reduced 'have', produced as /əv/	Tr_Sq_45_100 Tr_Sq_45_40 Tr_Dlg_45_100 Tr_Dlg_45_40
Tr46_100 Tr46_40	Topic: Marc tells a funny story about his cousin making his first confession (religious ceremony).	Marc: I have no idea I think he that is the only thing that came into his head it was probably one of those words ah he heard it yeah yeah [inhales] well we had a we had a priest in that school at the same time who was a a	Darragh: he had heard it probably he ma- he he may may have heard it	1. collocation: 'come into', 2. huge reduction of 'only', 3. weak form of 'his', produced as /iz/	Tr_Sq_46_100 Tr_Sq_46_40 Tr_Dlg_46_100 Tr_Dlg_46_40
Tr47_100 Tr47_40	Topic: Marc complains that some	Marc: somebody parked his car	Darragh: oh you hate	1. weak form of 'what', produced as	Tr_Sq_47_100 Tr_Sq_47_40 Tr_Dlg_47_100 Tr_Dlg_47_40

	people park their cars in an inappropriate way.	[inhales] in such a way that I I had to walk out on onto the main car ferry road to get around this car so say I am not going to let it go this is just too too much too often and [inhales]	this I know what you are going to say yeah yeah you hate this yeah yeah emh	/ət/, 2. reduction of 'you', produced as /je/, weak form of 'are', then 'you are' produced as /je ə/, 3. 'going to' reduced to 'gonna', 4. grammatical paradigm FS 'are going to'	
Tr48_100 Tr48_40	Topic: Making a call. (They are reading out a dialogue exacted from the textbook of junior secondary school, but in a more natural way).	Dermot: yes please I would like to book two rooms for the coming weekend oh just two Friday	Marty: hello the Rose Hotel can I help you next weekend you say well how many nights OK two nights from Friday to Saturday	1. weak form on the subject 'I', produced as /ə/, 2. weak form on the non-lexical word 'for', produced as /fə/, 3. collocation 'would like to'	Tr_Sq_48_100 Tr_Sq_48_40 Tr_Dlg_48_100 Tr_Dlg_48_40

		<p>two double rooms please</p> <p>what's the price of a double room</p> <p>I see can I book the rooms now please</p> <p>Rick Smith</p>	<p>Friday the 20th and do you need single rooms or double</p> <p>right two doubles for Friday and Saturday July 20th and 21st</p> <p>well it is 400 <i>yuan</i> per night with breakfast</p> <p>certainly what is your full name please</p>		
Tr49_100 Tr49_40	The same topic as No. 48, but a more natural interactive dialogue is produced.	<p>Dermot:</p> <p>yes ah I would like to book two rooms for the ah weekend please</p> <p>we would like one double room and one twin</p>	<p>Marty: hello the rose hotel can I help you</p> <p>right OK and what kind of a room would you like</p> <p>one- OK one double and one twin</p>	1. reduced 'breakfast' to /brekst/	<p>Tr_Sq_49(50)_100</p> <p>Tr_Sq_49(50)_40</p> <p>Tr_Dlg_49(50)_100</p> <p>Tr_Dlg_49(50)_40</p>

		<p>oh definitely non- smoking we are both non- smokers yeah</p> <p>both rooms yes yeah</p> <p>yeah Friday and Saturday</p> <p>[inhales] yes I would like a continenta l breakfast please</p> <p>for the other party as well</p> <p>yes</p> <p>thank you very much</p>	<p>do you have any preferences for smoking or non- smoking or a view non-smoking is that for both of the rooms</p> <p>OK and you said this weekend so that is the 20th and 21st</p> <p>right OK would you know whether you want breakfast or not included</p> <p>continental breakfast for ah ah also for both th- parties yeah great OK well let me just write that down so Friday Saturday July 20th and 21st one twin one double great thank you</p>		
--	--	-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--	--

		thank you very much			
		bye			
Tr50_100 Tr50_40	The same dialogue as No. 49.	Dermot: I would like a continenta l breakfast please for the other party as well yes	Marty: continental breakfast for ah ah also for both th- parties yeah	1. reduced 'as' to /əz/, then 'party as' is produced as /pa:ti əz/, 2. collocation 'as well'	Tr_Sq_49(50)_ 100 Tr_Sq_49(50)_ 40 Tr_Dlg_49(50)_ 100 Tr_Dlg_49(50)_ 40

Appendix 18: Transcripts and Analysis of Training Materials 04/2008 for Control Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr51_100	Marc is talking about his teaching timetable.	Marc: well ah ah I teach you see I have been teaching about four hours a week there but	Donal: ah once a week yeah	1. reduced 'I have', produced as 'Iv', 2. grammatical paradigm FS 'have been doing', 3. semi-fixed frame '... a week'	Tr_Sq_51_100 Tr_Dlg_51_100
Tr52_100	Topic: The benefits of living in the countryside.	Marc: can li- b- b- mainTAINS that she can live quite comfortably in the countryside and be and be anonymous do you know what I mean ahm	Donal: yeah well yeah well I suppose you can yeah yeah	1. reduced 'do you know', produced as 'dje know', 2. weak form of 'what I mean', 3. chunk 'do you know what I mean'	Tr_Sq_52_100 Tr_Dlg_52_100
Tr53_100	Marc is talking about the guy who shares the same flat with him.	Marc: when sitting at the table with sort of five knives and five forks and whatever you know his was a	Donal: [laughs]	1. two 'you know' are chunks, used as a word-filler, 2. the first 'you know' reduced to 'je know', 3. 'was a' reduced to	Tr_Sq_53_100 Tr_Dlg_53_100

		a you know good ould I do not know		/wəzə/	
Tr54_100	Topic: Donal's boss.	Marc: wow	Donal: and he is basically erm you know he travels out to india travels out to out to pakistan and and then he gets a lot of irish students as well it is pretty much fifty fifty you know he is he is doing very well	1. chunk 'you know', 2. weak form of 'he', produced as /i/, 3. 'out' reduced to /əu/, 4. collocation 'travel out to'	Tr_Sq_54_100 Tr_Dlg_54_100
Tr55_100	People's attitude to a man who insists on speaking only Irish (only 10% of population speaks Irish fluently).	Marc: really are you serious yeah	Donal: they they said if you do not speak fuckin english well you are out you are out of the pub and th- this is the front page of ehm of of th- the evening herald	1. slow speed of delivery, 2. wide tonal range, 3. question intonation pattern, 4. 'you' reduced to 'jə', 5. chunk 'are you serious'	Tr_Sq_55_100 Tr_Dlg_55_100
Tr56_100	Discussion of Google Map views	Marc: emhm	Darragh: somebody has taken eh	1. reduced 'it is', produced as	Tr_Sq_56_100 Tr_Dlg_56_100

	of geological features.	It is fantastic is not it yeah well I used to love going over at visit Elena	the Massif Central and and tipped it vertically I can just see rows and rows of eh	'its', 2. tag question intonation, and answer	
Tr57_100	Topic: Seeing a doctor. (They are reading out a dialogue extracted from the textbook of junior secondary school, but in a more natural way).	Dermot: Good morning, doctor. I have got a bad cold, I have a headache and a running nose. OK. Do you have any quicker ways? Oho, no! It is a little painful to get an injection. I prefer to take medicine. eh Is it good to drink much water?	Marty: Good morning. What is your trouble, young man? Oh, I see. Let me check you. Oh, it is necessary for you to take some Chinese medicine. em, yes. You can get an injection. Oh, sure! It is It is also important to get a lot of rest.	1. weak form on the subject pronoun 'I', reduction of 'have', thus 'I have' is produced as 'əv', 2. weak form on three occurrences of 'a', 3. collocation 'have got a bad cold', 'have a headache' and 'have a running nose'	Tr_Sq_57(58)_100 Tr_Dlg_57(58)_100

		Thank you, doctor. I will do as you say.			
Tr58_100	The same dialogue as No. 57.	Dermot: eh Is it good to drink much water? Thank you, doctor. I will do as you say.	Marty: Oh, sure! It is It is also important to get a lot of rest.	1. weak form on 'I will', produced as /əl/, 2. weak form on 'as', produced as /əs/	Tr_Sq_57(58)_ 100 Tr_Dlg_57 (58)_100
Tr59_100	The same topic as No. 57 and 58, but a more natural interactive dialogue is produced.	Dermot: good morning, doctor. oh, I woke up yesterday with a pain in my throat. and I am afraid it is going to go down to my chest. so I would like to take some preventive measures. 'ah'	Marty: oh, good morning. how are you today? what can I do for you? aw right. yeah, sure. well, first of all, let me just have a look in your in your throat if I may can you say 'ah' OK yeah oh, I I see it is it is quite	1. weak form on the article word 'an' – joined with 'anti-biotic', 2. tag question intonation	Tr_Sq_59(60)_ 100 Tr_Dlg_59(60)_ 100

		<p>you could not give me an anti- biotic, could you</p> <p>ooh ooh erm that sounds painful but if it will speed things up I will take your advice</p>	<p>infected actually ah very red so erm yeah, you you need to take some Chinese medicine</p> <p>no, I I prefer not to it is not very healthy for you so well, I I suppose I could give you erm an injection that would work a bit faster</p> <p>[laughs]</p> <p>OK all right, then</p>		
Tr60_100	The same dialogue as No. 59.	<p>Dermot: ooh erm that sounds painful but if it will speed things up I will take your advice</p>	<p>Marty:</p> <p>[laughs]</p> <p>OK all right, then</p>	<p>1. reduced non-lexical (syntactical) word ‘will’ to /l/, then ‘it will’ is produced as /itl/, 2. collocation ‘speed up’</p>	<p>Tr_Sq_59(60)_ 100 Tr_Dlg_59(60)_ 100</p>

Appendix 19: Transcripts and Analysis of Training Materials 04/2008 for Test Group

Name of File	Description of Theme	Orthographic Text		Linguistic Features of SNIPPETS	Related Files (Optional)
Tr51_100 Tr51_40	Marc is talking about his teaching timetable.	Marc: well ah ah I teach you see I have been teaching about four hours a week there but	Donal: ah once a week yeah	1. reduced 'I have', produced as 'Iv', 2. grammatical paradigm FS 'have been doing', 3. semi-fixed frame '... a week'	Tr_Sq_51_100 Tr_Sq_51_40 Tr_Dlg_51_100 Tr_Dlg_51_40
Tr52_100 Tr52_40	Topic: The benefits of living in the countryside.	Marc: can li- b- b- mainTAINS that she can live quite comfortably in the countryside and be and be anonymous do you know what I mean ahm	Donal: yeah well yeah well I suppose you can yeah yeah	1. reduced 'do you know', produced as 'dje know', 2. weak form of 'what I mean', 3. chunk 'do you know what I mean'	Tr_Sq_52_100 Tr_Sq_52_40 Tr_Dlg_52_100 Tr_Dlg_52_40
Tr53_100 Tr53_40	Marc is talking about the guy who shares the same flat with him.	Marc: when sitting at the table with sort of five knives and five forks and whatever you know his was a	Donal: [laughs]	1. two 'you know' are chunks, used as a word-filler, 2. the first 'you know' reduced to 'je know', 3. 'was a' reduced to	Tr_Sq_53_100 Tr_Sq_53_40 Tr_Dlg_53_100 Tr_Dlg_53_40

	of geological features.	It is fantastic is not it yeah well I used to love going over at visit Elena	the Massif Central and and tipped it vertically I can just see rows and rows of eh	‘its’, 2. tag question intonation, and answer	Tr_Dlg_56_40
Tr57_100 Tr57_40	Topic: Seeing a doctor. (They are reading out a dialogue extracted from the textbook of junior secondary school, but in a more natural way).	Dermot: Good morning, doctor. I have got a bad cold, I have a headache and a running nose. OK. Do you have any quicker ways? Oho, no! It is a little painful to get an injection. I prefer to take medicine. eh Is it good to drink much water?	Marty: Good morning. What is your trouble, young man? Oh, I see. Let me check you. Oh, it is necessary for you to take some Chinese medicine. em, yes. You can get an injection. Oh, sure! It is It is also important to get a lot of rest.	1. weak form on the subject pronoun ‘I’, reduction of ‘have’, thus ‘I have’ is produced as ‘əv’, 2. weak form on three occurrences of ‘a’, 3. collocation ‘have got a bad cold’, ‘have a headache’ and ‘have a running nose’	Tr_Sq_57(58)_100 Tr_Sq_57(58)_40 Tr_Dlg_57(58)_100 Tr_Dlg_57(58)_40

		Thank you, doctor. I will do as you say.			
Tr58_100 Tr58_40	The same dialogue as No. 57.	Dermot: eh Is it good to drink much water? Thank you, doctor. I will do as you say.	Marty: Oh, sure! It is It is also important to get a lot of rest.	1. weak form on 'I will', produced as /əl/, 2. weak form on 'as', produced as /əs/	Tr_Sq_57(58)_ 100 Tr_Sq_57(58)_ 40 Tr_Dlg_57 (58)_100 Tr_Dlg_57 (58)_40
Tr59_100 Tr59_40	The same topic as No. 57 and 58, but a more natural interactive dialogue is produced.	Dermot: good morning, doctor. oh, I woke up yesterday with a pain in my throat. and I am afraid it is going to go down to my chest. so I would like to take some preventive measures. 'ah'	Marty: oh, good morning. how are you today? what can I do for you? aw right. yeah, sure. well, first of all, let me just have a look in your in your throat if I may can you say 'ah' OK yeah oh, I I see it is it is quite	1. weak form on the article word 'an' – joined with 'anti-biotic', 2. tag question intonation	Tr_Sq_59(60)_ 100 Tr_Sq_59(60)_ 40 Tr_Dlg_59(60)_ 100 Tr_Dlg_59(60)_ 40

		<p>you could not give me an anti-biotic, could you</p> <p>ooh ooh erm that sounds painful but if it will speed things up I will take your advice</p>	<p>infected actually ah very red so erm yeah, you you need to take some Chinese medicine</p> <p>no, I I prefer not to it is not very healthy for you so well, I I suppose I could give you erm an injection that would work a bit faster</p> <p>[laughs]</p> <p>OK all right, then</p>		
Tr60_100 Tr60_40	The same dialogue as No. 59.	<p>Dermot: ooh erm that sounds painful but if it will speed things up I will take your advice</p>	<p>Marty:</p> <p>[laughs]</p> <p>OK all right, then</p>	<p>1. reduced non-lexical (syntactical) word ‘will’ to /l/, then ‘it will’ is produced as /itl/, 2. collocation ‘speed up’</p>	<p>Tr_Sq_59(60)_ 100 Tr_Sq_59(60)_ 40 Tr_Dlg_59(60)_ 100 Tr_Dlg_59(60)_ 40</p>

**Appendix 20: Questionnaire 2 on Training Process for Control Group
(Chinese Teachers of English)**

1	Which control group are the students in (GpAc, GpBc, GpCc)?	
2	For how long are the students exposed to the training materials per week? (hours)	
3	Where do the students access these training materials – in the classroom or the language lab?	
4	Do they wear headphones or not?	
5	Do the students first access the snippets, the sequences or the dialogues?	
6	In what order do students listen to these recorded materials?	
7	What instructions are given by you before each training session?	
8	Are the students asked to imitate or write what they heard?	
9	If the students are asked to imitate, what forms of the words are they asked to produce – citation forms or connected speech?	
10	Are the students asked to imitate only (intelligibility) or interpret as well (comprehension)?	
11	Is the orthographic text explained or displayed on the blackboard?	
12	If yes, when is it explained or displayed – before the students perform the tasks, or afterwards?	
13	Do you also ask the students to perform these tasks after the orthographic text is explained or displayed?	

14	If yes, is there any difference between the earlier performance (before the orthographic text is explained or displayed) and the later performance (after the orthographic text is explained or displayed)?	
15	If there is a difference, which performance is better?	
16	Are the linguistic features also laid out on the board and explained to the students?	
17	If so, at what stage?	
18	In addition to the training given by you, can the students access the training materials at any time they wish?	
19	Are there any particular phonemes the students find difficult?	
20	Are the snippets, sequences and dialogues of a suitable length?	
21	Are the training materials suitable for the students?	
22	Are the students interested in the topics?	
23	What suggestion do you have for further training?	
24	Any other comment?	

**Appendix 21: Questionnaire 2 on Training Process for Test Group
(Chinese Teachers of English)**

1	Which test group are the students in (GpAt, GpBt, GpCt)?	
2	For how long are the students exposed to the training materials per week? (hours)	
3	Where do the students access these training materials – in the classroom or the language lab?	
4	Do they wear headphones or not?	
5	Do the students first listen to the normal speed or the slowed-down speed?	
6	If the normal speed is played first, how many times will the normal speed be played before the students hear the slowed-down speed?	
7	Do the students first access the snippets, the sequences or the dialogues?	
8	In what order do students listen to these recorded materials?	
9	What instructions are given by you before each training session?	
10	Are the students asked to imitate or write what they heard?	
11	If the students are asked to imitate, what forms of the words are they asked to produce – citation forms or connected speech?	
12	Are the students asked to imitate only (intelligibility) or interpret as well (comprehension)?	
13	When are the students asked to perform these tasks – during the normal speed, after the slowed-down	

	speed, or both?	
14	If both, which speed leads to a better performance?	
15	Is the orthographic text explained or displayed on the blackboard?	
16	If yes, when is it explained or displayed – during the normal speed or after the slowed-down speed?	
17	Are the linguistic features also laid out on the board and explained to the students?	
18	If so, at what stage?	
19	In addition to the training given by you, can the students access the training materials at any time they wish?	
20	Are there any particular phonemes the students find difficult?	
21	Are the snippets, sequences and dialogues of a suitable length?	
22	Are the training materials suitable for the students?	
23	Are the students interested in the topics?	
24	What suggestion do you have for further training?	
25	Any other comment?	

Appendix 22: Feedback 1 on Questionnaire 2 on Training Process for Test Group (12/10/2007)

1	Which test group are the students in (GpAt, GpBt, GpCt)?	GpAt
2	For how long are the students exposed to the training materials per week? (hours)	20 minutes
3	Where do the students access these training materials – in the classroom or the language lab?	In the language lab
4	What equipment does the teacher use – PC or tape recorder?	PC
5	Do the students wear headphones or not?	Yes
6	Do the students first listen to the normal speed or the slowed-down speed?	Normal speed
7	If the normal speed is played first, how many times will the normal speed be played before the students hear the slowed-down speed?	Three to four times
8	Do the students first access the snippets, the sequences or the dialogues?	Snippets
9	In what order do students listen to these recorded materials?	Listen for three to four times with normal speed, then to slow speed. Write relative orthotext and linguistic features on the blackboard and explain them. Then listen again.
10	What instructions are given by you before each training session?	Listen and repeat
11	Are the students asked to imitate or write what they heard?	Imitate
12	If the students are asked to imitate, what forms of	Citation forms

	the words are they asked to produce – citation forms or connected speech?	
13	Are the students asked to imitate only (intelligibility) or interpret as well (comprehension)?	Imitate only
14	When are the students asked to perform these tasks – during the normal speed, after the slowed-down speed, or both?	Both
15	If both, which speed leads to a better performance?	Slow speed
16	Is the orthographic text explained or displayed on the blackboard?	Yes
17	If yes, when is it explained or displayed – during the normal speed or after the slowed-down speed?	After the slow speed
18	Are the linguistic features also laid out on the board and explained to the students?	Yes
19	If so, at what stage?	After slow speed
20	In addition to the training given by you, can the students access the training materials at any time they wish?	No
21	Are there any particular phonemes the students find difficult?	Reduction and weak forms
22	Are the snippets, sequences and dialogues of a suitable length?	Yes
23	Are the training materials suitable for the students?	Some of them are not.
24	Are the students interested in the topics?	Yes, most of them
25	What suggestion do you have for further training?	Students like the materials which close to their study and life.
26	Any other comment?	No

Appendix 23: Feedback 2 on Questionnaire 2 on Training Process for Control Group (06/11/2007)

1	Which control group are the students in (GpAc, GpBc, GpCc)?	GpBc
2	For how long are the students exposed to the training materials per week? (hours)	One hour
3	Where do the students access these training materials – in the classroom or the language lab?	In the classroom
4	What equipment does the teacher use – PC or tape recorder?	Tape recorder
5	Do the students wear headphones or not?	No
6	Do the students first access the snippets, the sequences or the dialogues?	Snippets
7	In what order do students listen to these recorded materials?	Listen first, then write the orthotext onto the blackboard, and then explain the linguistic features
8	What instructions are given by you before each training session?	Listen and repeat
9	Are the students asked to imitate or write what they heard?	Imitate
10	If the students are asked to imitate, what forms of the words are they asked to produce – citation forms or connected speech?	Connected speech
11	Are the students asked to imitate only (intelligibility) or interpret as well (comprehension)?	Imitate only
12	Is the orthographic text explained or displayed on the blackboard?	Yes
13	If yes, when is it explained or displayed – before	Afterwards

	the students perform the tasks, or afterwards?	
14	Do you also ask the students to perform these tasks after the orthographic text is explained or displayed?	Yes
15	If yes, is there any difference between the earlier performance (before the orthographic text is explained or displayed) and the later performance (after the orthographic text is explained or displayed)?	Yes
16	If there is a difference, which performance is better?	Later performance
17	Are the linguistic features also laid out on the board and explained to the students?	Yes
18	If so, at what stage?	After listening for 3 to 4 times
19	In addition to the training given by you, can the students access the training materials at any time they wish?	No
20	Are there any particular phonemes the students find difficult?	No
21	Are the snippets, sequences and dialogues of a suitable length?	Snippets are OK, but the sequences and dialogues seem too long.
22	Are the training materials suitable for the students?	No
23	Are the students interested in the topics?	Some of them
24	What suggestion do you have for further training?	Choosing the suitable materials
25	Any other comment?	No

Appendix 24: Feedback 3 on Questionnaire 2 on Training Process for Test Group (20/11/2007)

1	Which test group are the students in (GpAt, GpBt, GpCt)?	GpBt
2	For how long are the students exposed to the training materials per week? (hours)	One hour
3	Where do the students access these training materials – in the classroom or the language lab?	In the classroom
4	What equipment does the teacher use – PC or tape recorder?	Tape recorder
5	Do the students wear headphones or not?	No
6	Do the students first listen to the normal speed or the slowed-down speed?	Normal speed
7	If the normal speed is played first, how many times will the normal speed be played before the students hear the slowed-down speed?	Three to four times
8	Do the students first access the snippets, the sequences or the dialogues?	Snippets
9	In what order do students listen to these recorded materials?	1. First listen, then write the orthotext onto the blackboard, and then explain the meaning of new words and linguistic features. 2. Listen to the snippets first, then to the sequences and to dialogues.
10	What instructions are given by you before each training session?	Listen and repeat
11	Are the students asked to imitate or write what they heard?	Imitate

12	If the students are asked to imitate, what forms of the words are they asked to produce – citation forms or connected speech?	Both
13	Are the students asked to imitate only (intelligibility) or interpret as well (comprehension)?	Imitate only
14	When are the students asked to perform these tasks – during the normal speed, after the slowed-down speed, or both?	Both
15	If both, which speed leads to a better performance?	Slow speed
16	Is the orthographic text explained or displayed on the blackboard?	Yes
17	If yes, when is it explained or displayed – during the normal speed or after the slowed-down speed?	During the normal speed
18	Are the linguistic features also laid out on the board and explained to the students?	Yes
19	If so, at what stage?	After slow speed
20	In addition to the training given by you, can the students access the training materials at any time they wish?	No
21	Are there any particular phonemes the students find difficult?	[v], [n], [ŋ], [θ], [ð], and [ʒ], etc.
22	Are the snippets, sequences and dialogues of a suitable length?	No
23	Are the training materials suitable for the students?	Not really
24	Are the students interested in the topics?	Not really. They are too far from what they learn in class and the student's life.
25	What suggestion do you have for further training?	If it is possible to include some topics, like shopping, making a call, booking a ticket, seeing a doctor and

		asking the way, etc., in the training materials since they are what the students are more familiar with and more interested in.
26	Any other comment?	Slowed-down version sounds a little different and seems a bit change.

Appendix 25: Feedback 4 on Questionnaire 2 on Training Process for Test Group (23/12/2007)

1	Which test group are the students in (GpAt, GpBt, GpCt)?	GpCt
2	For how long are the students exposed to the training materials per week? (hours)	45 minutes
3	Where do the students access these training materials – in the classroom or the language lab?	In the classroom
4	What equipment does the teacher use – PC or tape recorder?	PC
5	Do the students wear headphones or not?	No
6	Do the students first listen to the normal speed or the slowed-down speed?	Normal speed
7	If the normal speed is played first, how many times will the normal speed be played before the students hear the slowed-down speed?	Four to five times
8	Do the students first access the snippets, the sequences or the dialogues?	The dialogues
9	In what order do students listen to these recorded materials?	1. Dialogue is played for 4-5 times at normal speed, then 2-3 times at slowed-down speed, and then back to normal speed again for 1-2 times 2. Then sequences, and the snippets. First at normal speed, then go into the slowed-down speed.
10	What instructions are given by you before each training session?	Listen and repeat

11	Are the students asked to imitate or write what they heard?	Imitate
12	If the students are asked to imitate, what forms of the words are they asked to produce – citation forms or connected speech?	Citation forms
13	Are the students asked to imitate only (intelligibility) or interpret as well (comprehension)?	Imitate only
14	When are the students asked to perform these tasks – during the normal speed, after the slowed-down speed, or both?	Both
15	If both, which speed leads to a better performance?	Slowed-down speed
16	Is the orthographic text explained or displayed on the blackboard?	Yes
17	If yes, when is it explained or displayed – during the normal speed or after the slowed-down speed?	After the slowed-down speed
18	Are the linguistic features also laid out on the board and explained to the students?	Yes
19	If so, at what stage?	After the slowed-down speed
20	In addition to the training given by you, can the students access the training materials at any time they wish?	No
21	Are there any particular phonemes the students find difficult?	Weak forms
22	Are the snippets, sequences and dialogues of a suitable length?	Yes
23	Are the training materials suitable for the students?	Yes
24	Are the students interested in the topics?	Yes
25	What suggestion do you have for further training?	No
26	Any other comment?	The normal speed sounds more natural than the

		slowed-down speed.
--	--	--------------------

Training procedures about Q9 are recommended to the teacher:

Start with snippets, then sequences, and at last go into the wider context – dialogue.

Every procedure begins with normal speed first, then slowed-down speed, then back to normal speed again.

But the teacher can decide which the better way is for the students.

Appendix 26: Dialogues Provided by Chinese English Teacher for Reference

Topic 1: shopping

1. MAN: Good morning. Can I help you?

GIRL: Yes, please. I'd like to buy a camera.

MAN: We have many cameras here. Some are made in China, and some are made in other countries.

GIRL: This one looks very nice, where is it made?

MAN: In Japan.

GIRL: How much is it?

MAN: 5,000 *yuan*. This is a digital camera.

GIRL: Wow, that's too expensive! I can't afford it. Do you have an ordinary one?

MAN: Sure. What about those ones?

GIRL: Well, this one looks very good. And the price is OK. Is it made in China?

MAN: Yes, it's made in Shanghai.

GIRL: OK. I'll take this one.

2. MR. SHAW: Good afternoon. What can I do for you?

MRS. CLARK: I'd like to buy a washing machine.

MR. SHAW: This is the newest washing machine, madam.

MRS. CLARK: Is it made in Sweden?

MR. SHAW: No, it is made in Germany.

MRS. CLARK: Please show me how to use it.

MR. SHAW: Sure. Here are some sheets and shirts. You put them in the machine.

You close the door and press the button.

MRS. CLARK: Oh, but the machine shouldn't sound like that. It's shaking.

MR. SHAW: The ground is not so flat, madam.

MRS. CLARK: It makes so much noise.

MR. SHAW: Washing machines always make noise.

MRS CLARK: Well, let me think about it. Thank you.

3. A: What can I do for you, Miss?

B: I'm looking for a light yellow suit.

A: What size do you want, please?

B: Size M.

A: Yes, we've got Size M, but the yellow ones are sold out.

B: Have you got any other colours?

A: Yes, what about those over there?

B: They look nice. How much does it cost?

A: 298 dollars.

B: Hmm! That's very expensive. Can I try it on?

A: Of course.

B: It's just right. I'll take it.

4. TIM: Excuse me. How much are the shirts?

GIRL: Which one?

TIM: The woollen one.

GIRL: Oh, it costs 49.88 dollars.

TIM: Almost fifty dollars. That's too expensive.

GIRL: How about the cotton one? It's 29 dollars.

TIM: That sounds reasonable.

GIRL: What size do you want?

TIM: Large. May I try it on?

GIRL: Sure. Oh, it fits you well.

TIM: I think so. I'll take it. Can I have a discount?

GIRL: Sorry, I'm afraid not.

5. A: Can I help you, Madam?

B: Yes, I'm looking for a leather jacket.

A: What colour do you want?

B: I want a warm colour.

A: Let me see what colours I have ... Oh, how about the red one? Here you are.

B: Hmm. It looks too bright. Do you have any other colours?

A: What about the brown one and the black one?

B: Well, I prefer the brown one. How much is it?

A: 880 *yuan*.

B: That's too expensive. I don't think I'll take it.

A: Maybe you can try it on and have a look before you decide.

B: OK.

A: It looks very nice on you and it fits you so well. If you take it, I'll give you a 10% discount.

B: Hmm. It's really very nice. Even though it's still a little expensive.

6. WOMAN: Good morning. Can I help you?

BILL: Yes, please. I'd like to buy a mobile phone, but I don't know much about mobile phones.

WOMAN: Well, what brand do you like, Motorola, Nokia or Siemens?

BILL: I have no idea.

WOMAN: Motorola and Nokia are more popular. You can buy a Nokia if you like.

BILL: How much is a Nokia?

WOMAN: It depends on what types you buy. The old type is cheaper and the new type is expensive.

BILL: I don't want the newest. A practical one will do.

WOMAN: You can have this one. It looks nice and works well, and it's not expensive.

BILL: How much is it?

WOMAN: 1,200 *yuan*.

BILL: Does it include everything?

WOMAN: No. If you choose a number, you need to pay another 200 *yuan*.

BILL: OK. I'll take this one. Thank you for your help.

WOMAN: It's my pleasure.

Topic 2: making a call

7. WOMAN: Hello. The Rose Hotel, can I help you?

MR. SMITH: Yes, please. I'd like to book two rooms for the coming weekend.

WOMAN: Next weekend, you say? Well, how many nights?

MR. SMITH: Oh! Just two.

WOMAN: OK. Two nights. From Friday or Saturday?

MR. SMITH: Friday.

WOMAN: Friday, the 20th ... And do you need single rooms or double?

MR. SMITH: Two double rooms, please.

WOMAN: Right. Two doubles for Friday and Saturday, July 20th and 21st.

MR. SMITH: What's the price of a double room?

WOMAN: Well, it's 400 *yuan* per night, with breakfast.

MR. SMITH: I see. Can I book the rooms now, please?

WOMAN: Certainly. What's your full name, please?

MR. SMITH: Rick Smith.

8. A: Hello! May I speak to Miss Zhao?

B: Hold on for a moment, please! (A moment later) I'm sorry, she isn't here right now. Could I take a message?

A: Certainly, that's very kind of you. I want to speak to her about my son David Smith. He has got a bad cold. I'm afraid he isn't able to go to school today.

B: I'm sorry to hear that. I hope he'll be all right very soon.

A: Thank you! Could I leave my telephone number to you?

B: OK. Wait a moment, please. I'll go and get a pen and a piece of paper. OK! Please give me your phone number.

A: 3579860. That's my office telephone number. My name is Peter Smith.

B: All right. I'll ask her to call you as soon as she is back.

A: That'll be fine. Thank you very much. Goodbye!

B: Goodbye!

9. MR. GREEN: Could I speak to Mr. Song Jia, the headmaster, please?

TEACHER: I'm sorry he isn't here right now. May I help you?

MR. GREEN: That's very kind of you, but I want to speak to him about my son, Jim Green. We are going to Mount Emei on Friday. I hope to see him as soon as possible.

TEACHER: Are you free later today, Mr Green?

MR. GREEN: Sorry. I'm free every day except today.

TEACHER: Can you come tomorrow?

MR. GREEN: I think so. What time?

TEACHER: Between 8:00 and 9:00.

MR. GREEN: Yes, that would be fine.

TEACHER: I'll leave a message on his desk.

MR. GREEN: Many thanks. Goodbye!

10. ANN: Hello. Is Tina there?

TINA: Hello, Ann. This is Tina.

ANN: Could I borrow your CD player tomorrow, please?

TINA: Sorry. It's not a very good line. Could you speak more loudly, please?

ANN: Sure. I said, could I borrow your CD player, please? Mine's broken.

TINA: Of course. When do you want it?

ANN: Tomorrow if possible.

TINA: OK. I'll bring it to school in the morning.

ANN: Thanks! Bye.

MOM: Who was that?

TINA: It was ANN.

MOM: Why were you shouting?

TINA: The line was bad. We couldn't hear each other clearly.

MOM: What did she want?

TINA: She wanted to borrow my CD player.

MOM: Why didn't you tell her to use her own?

TINA: She said hers was broken.

Topic 3: booking a ticket

11. MR. SMITH: Could you tell me how much it costs to fly to Hainan?

WOMAN: Where are you flying from?

MR. SMITH: From Beijing.

WOMAN: The price of a ticket from Beijing to Haikou is 1,000 *yuan* one-way.

MR. SMITH: Can you tell me if there is a flight in the morning on November, 26th?

WOMAN: Just a minute, please. I have to check my computer. Yes. There's a flight at 8:15 in the morning.

MR. SMITH: Great! I'd like to book four tickets, please.

WOMAN: Would you want one-way or round-trip?

MR. SMITH: Round-trip. We'll return on the 8th of December. Is that possible?

WOMAN: Yes. Four tickets on November 26th to Haikou and returning to Beijing on December 8th.

Topic 4: seeing a doctor

12. TIM: Good morning, doctor.

DOCTOR: Good morning. What's your trouble, young man?

TIM: I've got a bad cold. I've a headache and a running nose.

DOCTOR: I see. Let me check you. It's necessary for you to take some Chinese medicine.

TIM: OK. Do you have any quicker ways?

DOCTOR: Yes. You can get an injection.

TIM: Oh, no! It's a little painful to get an injection. I prefer to take medicine. Is it good to drink much water?

DOCTOR: Sure! It's also important to get a lot of rest.

TIM: Thank you, doctor. I'll do as you say.

13. BOY: Doctor, I'm too heavy. Can you help me to lose some weight?

DOCTOR: It's necessary to do morning exercises every day.

BOY: It isn't good to eat too much sweet food, is it?

DOCTOR: That's right. It's bad for you to eat too much sweet.

BOY: Is it OK to swim for one hour every day?

DOCTOR: Yes, that's a very good exercise.

BOY: Anything else?

DOCTOR: Yes. It's OK to eat some fruit before meals.

14. A: I'm not feeling well. I have got a cold.

B: Have you been to the hospital yet?

A: Yes, I went to see Doctor Li this morning and he gave me some medicine.

B: Have you taken the medicine?

A: Yes, I took it just a moment ago.

B: Mr Li is a good doctor. I think you'll be all right soon.

A: I hope so. By the way, where is our teacher? I have to ask for sick leave.

B: She has gone to the office.

Appendix 27: Questionnaire 3 for Test Group (English Version)

(1) not at all (2) not much (3) undecided (4) somewhat (5) very much

1. Do you like listening to the training materials?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

2. Can you understand the materials when listening at normal speed?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

3. Does the slow-down help you hear what was said?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

4. Did the training period help you understand L1 speakers better?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

5. If the answer to Q4 is '(4)' or '(5)', was the improvement due to the slow-down facility?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

6. Any other comment: _____

Please tick the comment you agree with. It is possible to tick more than one box.

The slow-down facility:

☐ gives me more time to listen and understand

☐ makes the vowels clearer

☐ makes the consonants clearer

☐ is too slow

☐ sounds unnatural

□ helps me follow NS intonation patterns

Thank you for your cooperation!

Appendix 28: Questionnaire 4 for Control Group (English Version)

(1) not at all (2) not much (3) undecided (4) somewhat (5) very much

1. Do you like listening to the training materials?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

2. Did the training period help you understand L1 speakers better?

☐ (1) ☐ (2) ☐ (3) ☐ (4) ☐ (5)

3. Any other comment: _____

Thank you for your cooperation!

Appendix 29: Handout of Training Samples Practised by GpB

(1) which one are you thinking of

1. reduced 'you' to /je/, 2. clearly pronounced /v/ in 'of', 3. weak form of 'are', produced as /ə/, 4. collocation 'think of', 5. grammatical paradigm FS 'are you'

(2) certainly made up for that

1. lively stress intonation pattern, 2. collocation 'make up for'

(3) I moved out of home when I was eighteen

1. double word stress on 'eighteen', 2. weak form of 'I', produced as /ə/, 3. collocation 'move out of', 4. grammatical paradigm FS 'I was'

(4) fifty sixty percent of the people

1. intonation indicates a range or a more exact figure, 2. big intonational change, 3. barely distinguishable 'of the' (= uh th), 4. collocation 'percent of'

(5) I can not walk down the street without having to walk on the road

1. intonation and extension of 'road' indicate further phrase to follow, 2. huge intonational change, 3. it is hard to hear that 'can't' is negative (and not 'can'. Negativity is made clear to the L1 listener by extending the '-n-' and following with 'without ...', NOT by articulating the '-t' in 'can't'), 4. extreme reduction of 'without' to /wiə-/, 5. reduction of 'having to' to /havn to/, 6. grammatical paradigm FS 'have to'

(6) well that is exactly what the Italians would have been doing

1. really fast speed of delivery, 2. huge intonational change, 3. reduced 'would have been' to /wudəbin/, 4. weak form of 'what', produced as /wə/, 5. grammatical paradigm FS 'would have been'

(7) you are scraping the bottom of the barrel sir

1. idiom 'scrap the bottom of the barrel', 2. fast speed in unstressed elements, 3. expressiveness of this idiom fits into the expressive envelope: 'Aw, Donal - How could you tell such a sick joke?!?', 4. weak form of 'of', produced as /ə/

(8) particularly when a woman asks you

1. fast speed, 2. reduced 'particularly' to /pətɪkəlju/

(9) but I have also had five hours a week at the DBS

1. heavily reduced 'I have also had' to /ɪvəlsəd/, 2. semi-fixed frame '... a week', 3. weak form of 'a' and 'at the', produced respectively as /ə/ and /əd/, 4. grammatical paradigm FS 'have had'

(10) thanks for very much for taking the time to come in

1. reduced 'for' to /fə/ and 'taking' to /təkin/, 2. collocation 'come in' and 'thanks (very much) for ...'

(11) have I sent you the erm you know the mistakes that come up during the exam

1. reduced 'have', 2. reduced 'you know', 3. weak form of 'that', produced as /ðə/, 4. word stress on key words, 5. chunk 'you know', 6. collocation 'come up', 7. grammatical paradigm FS 'have sent'

(12) especially when you are working three nights a week

1. reduced 'especially' to 'specially', 2. huge reduction of 'when you are', 3. indistinguishable pronoun 'you', 4. weak form of 'a', 5. individual word stress, 6. semi-fixed frame '... a week', 7. grammatical paradigm FS 'are working'

(13) I do not want to go back

1. reduced 'want to' to 'wanna', 2. fast speed of delivery, 3. collocation 'go back'

(14) I do not know th-what you were doing in the eighties

1. reduced 'do not know' to 'dunno', 2. reduced 'what you were' to /wə tjuə/, 3. weak form of 'in', pronounced as /n/, 4. grammatical paradigm FS 'were doing', 5. collocation 'in the eighties'

(15) something to do with the fact that I am I am growing a a bit older

1. large reduction of 'i' in 'with', 2. reduction of 'that I am' to 'that am', 3. individual word stress, 4. semi-fixed frame 'something to do with', 5. collocation 'a bit', 6. grammatical paradigm FS 'am growing'

(16) people just were not able to talk to him

1. reduction of 'were not able to' to 'wənt ebl tə', 2. weak form of 'to him' produced as /tə im/, 3. rejectionist tone in retelling a story, especially in 'talk', 4. grammatical paradigm FS 'were not able to', 5. collocation 'talk to'

(17) ah one of my neighbours one day about six or seven years ago

1. weak form of 'or' produced as /ə/

(18) and of course he goes for the weird and wonderful and the way-out

1. reduction of 'and of course' to 'ənə course', 2. slight reduction of 'and' to 'an' in the idiom 'weird and wonderful', 3. collocation 'of course' and 'go for', 4. collocation 'weird and wonderful'

(19) would not you think that it would ah- translate into the Belgian community as well

1. reduction of 'would not you' to 'wouldn you', 2. vowel in second 'would' is reduced (despite pause) because it is a grammatical paradigm FS 'would + V.', 3. collocation 'translate ... into ...' and 'as well'

(20) What do you call him

1. reduced 'do you' to 'dju', 2. weak form of 'him' produced as /im/

(21) is that what you mean

1. reduced 'what you mean' to /wə tju min/, 2. weak form of non-lexical words 'is' and 'that', produced as /zðe/

(22) it is going to continue existing

1. reduced 'going to' to 'gonna', 2. grammatical paradigm FS 'is going to', 3. terminal stress via slowing-down speed

(23) well I do have the occasional sleepless night

1. sentence stress on the word 'do', 2. collocation: 'sleepless night'

(24) he could not have known what it meant

1. reduced the syntactic paradigm 'could not have known' to 'couldna', 2. weak form of 'what it', produced to /ə i/, 3. grammatical paradigm FS 'could not have known'

(25) are we on the see-food diet

1. reduced 'are we', produced as /ər wi/, 2. (joking) question intonation, 3. collocation 'on a diet'

(26) you heard about the award winning farmer did not you

1. question intonation, 2. reduction: 'did not you' to: 'didn't yə', 3. collocation 'hear about'

(27) and I got back to the hotel on Sunday evening

1. reduced 'and I' to 'anna', 2. collocation 'get back to'

(28) we will put an end to that

1. weak form on the grammatical word 'will', 2. 'put' becomes 'pud' and there is a large reduction of 'put an end' to: 'pud an en', 3. intonation indicates a threatening intention, 4. collocation 'put an end to'

(29) I am going to hide

1. reduced 'I am going to' to 'ain gonna', 2. grammatical paradigm FS 'am going to'

(30) I am right behind you

1. 'you' is reduced to 'je'

(31) that are particularly interesting particularly fascinating

1. reduced 'particularly' to 'particuly' – in both examples, 2. weak form of non-lexical words 'are', produced as /ə/

(32) it is a pity you do not do a bit more

1. weak form of non-lexical word 'a', produced as /ə/, 2. reduced 'you', 3. reduced 'do not do' to 'dondo', 4. semi-fixed frame 'it is a pity ...', 5. collocation 'a bit'

(33) it was enjoyable

1. stress on the non-lexical word 'was'

(34) and the(y) people have come in in the last say five six seven or eight years

1. reduction of the non-lexical word 'have', 2. 'seven or eight' reduced to /sevn ə e/, 3. five, six, seven and eight are parallel structure to try to find the correct time, 4. collocation 'come in', 5. grammatical paradigm FS 'have come'

(35) and basically people were generally quite nice

1. reduced 'basically', produced as 'basicly', 2. 'generally' is reduced to 'generly', 3. word and stress patterns for emphasis

(36) well do you have a back garden

1. question intonation, 2. reduction of the non-lexical word 'do'

(37) would you believe it

1. reduced 'would you' to 'wouldje', 2. chunk 'would you believe it'

(38) some are made in China and some are made in other countries

1. weak form on the grammatical word 'are', produced as /ə/, then 'some are' is reduced to /sʌmə/, 2. weak form on the non-lexical word 'in', 'made in' produced as 'madn', especially in the second example, 3. collocation 'be made in'

(39) I would recommend a moderate zoom lens

1. reduced 'moderate' to 'modret'

(40) and and how many pixels does that have

1. followed by the ending phoneme –'s' in 'pixels', 'does' is produced as /sʌz/. It is the NON-lexical word which is reduced

(41) they are often looking for people

1. reduced 'they are', produced as 'the', 2. hyper-intensity on 'often', 3. collocation: 'look for', 4. reduced form of 'for' with schwa, produced as /fə/

(42) well do you know what was really shocking

1. reduced 'do you', produced as 'dje', 2. weak form of 'what' and 'was', produced as /wəz/

(43) it is your busiest time of the year

1. reduction of the non-lexical word 'of'

(44) it would be to my favour

1. collocation: 'to one's favour', 2. rising intonation indicates further information to follow

(45) no I would have no questions

1. reduced 'have', produced as /əv/

(46) that is the only thing that came into his head

1. collocation: 'come into', 2. huge reduction of 'only', 3. weak form of 'his', produced as /ɪz/

(47) I know what you are going to say

1. weak form of 'what', produced as /ət/, 2. reduction of 'you', produced as /je/, weak form of 'are', then 'you are' produced as /je ə/, 3. 'going to' reduced to 'gonna', 4. grammatical paradigm FS 'are going to'

(48) I would like to book two rooms for the coming weekend

1. weak form on the subject 'I', produced as /ə/, 2. weak form on the non-lexical word 'for', produced as /fə/, 3. collocation 'would like to'

(49) continental breakfast

1. reduced 'breakfast' to /brekst/

(50) for the other party as well

1. reduced 'as' to /əz/, then 'party as' is produced as /pa:ti əz/, 2. collocation 'as well'

(51) I have been teaching about four hours a week there

1. reduced 'I have', produced as 'Iv', 2. grammatical paradigm FS 'have been doing', 3. semi-fixed frame '... a week'

(52) do you know what I mean

1. reduced 'do you know', produced as 'dje know', 2. weak form of 'what I mean', 3. chunk 'do you know what I mean'

(53) you know his was a a you know good ould

1. two 'you know' are chunks, used as a word-filler, 2. the first 'you know' reduced to 'je know', 3. 'was a' reduced to /wəzə/

(54) you know he travels out to India travels out to

1. chunk 'you know', 2. weak form of 'he', produced as /i/, 3. 'out' reduced to /əu/, 4. collocation 'travel out to'

(55) are you serious

1. slow speed of delivery, 2. wide tonal range, 3. question intonation pattern, 4. 'you' reduced to 'jə', 5. chunk 'are you serious'

(56) It is fantastic is not it yeah

1. reduced 'it is', produced as 'its', 2. tag question intonation, and answer

(57) I have got a bad cold, I have a headache and a running nose

1. weak form on the subject pronoun 'I', reduction of 'have', thus 'I have' is produced as 'əv', 2. weak form on three occurrences of 'a', 3. collocation 'have got a bad cold', 'have a headache' and 'have a running nose'

(58) I will do as you say.

1. weak form on 'I will', produced as /əl/, 2. weak form on 'as', produced as /əs/

(59) you could not give me an anti-biotic, could you

1. weak form on the article word 'an' – joined with 'anti-biotic', 2. tag question intonation

(60) but if it will speed things up

1. reduced non-lexical (syntactical) word 'will' to /l/, then 'it will' is produced as /itl/, 2. collocation 'speed up'

Appendix 30: Instructions for Panel Evaluation – Step 1

Listen to the two exemplar samples in sequence and rank the four student versions from most native-like to least native-like.

Ranked order (best → worst):

Sample 3: When did you move out of home?

most native-like	Student Production No. _____
↓	Student Production No. _____
↓	Student Production No. _____
least native-like	Student Production No. _____

Sample 4: You have a list of questions?

most native-like	Student Production No. _____
↓	Student Production No. _____
↓	Student Production No. _____
least native-like	Student Production No. _____

Appendix 31: Instructions for Panel Evaluation – Step 2

Listen to the two exemplar samples in sequence and give a score to each of the students, based on the evaluation system 0-4 given below, judging word by word. The highest possible score for each word is indicated in the spreadsheet.

‘0’ = when the word is not recognisable, and no individual sounds can be perceived

‘1’ = when a word is not recognisable, but some individual sounds are correct

‘2’ = when a word is recognisable, but there are some errors in individual sounds

‘3’ = when the word is clearly spoken

‘4’ = when a convincing, native-like flow or (where appropriate) intonation pattern is produced, i.e., a **principled**, native-like blur – not an **accidentally** produced blur. If the blur covers two or more words, e.g., ‘did you’, ‘move out of’ and ‘a list of’, then **ALL** the words within the blur are given a ‘4’. If there is no ‘blur’ recognisable, then judge **word by word** and the highest score for each word should be ‘3’.

	recognise word ?	individual sounds correct ?	some errors ?	deliberately produce blur ?
‘0’	×	×		
‘1’	×	√		
‘2’	√		√	
‘3’	√		×	×
‘4’	√		×	√

Sample 3:

	when	did you		move out of			home
Max score	3	4	4	4	4	4	3
Student A							

Student B							
Student C							
Student D							

Sample 4:

	you	have	a list of			questions
Max score	3	3	4	4	4	3
Student A						
Student B						
Student C						
Student D						

Appendix 32: Acceptability of Slow-down Facility – Questionnaire for Linguistic Researchers at BAAL Conference 2008

In order to get feedback on the effectiveness of the slow-down technology on language learning and teaching so as to verify the expectation of the present author that the slow-down facility, to some extent, contributed to the success of learning and acquisition of formulaic language, a questionnaire (see below) was designed for linguistic researchers at the 41st BAAL (British Association for Applied Linguistics) Annual Meeting on 11 September, 2008.

A. Text of Questionnaire

Reaction to Slow-down as a Tool in Language Teaching and Learning

1. Are you a L1 English speaker?

☐ Yes

☐ No

If **NO**, what is your mother tongue? _____

2. Do you have a specialist knowledge of phonetics?

☐ Yes

☐ No

3. Please tick the comments you agree with. It is possible to tick more than one box.

The slow-down facility:

☐ gives me more time to listen and understand

☐ makes the vowels clearer

☐ makes the consonants clearer

- ☐ is too slow
 - ☐ sounds unnatural
 - ☐ helps me follow L1-speaker intonation patterns
4. Do you think the slow-down facility would be helpful in language learning?
- ☐ not at all ☐ not much ☐ undecided ☐ somewhat ☐ very much
5. If you are a language teacher, would you like to use slow-down in your teaching?
- ☐ not at all ☐ not much ☐ undecided ☐ somewhat ☐ very much

If **YES**, how could it best be used?

If **NO**, why not?

Thank you very much for your cooperation!

B. Discussion

A group of 21 linguistic researchers were at the presentation co-presented by the present author in which the slow-down facility was demonstrated and discussed as part of the presentation. After the presentation, the questionnaire was given to the researchers, and ten were returned. The evaluation is as follows.

The first two questions were about the background information of the linguistic researchers. 50% (5 out of 10) of those researchers are L1 English speakers, two are Japanese L1 speakers, two are Chinese (one is Cantonese, the other possibly a Mandarin speaker), and one is a Russian mother tongue speaker. 50% of them have specialist knowledge of phonetics, 4 of them gave a negative answer, and the last one skipped this question.

The following discussions were about the effectiveness of the slow-down facility. 60% of researchers thought that slow-down facility could give them more time to listen and understand. 40% of them thought the slow-down could make the vowels clearer, and 30% thought that it could make the consonants clearer. Also there were 40% of them thought that slow-down technology could help them follow L1-speaker intonation patterns. There were 30% of them found the slow-down was too slow, and there were 70% of the researchers thought the slow-down sounded unnatural. One comment given by one of the researchers (who is an L1 English speaker) was that, even though it sounded unnatural, yet, ‘that does not matter’.

The last two questions were about the acceptability of slow-down technology in language learning and teaching. The answers were based on a 5-scale multiple choice, from very negative to very positive. 9 of the researchers provided their answers. As to whether the slow-down facility would be helpful in language learning (see Table 41 below), 1 out of 9 responses was very negative, 1 was undecided, 4 were somewhat positive and 3 were very positive. Among the answers to whether they would like to use slow-down in their teaching (supposing they are language teachers), 4 out of 9 were

somewhat negative, 2 were a bit positive and 3 very positive, nobody was very negative (see Table 42 below).

Q: Do you think the slow-down facility would be helpful in language learning?

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
Total (9)	1	0	1	4	3

Table 41: Acceptability of the slow-down technology by linguistic researchers (1)

Q: If you are a language teacher, would you like to use slow-down in your teaching?

	(1) not at all	(2) not much	(3) undecided	(4) somewhat	(5) very much
Total (9)	0	4	0	2	3

Table 42: Acceptability of the slow-down technology by linguistic researchers (2)

Some very useful comments were also given by the linguistic researchers. One researcher (an L1 English speaker) thought that the slow-down facility could best be used for language training courses. One of them, who was also an L1 English speaker, commented that 40% speed could help listeners with word segmentation, i.e., to separate where the words began and ended. Two of them (both non-L1 English speakers) thought that it could best be used in L2 listening comprehension practice, especially for clarifying phoneme clusters. There was also one L1 English speaker who thought that the use of the slow-down facility in language teaching could be very flexible. Another researcher, also an L1 English speaker, commented that the slow-down technology was a very interesting research and drew a large ‘smiley’ on the questionnaire to indicate a very positive attitude.

Therefore, generally speaking, even though the amount of feedback given by linguistic researchers was limited, an overall positive attitude could be detected in the questionnaire, especially when combined with some very useful comments which were provided as well. This corroborates the evaluation done by the present author in the consideration of slow-down technology in facilitating Chinese language learners in coping with natural authentic English speech, particularly in the area of formulaic language.

Appendix 33: Presentations and Publications

A. Conference Papers

- Campbell, D. Wang, Y. Meinardi, M. Richardson, B. McDonnell, C. & Pritchard, C. (2010) Dialogic fluency — why it matters. In: TISLID (Technological Innovation for Specialized Linguistic Domains), *theoretical and methodological perspectives*. Madrid, Spain 21-22 October 2010.
- Kilfeather, E. Campbell, D. Wang, Y. McDonnell, C. Meinardi, M. & Richardson, B. (2010) HTML5 and the learner of spoken languages. In: BAAL (British Association for Applied Linguistics), *43rd BAAL annual meeting: applied linguistics: global and local*. Aberdeen, UK 9-11 September 2010.
- Campbell, D. Meinardi, M. Richardson, B. Wang, Y. & McDonnell, C. (2009) DIT's Dynamic Speech Corpus and dialogic fluency. In: EUROCALL (European Association for Computer-Assisted Language Learning), *EuroCALL 2009: new trends in CALL: working together*. Gandia, Spain 9-12 September 2009.
- Campbell, D. Meinardi, M. Richardson, B. & Wang, Y. (2009) DIT's Dynamic Speech Corpus and the new learning paradigm. In: IATEFL (International Association of Teachers of English as a Foreign Language), *43rd IATEFL Conference*. Cardiff, UK 1-4 April 2009.
- Campbell, D. Wang, Y. McDonnell, C. Meinardi, M. & Richardson, B. (2009) The DIT Dynamic Speech Corpus. In: IVACS (Inter-Varietal Applied Corpus Studies), *IVACS Symposium Conference: corpus and interaction*. Edinburgh, UK 22nd January 2009.

- Campbell, D. Wang, Y. & McDonnell, C. (2008) FS \neq FS: formulaicity and prosody. In: BAAL (British Association for Applied Linguistics), *41st BAAL annual meeting*. Swansea, UK 11-13 September 2008.
- Kousidis, S. Dorran, D. Wang, Y. Vaughan, B. Cullen, C. Campbell, D. McDonnell, C. & Coyle, E. (2008) Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues. In: ASSTA (Australian Speech Science and Technology Association) and ISCA (International Speech Communication Association), *INTERSPEECH 2008 incorporating SST 2008*. Brisbane, Australia 22-26 September 2008.
- Campbell, D. Wang, Y. & McDonnell, C. (2007) A prototype speech corpus. In: EUROCALL (European Association for Computer-Assisted Language Learning), *EuroCALL 2007: mastering multimedia: teaching languages through technology*. Coleraine, UK 5-8 September 2007.
- Campbell, D. McDonnell, C. Wang, Y. Meinardi, M. Richardson, B. & Pritchard, C. (2007) DITCall-Slow: slowing native speech for language learners. In: ILTA (the Irish Learning Technology Association), *EdTech2007: 8th annual Irish technology users conference*. Dublin, Ireland 23-25 May 2007.
- Vaughan, B. Kousidis, S. Cullen, C. Wang, Y. McDonnell, C. Campbell, D. & Pritchard, C. (2007) Task-Based mood induction procedures for the elicitation of natural emotional responses. In: *ICASSP 2007 Conference*. Hawaii, USA 15-20 April 2007.
- Cullen, C. Vaughan, B. Kousidis, S. Wang, Y. MacDonnail, C. & Campbell, D. (2006) Generation of high quality audio natural emotional speech corpus using task based mood induction. In: *InSciT2006 Conference*, Merida, Spain 25-28 October 2006.

- Vaughan, B. Cullen, C. Kousidis, S. Wang, Y. McDonnell, C. & Campbell, D. (2006) The Use of task based mood-induction procedures to generate high quality emotional assets. In: *IT&T2006 Conference*. Carlow, Ireland 25-26 October 2006.
- Campbell, D. Wang, Y. Kelleher, J. Meinardi, M. & Richardson, B. (2006) DIT Speech Corpus. In: IVACS (Inter-Varietal Applied Corpus Studies), *3rd IVACS international conference: language at the interface*. Nottingham, UK 23-24 June 2006.
- Campbell, D. Meinardi, M. Richardson, B. & Wang, Y. (2006) Natural English in speaking and listening activities. In: ACELS (The Advisory Council for English Language Schools), *MATSDA/ACELS conference: what we know and what we do: connecting theory and practice in materials development for language teaching*. Dublin, Ireland 21-22 January 2006.

B. Publications

- Campbell, D. Wang, Y. Meinardi, M. Richardson, B. McDonnell, C. & Pritchard, C. (2010) Dialogic fluency — why it matters. Proceedings of the TISLID 2010 Conference.
- Campbell, D. McDonnell, C. Meinardi, M. Pritchard, C. Richardson, B. & Wang, Y. (2009) DIT's Dynamic Speech Corpus. *Speak Out!*, 41, pp.8-11.
- Campbell, D. McDonnell, C. Meinardi, M. Pritchard, C. Richardson, B. & Wang, Y. (2009) Dublin Institute of Technology's (DIT) Dynamic Speech Corpus. IATEFL 2009 Cardiff Conference Selections, pp.108-10.

- Campbell, D. Wang, Y. & McDonnell, C. (2008) FS \neq FS (Formulaicity and Prosody). Proceedings of the 41st BAAL 2008 Conference.
- Campbell, D. McDonnell, C. Wang, Y. Meinardi, M. Richardson, B. & Pritchard, C. (2007) DITCall-Slow: slowing native speech for language learners. Proceedings of the EdTech 2007 Conference.