

# Evaluation of authenticity: Comparison of dialogues in a Chinese coursebook and authentic interactions

by Xu Han

British Council's Master's Dissertation Awards 2020  
Special Commendation



UNIVERSITY OF  
LIVERPOOL

ENGLISH

**MA in TESOL**

**Evaluation of Authenticity: Comparison of  
Dialogues in a Chinese Coursebook and  
Authentic Interactions**

By Xu Han (ID:201255821)

Dissertation submitted in part fulfilment of the degree of  
Master of Arts of The University of Liverpool

September 2019

## **Abstract**

This study surveyed spoken language from a set of Chinese University coursebooks in order to examine the extent to which the language in the conversations represents common features of authentic spoken language, particularly interaction features. It was found that the spoken English used in the coursebooks lacks similarity to unscripted conversational language in use. This is mainly reflected in that the language in coursebooks is short of relevant interactive functions of spoken language: listenership (e.g. response), relational language (e.g. discourse marking and hedging) and overlaps. The results were achieved by a comparison of the frequency of the most common words and their interactive functions in the coursebooks and authentic materials. The findings suggest that unscripted, naturally occurring conversations can be used as supplementary materials of coursebooks for the teaching of spoken language because they represent important aspects of authentic interaction which are lacking in the coursebooks.

## **Declaration**

This work is original and has not been submitted previously in support of any degree, qualification or course.

## **Acknowledgements**

I gratefully acknowledge the advice and guidance of dissertation supervisor Dr Hitomi Masuhara and encouragement of Dr John Graham Ethelston and Dr Christian Jones. I should also mention my appreciation of tutor Clive Newton, friend Michael, and my sister in law for their help when I was studying at the University of Liverpool.

Great thanks also go to my parents, my husband, my son for their love and support in the course of my living in the United Kingdom.

## Table of Content

Abstract .....	2
Chapter 1 Introduction.....	7
1.1 Motivation for this research.....	7
1.2 The aims of the research .....	8
1.3 The scope of the research .....	9
1.4 Definitions of terms .....	10
1.5 Outline of the structure of the dissertation.....	11
Chapter 2 Literature review .....	12
2.1 Definitions of authenticity .....	12
2.2 Previous research on authenticity .....	13
2.2.1 Previous research on authenticity of coursebooks .....	13
2.2.2 Previous research on authentic materials.....	18
2.2.3 Interactive and interpersonal functions of spoken language revealed by previous research on corpora.....	25
Chapter 3 Methodology.....	29
3.1 Outline of the study .....	29
3.2 Data collection and analysis .....	30
3.2.1 Research design .....	30
3.2.2 Data source .....	30
3.3.3 Data analysis.....	34
Chapter 4 Findings and discussions.....	37
4.1 RQ1. Does frequency of the most common words and chunks testified by corpora of general spoken English differ between the coursebooks and authentic data? .....	37
4.1.1 Findings .....	37
4.1.2 Discussions .....	40

4.2	RQ2. Are the common features of spoken discourse, particular interaction features comparable between the coursebooks and authentic data?.....	44
4.2.1	Findings .....	44
4.2.2	Discussions .....	46
4.3	Other findings.....	65
Chapter 5	Conclusion .....	70
5.1	Conclusion of the research.....	70
5.1.1	Summary of the findings .....	70
5.1.2	What have researchers said.....	71
5.1.3	Suggestions.....	73
5.2	Implications of the research.....	74
5.2.1	To teachers .....	75
5.2.2	To teaching material designers .....	75
5.3	Limitations of the research .....	76
5.3.1	Limitation of the sample authentic materials for comparison.....	76
5.3.2	Limitation of the corpus CANCODE (5m).....	76
5.3.3	Limitation of the size of the research.....	77
5.4	For the future .....	77
5.4.1	About application of authentic materials .....	77
5.4.2	About transcription of authentic spoken language.....	78
5.4.3	About corpus and language in use .....	79
	References .....	80
	Appendix 1 Sources and sample texts for coursebook conversations: .....	85
	Appendix 2 Sources and transcripts for conversations (IViE):.....	95
	Appendix 3 Sources and transcripts for Ted Talks: .....	104

## **Chapter 1 Introduction**

### **1.1 Motivation for this research**

Speaking serves important functions. These includes transactional and interpersonal functions which facilitate the exchange of information, goods and services, as well as establishing and maintaining social relations (Thornbury, 2008, p. 28). Therefore, development of speaking competence is important as it is crucial to achieve the purposes of communication.

Chinese students are weak in spoken English compared with other English language skills. According to IELTS test statistics ‘Test taker performance 2017’ (on the IELTS official website), Chinese learners show lower scores in speaking and writing compared with reading and listening (listening 5.90, reading 6.11, writing 5.37, speaking 5.39). Thornbury (2008, p. 28) states that one of the important reasons of speaking failure is lack of practice. This is because most Chinese students learn English in a non-native environment, they do not need to speak English outside classroom.

Another reason for the weakness of spoken English might be the course setting in universities in China, where English lessons have been set as a course of general English for students who are not learning English as one of their major academic disciplines. According to my informal interviews with university students and teachers, it appears



Chinese students have few opportunities to speak English in the classroom as most of the time are spent on reading comprehensions, vocabularies and grammars, though the aim of the courses is to improve students' language competence through listening, speaking, reading and writing. Also, when trying to have conversations in English outside the classroom, they found their knowledge to be inadequate. This was also my experience when I was in university in 1990s.

This has led me to think about and examine how the coursebooks of spoken language used in the classroom could do more to help students with their speaking development.

## **1.2 The aims of the research**

The aims of the dissertation are:

- (1) to explore the extent to which the spoken language in coursebooks represents common features of authentic spoken language, particularly interaction features, and
- (2) how far the conversations in coursebooks are good examples of spoken English for speaking development. (The good examples of spoken English should represent actual features of interactions of authentic conversations which can prepare the students to interact with others in spoken English and to cater for a wide variety of communication purposes and circumstances in their lives and studies).

To achieve these aims, the following research questions will be answered:

RQ1. Does frequency of the most common words and chunks testified by corpora of general spoken English differ between the coursebooks and authentic data?

RQ2. Are the common features of spoken discourse, particular interaction features, comparable between the coursebooks and authentic data?

The coursebooks are a set of contemporary general English coursebooks which are used broadly throughout the universities of China. The authentic data is sourced from two websites: one is speech data (conversations) from the IViE Corpora which is downloaded and transcribed into spoken discourses; the other is interviews (conversations between interviewer and interviewee) selected from TED Talks with transcripts provided on the website. The corpora of general spoken English refer to the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) (5 million words of spoken English discourse as described in O’Keeffe, McCarthy & Carter, 2007, chapters two, three, seven and eight). By comparing differences of spoken discourses between coursebooks and authentic conversations in terms of features of spoken language advised by the corpora, the extent to which the coursebooks reflect common features and language patterns of authentic spoken language is examined.

### **1.3 The scope of the research**

The scope of the research covers the spoken discourses of the coursebooks. That means tasks or activities included in the coursebooks are not considered. The voice of the

language such as pitch, intonation, and accent are also out of consideration. The spoken texts refer to dialogues in the coursebooks - that is, there are at least two participants involved in the conversations. The research focuses on dialogue analysis, rather than monologue, because the former occurs more frequently than the latter in our day-to-day communications. The dialogue is more difficult and challenging to deal with as the speakers of a dialogue have less time to prepare for their speech which needs immediate response and interactions between interlocutors. If a learner of English shows an effective communication in a conversation, the learner shows his or her proficiency of spoken language.

#### **1.4 Definitions of terms**

In the above section of aim of the dissertation, authenticity is mentioned when introducing the research questions. In the dissertation, the definition of authenticity uses what Morrow (1977, p. 13) describes as: ‘an authentic text is a stretch of real language, produced by a real speaker or writer for a real audience and designed to convey a real message of some sort’. For a more thorough discussions of authenticity see 2.1 in Chapter 2.

Another term mentioned as one of the objectives of learning English is ‘speaking development’. Tomlinson (2007, p. 2) distinguishes between language acquisition and language development by emphasizing that language acquisition proceeds language development. Language acquisition is about obtaining basic communicative competence. Language development is about obtaining the ability to use the language successfully for

a wide variety of purposes.

Bygate (1987, p. 3) distinguishes language knowledge and language skill. Language knowledge is a certain amount of vocabulary and what we know about how to assemble sentences, whilst skill is the ability to act on the knowledge and produce language rapidly and smoothly for different circumstances. This view further expresses the idea that learning a language is not only obtaining the knowledge of it, but more importantly, it is about using the language. The implication is that spoken English teaching materials should serve the purpose of language learning which can help learners to develop their abilities to use spoken language for a wide range of situations.

## **1.5 Outline of the structure of the dissertation**

Chapter 2 is a literature review related to coursebooks analysis and authentic materials and corpora study. Chapter 3 introduces the methodology of the research. Chapter 4 deals with the results of the research and discusses these findings. Chapter 5 is about the conclusion of the research and its implications.

## **Chapter 2 Literature review**

### **2.1 Definitions of authenticity**

There has been a frequent and hot debate revolving around definitions of ‘authentic’. It is suggested by Breen (1985, p. 61) that authenticity is not about the text or the task but the learner’s interaction with it. In recent years, the scope of the debate has been extended from text authenticity and task authenticity to learner and teacher authenticity, and context authenticity (Tomlinson, 2017, p. 1).

Learners’ authenticity focusses more on the interactions with the materials. Teachers’ authenticity depends on their reactions to the materials and this authenticity may not be possible if they feel the materials are not relevant or interesting. Context authenticity emphasizes the materials being used have to be developed for the needs and wants of learners in classroom. That is, the contexts should be localized and relevant to learners’ profiles, as it might achieve authenticity in its target classrooms. But, if using the contexts in other classrooms where they are not suitable pedagogically or relevant to the learning objectives of the learners, they might be perceived as culturally alien (Tomlinson (2017, pp. 1-5).

The researchers’ views of authenticity should cover various aspects involved in teaching process: the materials, the participants, the relevance and interactions between

participants and materials, the context, etc. However, this research will mainly focus on text authenticity. Text authenticity means the language in texts is being primarily used for communication rather than produced for teaching purposes (the stock definition of authenticity of text) (Tomlinson, 2017, p. 1). At present, it is a topic many researchers pay attention to.

## **2.2 Previous research on authenticity**

### **2.2.1 Previous research on authenticity of coursebooks**

#### **2.2.1.1 General coursebooks are not always offering realistic examples of spoken language**

Coursebooks have not always offered realistic examples of spoken language as they have been over contrived. It has been argued that these coursebooks do not prepare students for the reality of language use outside the classroom (e.g. Gilmore, 2004; Cullen and Kuo, 2007; Angouri, 2010; Tomlinson, 2010).

Gilmore (2004) examined seven textbooks published in 1980s and 1990s with the intention of understanding whether they represented day to day spoken language. To do the research, he selected seven dialogues in seven textbooks with the context of service encounter. Then by choosing the same questions included in the dialogues, he acted as

one of the speakers to initiate seven equivalent authentic dialogues in real life and recorded and transcribed them to spoken discourses. The discourse features - lexical density, false starts, repetition, pauses, terminal overlap, latching, hesitation devices and back-channels, were analysed to answer his research questions. By comparing the frequency and percentage of the discourse features under two datasets, he found there were considerable differences across a range of discourse features and particularly many of the features of the authentic dialogues. These included hesitation, pausing and overlapping turns which were all excluded in the textbook dialogues. Quality analysis was undertaken by him for each of the features in his research. Following this, he also investigated the dialogues with the same topic in three textbooks published in year of 1996, 1999 and 2001. The results showed that more of the discourse features had been found in the recent textbooks compared with textbooks published previously. However, the number of instances of repetition, false starts, hesitation devices and pauses were still well below those expected. The textbooks were a production of interaction combined with a number of different contrived factors which had neat and tidy turns, no false starts, no hesitations, no repetitions and lack of responses to speaker. In addition, the topics shifted abruptly because they were not developed adequately to mirror natural discourse (Gilmore, 2004, p. 363). He stated that the pedagogic artifice of material deprived students' exposure to natural language to some extent. Gilmore (2007, p. 98) indicated that it had long been recognised that textbooks did not present real lifetime spoken language. Although work had been done to address the balance, numerous gaps remained.

Following this, Cullen and Kuo (2007) surveyed a selection of 24 general EFL mainstream coursebooks ranging from levels of beginner to advanced English (published from 2000 to 2006 in the United Kingdom), in order to explore the coverage of spoken grammar in spoken English and the extent to which the existing known knowledge of the spoken discourse was reflected in the textbooks. To undertake the research, they used three categories A, B and C to categorize spoken grammar items. Category A was identified as features which needed grammatical encoding: noun phrase prefaces, noun phrase tags, past progressive tense and ellipsis. Category B was identified as fixed lexico-grammatical units which did not change the forms grammatically: vagueness tags and discourse markers. Category C was identified as non-standard forms which appeared frequently in spoken English but might be considered informal or incorrect by traditional prescriptive grammars due to their usages contradicting standard written forms. Their findings showed that category A were almost ignored and was merely given to advanced levels for extra interest; Category B got some attentions in textbooks, but category C obtained little attention.

The above studies of Gilmore (2004) and Cullen and Kuo (2007) surveyed a large number of coursebooks published from 1980s to 2000s with different aspects of features of spoken language. The findings suggest that textbooks do not present adequate primary features of spoken language to learners, particularly for those who target speaking development as their objectives of learning English. The argument is that if learners' goals are to speak language for communication purposes in their everyday lives, they need to be exposed to



the true nature of spoken languages. Their study inspired this research which surveys a set of university coursebooks recently published in China. The relevant common features of spoken language, particular interactions are the focus of this research.

Tomlinson (2010) compared twelve EFL coursebooks at three different levels with authentic spoken English in a TV programme *Saturday Kitchen*, focusing only on one important function of imperative: to get people to do something. He found that the presenter James Martin did not often use the imperative to get people to do something and it seemed different chunks of language were used for slightly different purposes. Furthermore, he found similar utterances occur frequently on other unscripted television programmes. However, through examining the twelve EFL coursebooks, it was found that not much attention had been given to this important function of imperative in these coursebooks. Tomlinson (2010) suggested that coursebooks did not always seem to refer to authentic texts or corpora to make discoveries about how English was being used in every life, and the gap between the typical usage of spoken English in textbooks and real-life existed. Learners could not resort to textbooks as reliable sources to imitate authentic English because of the lack of certain features of spoken English in the textbooks.

#### **2.2.1.2 Previous research of authenticity on business textbooks**

Angouri (2010) examined six business textbooks and compared them with a particular dataset of twenty-one audio recordings of meetings in seven companies which consisted

of conversations lasting fifty-six minutes on average. As all the companies were multinational companies across a range of industries and considered leading companies in their fields, the set of data was assumed to be typical examples of authentic spoken language in business conferences. Her research was focused on ‘turn taking’ and ‘overlapping talk (OT)’ which normally occur at the beginning part of utterances. Findings suggested most of the linguistic strategies provided to learners in the six textbooks lacked similarity with that in real-life meetings, as they appear to never occur in everyday business meetings or only occur with a low frequency. For example, OT was not explicitly pre-announced (as implied in the textbooks) but was initiated with either a pragmatic device (uh, mmm, hmm), or with explicit agreement or disagreement with the current speaker’s utterance (yes, yeah, no). In addition, the likelihood of specific strategies (and of OT occurring) would depend on the discourse practices of each concept of ‘Communities of Practice’ (Cofp). Also, these strategies were contingent on a variety range of local factors such as the topics of interaction and the participants and their relations. The findings suggest that textbooks materials did not seem to capture the dynamic and complex nature of interactions in real life. Therefore, the textbooks alone could not prepare the students for the dynamic natures of work-related communication and contexts.

### **2.2.1.3 Impacts of contrived coursebooks to learners**

All the above studies suggest that many of the global coursebooks do not offer

representative examples of actual spoken language which can adequately replicate key discourse features of unscripted conversations. In contrast, the coursebooks are deliberately contrived for particular pedagogic aims of explicit teaching of language to help learners by focusing their processing energies on target features. For example, repeating the features to enhance students' opportunities for learning it. The justification often given for this is that such contrivance and simplification can emphasize the particular curriculum focus and enable learners to pay attention to the target features, However, 'this overprotects the learners and contradicts what is known about how languages are acquired and ... does not prepare them for the reality of language use outside the classroom' (Tomlinson & Masuhara, 2018, p. 31).

Nowadays, many textbooks and much teaching are informed by corpora linguistic data of English in actual use. This results in the gap between the language presented in coursebooks and actual English in real life is smaller. However, there is still a significant gap between 'textbook presentations of how English is used and the reality of actual use' (Tomlinson, 2010, p. 87).

### **2.2.2 Previous research on authentic materials**

Driven by dissatisfaction with coursebooks and in particular with the over contrived treatment of spoken language within them, using authentic materials in ELT classes is advocated by many researchers with arguments that authentic language show genuine use

of English (e.g., Allwright, 1979; Wong, Kwok, & Choi, 1995; Mishan, 2005; Gilmore, 2011; Watkins & Wilkins, 2011; Tomlinson, 2013b; Tomlinson, 2016, etc.). Authentic texts 'can provide rich and meaningful exposure to language in use which is a pre-requisite for language acquisition' (Tomlinson and Masuhara, 2018, p. 31).

### **2.2.2.1 Positive effects of using authentic materials**

Many researchers examined the positive effects upon using authentic materials and argued that it was often more motivating for learners to use them for language learning (e.g. Peacock, 1997; Gilmore, 2011).

Peacock (1997) investigated the effects of authentic materials for EFL learners at beginning level in Korea. He suggested that using authentic materials in class is more significantly motivating than traditional textbooks because students know they were learning the real-life English. Also, Peacock (1997) found that authentic materials appeared more motivating than coursebooks for students. However, it was not necessary that the sample learners found authentic learning materials to be more interesting than textbooks.

Gilmore (2011), compared the use of authentic materials with the use of textbooks for Japanese learners. He designed a test to measure students' communicative competence. His findings showed that students using authentic learning materials achieved significant

better results. The above studies show positive feedback from students using authentic materials.

Rather than tests, Wong, Kwok, & Choi (1995) had been using authentic learning materials in their teaching practice. They described their practical experience of using authentic materials in their ELT teaching with Year 1 students in the university of Hong Kong and concluded that their class activities were successful and supported the idea that authentic materials could serve as a bridge between the classroom and the outside world to enrich students' experiences in learning English. Authentic materials can sensitize students to the use of English in the real world. For example, students 'can see how English is used for presenting data, analyses, and recommendations' by learning a company report or laboratory report (Wong et al. , 1995, p. 319). Their teaching practice of using authentic materials and their positive comments on the practice and the achievement of them, have provided valuable data and might be motivating for others to use authentic materials in teaching.

#### **2.2.2.2 Various sources of authentic materials**

Some researchers raised questions of what authentic materials can be used and in what ways. It has been argued that authentic materials, such as academic program, social media on websites, films, soaps, etc. could be used as sources of authentic learning materials as they represent nature of actual spoken language. Some of these authentic materials have

been examined in below.

(1) Actual MBA program

Basturkmen (2001) made a study on a pervasive aspect of interactive speaking: question-response sequences in academic and work settings. She stated that, in some authentic sample texts sourced from a discussion taken place in an MBA program, the turns showed features that referred back to preceding utterances to signal the intentions or topics of their questioning turn. By analyzing another episode, she indicated the significant phenomenon of speakers' picking up each other's lexis demonstrated that 'talk' is a cooperative venture. Then she suggested that the study offers realistic models of interactive speaking, though it is based on a limited number of sample texts.

(2) Teachers' tales

Timmis (2010) advocated strongly to use naturally occurring texts rather than to design texts due to crucial features of spoken language that are encountered in a natural discourse context. To select spoken texts which represent typical spoken language features, he suggested using teachers' own tales for teaching since authentic spoken language can significantly motivate and engage students as it is related to their teachers' own experiences and spoken by their teachers themselves. He shared his own practical example of teacher-generated materials which was used in his teaching and motivated his students significantly for learning. The experience was also replicated by one of his colleagues.

### (3) YouTube

Additionally, Watkins & Wilkins (2011) advocated using online video website YouTube.com (or other online websites of streaming video) which could stimulate learner autonomy and raise various ways that could be used for teachers to teach English. They suggested that YouTube was a valuable resource to store, exhibit, view and download a great number of digital files which also could be used in class and after class to develop communicative competence and promote authentic lexis development. Based on existing literature review, they highlighted the potential value of application of YouTube in teaching, limitations and suggestions for future research.

### (4) Soaps

A recent research with regard to authentic teaching materials was done by Jones and Horak (2014) who analyzed features of spoken language from a small corpus of the popular UK soap opera EastEnders to investigate the extent to which the language used in the soap opera may be used as a model of spoken English for learners at intermediate levels and above. To do the analysis, they first used corpus data to discover whether the language corresponded to the findings of corpus research. They investigated the coverage of the first two thousand frequent common words in the soap opera as advised by British National Corpus (BNC, 2012). After that, they compared the most frequent words and chunks referring three corpora: the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) (as described in O’Keeffe, McCarthy and Carter, 2007), BNC, and

the American Soaps corpus (Davies, 2012). Following this, they analyzed keywords occurring with greater frequency than as advised by BNC which is more related to the features of the soap opera such as specific names. Finally, they adapted the framework (A, B and C categories) used by Cullen and Kuo (2007) to do analysis which was described in section 2.2.1.1 (except for category C, they examined repetition and overlapping).

The results showed nearly 95% coverage of the two thousand most frequent words of the spoken language section in the British National Corpus (BNC) were used in the opera. Also, it was found that there was the similarity of the most frequent words and two-word chunks between general spoken corpora and a larger soap opera corpus. Further, it was found that soap opera conversations shared at least some important characteristics of spoken language including ellipsis, discourse marking, etc. Jones (2017, p. 159) states it is surprising how few recordings of spoken language are available with transcripts. He raised the point that ‘the scripted spoken English of soap operas may therefore be a useful ‘halfway house’ between spoken English and textbook dialogues’ (Jones & Horak, 2014, p. 161).

##### (5) Films

In more recent years, Carmen (2016) shared the practical experience of development of The Film in Language Teaching Association (FILTA), which is an association to provide resources and share experience for learners and teachers using films for language learning



and teaching. She indicated that positive effects of applying multimodal audio-visual texts on learning and teaching English had been identified through the research into the area for the past 15 years.

Also, Berk (2009) examined multimedia teaching by using video clips in college class. Based on extensive literature review of the research and theories on methods of how videos had been used in the class, he argued that the research on multimedia learning had provided an empirical foundation for their application in teaching, especially for learners to increase their memories and understandings. He listed learning outcomes, suggested procedures and provided 12 specific techniques for video application in class. Though the research focused across the fields of education rather than English language teaching, the review of research on potential value of videos application can be considered as a source of authentic materials.

All the above researchers have done their studies on authentic materials with reference to language learning (except Berk, 2009, which is relevant to language leaning, but not directly about language learning). Some of them analyzed the positive effects of using authentic materials, some put them into their teaching practice (e.g. Gilmore, 2011; Peacock, 1997; Wong, Kwok, & Choi, 1995). Some explored the different available resources of authentic materials and ways or procedures of the applications in classroom which may inspire teachers to use them on their own classrooms and help them on their teaching practice (e.g. Basturkmen, 2001; Berk, 2009; Timmis, 2010; Tomlinson, 2010;

Watkins & Wilkins, 2011; Jones & Horak, 2014; Carmen, 2016). Their research has contributed to increasing of our existing known knowledges about sources of authentic materials and exploring of their applications in teaching practice.

### **2.2.3 Interactive and interpersonal functions of spoken language reveled by previous research on corpora**

When we talk about authentic materials, it is impossible not to mention corpora. ‘Corpus data have enabled the presentation to learners of actual samples of English as it has been used rather than examples of language as it is assumed to be typically used’ (Tomlinson, 2010, p. 87). Many researchers analysed corpora data to understand common features of English language in use (e.g. Thornbury, 2005; McCarthy & Carter, 1995; McCarten & McCarthy, 2010, O’Keeffe et al., 2007; etc.).

Thornbury (2005, p. 65) analysed a conversational extract from UTS Australian English Corpus and found the interactive features of spoken discourse very clear: taking turns and signaling their intentions to speak, paying attentions or responding when others speaking, interrupting at times and showing agreement or amusement by grunts, laughs and chuckles and also including asking and answering questions. ‘Conversation is not simply for the exchange of information but has a strong interpersonal function. That is, it serves to establish and maintain group solidarity’ (Thornbury, 2005, p. 66). He explained interactions referring to taking turns, interrupting and response to speakers, etc., and

mentioned that the linguistic interactive devices (e.g. back-channels, discourse markers, chunks) are used very significantly in conversations with different functions such as signposting the shift of turn or interconnection of the talks. The linguistic interactive devices and functions are also the core content which will be analyzed to examine features of coursebooks in this dissertation.

McCarthy & Carter (1995, p. 207) argued that “the interpersonal implications of spoken grammars are important”. The scrutiny of spoken English on corpus-based revealed that a certain grammatical choices or forms enable a greater degree of interactive and interpersonal language uses - uses of language which are in conformity with the goals of most projects of communicative language teaching. McCarthy & Carter (1995, p. 208) examined the core grammatical forms and significant patterns of spoken language referring to corpus data being collected at the University of Nottingham. The corpus data was constructed for the purposes of studying spoken grammar and targeted conversational and casual language rather than more varieties of formal spoken languages. They found that it was evident that speakers regularly made choices to have interactions with their collocutors and maintained relationship which reflected the interactive and interpersonal nature of the communication. McCarthy & Carter (1995, pp. 216 - 217) indicated that students raised conscious awareness of interactive properties by being exposed to the interpersonal uses of language and negotiation of meanings. This enabled them to obtain the knowledge of interpersonal and interactive functions of different lexico-grammatical options and gradually develop a capacity for noticing such features.

McCarten & McCarthy (2010, p. 15) examined the North American conversations from the Cambridge International Corpus and found some important functions of relational languages. In contrast to 'transactional' language which typically carries informational and factual content, relational language is concerned with establishing and maintaining relationships with interlocutors. Also, he suggested that the corpus was able to inform some language patterns with particular discourse functions, e.g. the three-word chunk 'I don't know' is located at the beginning of the utterances very often with the function of hedging. The corpus study revealed an area for materials writers - some conversational strategies facilitated by usage of chunks (e.g. 'you know what I mean', which appeals to common ground or understanding). Importantly, McCarthy (2010) raised some problems of using real language in conversational corpus because they may appear very messy. This can cause challenges for teachers if they intend to use authentic materials in their teaching classroom. He suggested ways of adapting authentic materials in pedagogic practice by providing guidelines and indicating relevant implications for teachers' guidance.

O'Keeffe et al (2007) identified the key findings in corpus linguistics, particularly on spoken language corpora, which is relevant and transferable in terms of how they can inform pedagogy. They hoped that might lead to pedagogical insights for language teachers about their teaching practice, or choices of teaching materials. They structured their book by using single words as a starting point and moving to strings of words (chunks) and subsequently into pragmatics and discourse level. To achieve their aims, in

5-million-word CANCODE spoken corpus, they surveyed the most frequently occurring words and two-, three-, four-word chunks, etc. By creating a word frequency list, they gained the constitution of core vocabulary of spoken language. It would be an important index of what words should be covered in a basic syllabus in order to produce competitive communicators. Also, some single words and strings of words revealed that they served listenership and relational functions in spoken discourse for communications (O’Keeffe et al, 2007, chapter 2 and chapter 3). These functions had also been explored in terms of interaction and pragmatics (O’Keeffe et al , 2007, chapter 7 and chapter 8).

The core content and structure of the book (O’Keeffe et al, 2007) has motivated the research and become the focus of this dissertation. Other research about functions of interactional and relational languages also inspired interest in looking at the authenticity of coursebooks.

## **Chapter 3 Methodology**

### **3.1 Outline of the study**

Because it is impractical to examine all dialogues in the coursebooks due to time limitation, sample conversations in the coursebooks are selected for this research. Also, sample authentic conversations are selected for comparing. The scope and the focus of the research will be further illustrated in the following section.

The research involves three steps:

1. authentic conversations were transcribed into spoken discourses.
2. the language items under the scope of survey were collected from the sample conversations of coursebooks and authentic data.
3. the collected datasets were examined closely by employing a combination of quantitative and qualitative techniques.

It was hoped the findings would answer the following questions:

RQ1. Does frequency of the most common words and chunks testified by corpora of general spoken English differ between the coursebooks and authentic data?

RQ2. Are the common features of spoken discourse, particular interaction features comparable between the coursebooks and authentic data?

## **3.2 Data collection and analysis**

### **3.2.1 Research design**

As mentioned above, a combination of qualitative and quantitative methods was used for this research. Quantitative analysis was used to compare frequency of some common words and chunks with reference to the findings of corpora and existing known knowledge of spoken discourse as advised by researchers. Qualitative analysis was used to look at the common features of spoken grammar, particularly interaction and relational features.

### **3.2.2 Data source**

#### **3.2.2.1 Coursebooks and the users of the coursebooks**

The set of *The New Standard College English* (Second edition) is four books which have been used widely in the universities in China. In each of the books, there are eight units with two or three dialogues in each unit. The transcripts of the dialogues are included in the coursebooks and Videos of the dialogues (DVDs) are also provided. The videos of book 1 and book 2 are filmed in Oxford. They involve three students Mark (English), Kate (American) and Janet (Chinese) talking about their lives and their studying at the university of Oxford in England. The videos of book 3 and book 4 are filmed in London

and these are about stories of three colleagues Joe (English), Andy (American) and Janet (Chinese) in London in England. They talk about their daily jobs, interests and concerns, and also provide their insights of the city of London.

The target learners of the coursebooks are students in universities of China who do not learn English as their major disciplines. The majority of them are young adults with ages ranging from 18 to 22 who have had six years' education in secondary schools. In general, their English level is equivalent to B2 level of CEFR in their year 1 and year 2. Most of the students pass College English Test 4 (CET4) in their year 1 or year 2 in the universities as required by university. (CET 4 is at the similar level of China Standards of English 5 (CSE 5). CSE 5 is equal to overall IELTS score 5.5. This is according to the article named 'IELTS successfully linked to China's Standards of English Language Ability', on official IELTS website).

Before the research for the dissertation was conducted, some informal interviews were undertaken with a few students. They were keen to acquire oral fluency in their English learning. It was also observed that they had obtained some abilities for independent learning and had motivations for speaking development. Based on the interviews, it is not unreasonable to suggest that it might be true in general of Chinese students in terms of English language leaning.



### **3.2.2.2 Authentic materials**

Regarding authentic materials in this research, two sets of authentic conversations were selected. One was audio clips of conversations quoted from the IViE Corpus (Grabe, E., Post, B. & Nolan, F. , 2001) which was created by Phonetics Laboratory in the Department of Linguistics, Philology and Linguistics of the University of Oxford. The other was two excerpts of conversations from TED Talks “The future we're building - and boring ” (April 2017) and “How Netflix changed entertainment - and where it's headed” (April 2018). The selection of the two sets of authentic data was for two reasons: (1) they are unscripted conversations which are expected to represent features of spoken language; (2) they have a variety of informal and formal contexts as the conversations (IViE) are in an informal context, whilst TED Talks are in a formal setting.

### **3.2.2.3 Data collection**

According to Littlejohn (2011, p. 186), the size of sample data is suggested as being a proportion of 10 percent to 15 percent of the total material for research purposes, ideally chosen around the midpoint for analyzing general nature of a set of materials. This method of data collection was adopted in this research:

With regard to coursebooks data, one dialogue was collected from units 2, 5, 7 of each of four coursebooks respectively as sample data. Therefore, 12 dialogues in total (12 versus

65 , accounting for 18% of total number of dialogues of the coursebook) were selected for analysis. 1550 words of conversations were collected from coursebooks 1&2 and coursebooks 3&4 respectively. By selecting sample dialogues evenly from the 4 coursebooks, it was hoped that the research would obtain an overall picture of the features of the coursebooks by examining the sample data.

The coursebooks data were then divided into two sets of data: dataset one included sample conversations from book 1 and book 2; dataset two included sample conversations from book 3 and book 4. This aimed to see whether particular features of spoken grammar were covered at lower level or higher level, as the language level of four books were assumed to have a gradual increase from book 1 to book 4. The sources of sample data are shown in following table 1:

Table 1: New Standard English (2016)

Source	Topics
Coursebooks 1&2	
Book1 Unit 2	eating in an English restaurant
Book 1 Unit 5	talking about the news
Book 1 Unit 7	talking about dating a girl
Book 2 Unit 2	talking about homesick
Book 2 Unit 5	going to the theatre
Book 2 Unit 7	talking about emperor

## Coursebooks 3&4

Book 3 Unit 2	Andy introduces Janet his hometown
Book 3 Unit 5	Janet learns about London's Chinese community
Book 3 Unit 7	Joe talks to Janet about his London hero
Book 4 Unit 2	Janet and Andy discuss Charles Dickens
Book 4 Unit 5	Janet and Andy discuss gender stereotyping
Book 4 Unit 7	Janet does an interview with an expert on buildings

---

In terms of authentic data, the authentic conversations from IViE corpus were transcribed first. Then the transcripts from the TED Talks were taken as they are provided on TED Talks official website. In order to make a data comparison of same size between coursebooks data (two sets) and authentic data (two sets: IViE corpus and TED Talks), equivalent words were extracted from authentic materials and coursebooks. That meant each dataset of authentic materials (IViE corpus and TED Talks) included 1550 words to match with the number of words in coursebooks dataset.

### 3.3.3 Data analysis

In general, the four sets of data were surveyed line by line, recorded, labeled for specific features of spoken language and the relevant language items were counted. The process of data analysis was as follows:

- 1) The most frequent words and chunks advised by corpora were compared among the datasets.
- 2) The similarities and differences of the frequency of common discourse markers and

chunks (single -word, two words, multiple words), as advised by corpus, were compared. Frequency comparisons were made with reference to corpora: the Cambridge and Nottingham Corpus of Discourse in English (CANCODE) (5 million words as described in O’Keeffe, McCarthy and Carter, 2007, chapters two and three). Because the CANCODE corpora has been designed to represent spoken discourse, it was suited to serve as a comparison tool to survey the spoken discourse features in the coursebooks.

3) Quantitative and qualitative examinations were used to explore which typical features of spoken English occurred frequently and which did not.

To do this analysis, spoken language features were analyzed from three categories, A, B and C. Category A and B included fixed lexico-grammatical units, the language forms of which cannot be changed grammatically (This is according to O’Keeffe, McCarthy, and Carter, 2007, chapter 2, 3, 7 and 8). Category A, which is ‘listening and response’, focused on the language items in the form of short utterances (e.g. yeah, mm, right, etc.) which are produced by listeners to signal that they are listening to speakers and wish the speaking continue, without intentions to take over the speaking turn. Category B, which is ‘relational language’, focused on language items used by speakers in conversations which have pragmatic functions in management of conversations, particular in the aspect of maintaining good relationship with listeners. The meaning of relational language talked about here is opposed to transactional language. The latter serves for information or service exchange, but the former shows speakers’ engagement with listeners when managing their utterances (e.g. well, you know, I mean, do you think, just, etc. ). Although

language cannot be differentiated strictly to either relational or transitional, the research focused on language relational functions – for example, discourse marking and hedging. Discourse markers (e.g. ‘right’) have different functions. They were categorized under either category A or B. The frequency of discourse markers was counted according to their functions in the conversations. Category C refers to the feature of overlapping which reflects both listeners and speakers’ engagement and contributions to conversations. The three categories were chosen with the intention of examining typical features of spoken language, and particular interactions in the interlocuters at the level of discourse.

## Chapter 4 Findings and discussions

In this chapter, two research questions will be answered in section 4.1 and 4.2. Section 4.1 is focused on comparison of the frequency of the most common words and chunks. The common words and chunks are based on CANCODE (5m) (O’Keeffe et al., 2007, p. 64). Section 4.2 is focused on interaction features (see section 3.3.3)

4.1 RQ1. Does frequency of the most common words and chunks testified by corpora of general spoken English differ between the coursebooks and authentic data?

### 4.1.1 Findings

- 1) Each of the sample discourses involves 1550 words in all the following tables
- 2) The frequent rank list of CANCODE (5m) is based on automatically extracted strings, referring to an occurrence of at least twenty times in the five-million-word (O’Keeffe et al., 2007, p. 64)

Table 2: Frequency of top ten most common words of CANCODE (5m words)

CANCODE Rank	CANCODE (5m words)	Coursebooks book1&2	Coursebooks book3&4	TED	Conversation Talks	Conversation (IViE corpus)
1.	the (169,335)	63	76	56	60	
2.	I (150,989)	56	51	11	39	
3.	and (141,206)	32	31	45	41	
4.	you (137,522)	55	42	44	65	
5.	it (106, 249)	43	49	40	76	

6.	to (105,854)	41	37	46	39
7.	a (103, 524)	32	23	38	38
8.	yeah (91,481)	0	0	14	16
9.	that (84,930)	22	28	41	28
10.	of (78,207)	16	31	29	31
<hr/>					
Total:	1,169,297(24%)	360(23%)	368 (24%)	334(22%)	433(28%)

Figure 2: Frequency of top ten most common words of CANCODE (5m words)

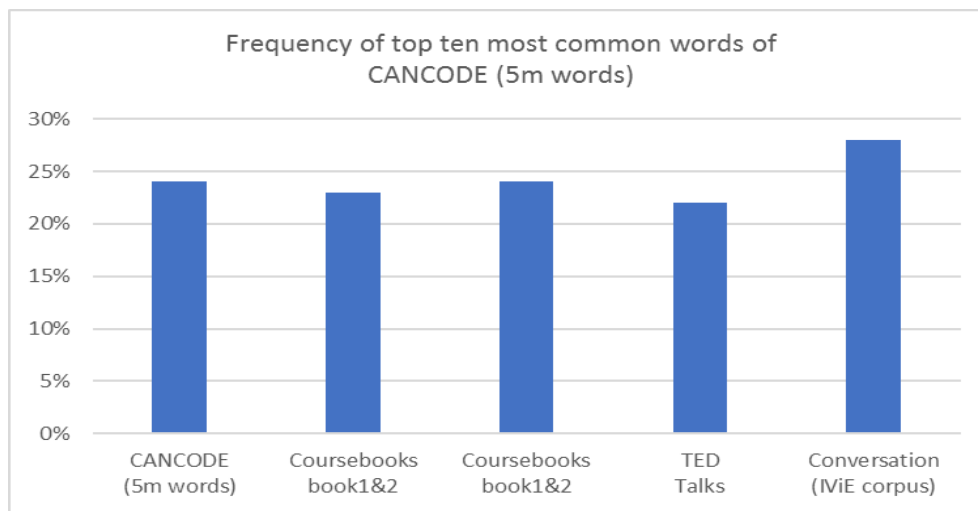


Table 3: Frequency of top ten chunks (two-word) of CANCODE (5 m words)

Rank	CANCODE (5m words)	Coursebooks book1&2	Coursebooks book3&4	TED Talks	Conversation (IViE corpus)
1.	you know	5	4	5	11
2.	I mean	0	1	3	4
3.	I think	1	6	3	11
4.	in the	6	7	3	2
5.	it was	2	1	2	1
6.	I don't	2	2	2	6

7.	of the	4	9	3	7
8.	and I	2	2	0	4
9.	sort of	1	0	1	2
10.	do you	4	4	1	3
<hr/>					
Total	137,247 (2.74%)	27(1.74%)	34 (2.19%)	23 (1.48%)	51 (3.29%)

Figure 3: Frequency of top ten chunks (two-word) of CANCODE (5m words)

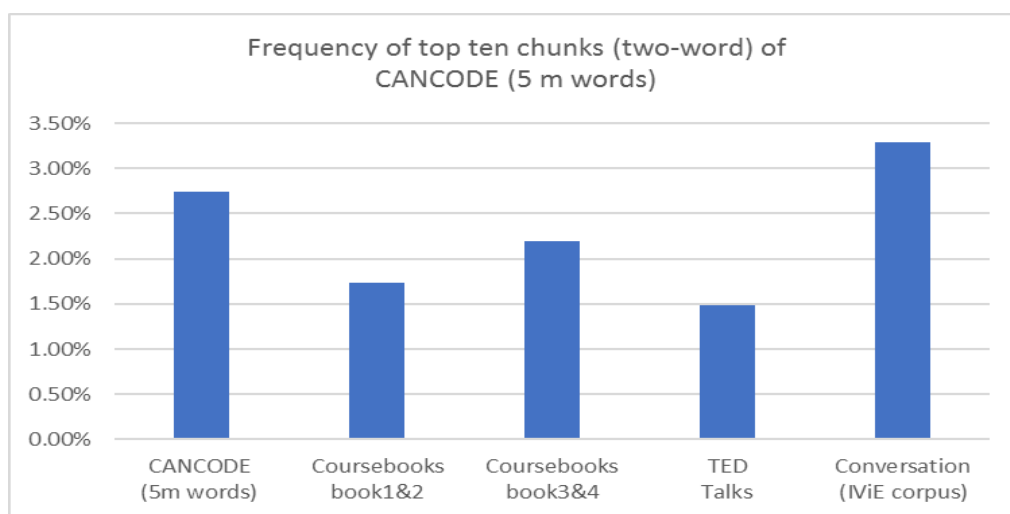


Table 4. Frequency of top ten chunks (three-word) of CANCODE (5 m words)

Rank	CANCODE (5m words)	Coursebook book1&2	Coursebook book3&4	Ted Talks	Conversation (IViE corpus)
1.	I don't know	1	1	2	2
2.	a lot of	0	1	2	2
3.	I mean I	0	0	0	0
4.	I don't think	0	1	0	4
5.	do you think	0	2	1	0
6.	do you want	2	1	0	0
7.	one of the	0	1	2	0



8.	you have to	0	0	0	0
9.	it was a	0	0	1	4
10.	you know I	0	1	0	0
<hr/>					
Total	20,613 (0.41%)	3 (0.19%)	8 (0.52%)	8 (0.52%)	10 (0.65%)

Table 5. Frequency of top ten chunks (four-word) of CANCODE (5 m words)

Rank	CANCODE (5m words)	Coursebook book1&2	Coursebook book3&4	Ted Talks	Conversation (IViE corpus)
1.	you know what I	0	0	0	1
2.	know what I mean	0	0	0	1
3.	I don't know what	0	0	0	0
4.	the end of the	0	0	2	0
5.	at the end of	0	0	0	0
6.	do you want to	2	0	0	0
7.	a bit of a	0	0	0	0
8.	do you know what	0	0	0	0
9.	I don't know if	0	1	0	0
10.	I think it was	0	0	0	1
<hr/>					
Total	4,982 (0.1%)	2 (0.13%)	1 (0.06%)	2 (0.13%)	3 (0.19%)

#### 4.1.2 Discussions

Table 2 shows the frequency of the top ten frequent words in the CANCODE (5m words) in the sample discourses of coursebooks, TED Talks and conversations (IViE corpus). From the data of coursebooks, it can be seen that the total percentage of the ten most frequent words account to 23% and 24% in coursebooks 1&2 and coursebooks 3&4

respectively, which are equivalent to the percentage (24%) of the top ten frequent words in CANCODE. They also show a similar trend from the highest frequency to lowest frequency, particular the frequency of the ten words in coursebooks 1&2. It is also observed that a few words may be used more frequently than others. This is probably due to the small sample collection of spoken texts, resulting in high frequent use of the words as they may connect to particular topics and contexts.

In addition, coursebooks show a lower frequency than expected in terms of use of “and’ and “yeah”. By observing the four sets of data, the percentage of use of “and” in TED Talks is closer to the data of CANCODE than the coursebooks. It is found that “and” occur 45 times with distribution of usage as follows: to link utterances by speakers for 36 times; to connect words or sentences serving grammar functions for 8 times; to work as a part of phrases for one time. By comparing this with usage in coursebooks, it is found that in coursebooks 1&2, “and’ is used for connecting words or sentence for 26 times which is mainly to serve for grammar functions, but it acts in function of linking utterances for 6 times. In coursebooks 3&4, “and” is used more than in coursebooks 1&2 (19 time for linking utterances and 12 for connecting words or sentences serving grammar functions), but is still less than TED Talks. To link utterances by using ‘and’ is a common usage in spoken language when speakers elaborate what they have said. The typical usage can be found in many places in TED Talks, whilst it does not occur very often in the coursebooks. Below are the examples of using ‘and’ for linking function.

Example 1 (from TED Talks)

“You started Tesla with the goal of persuading the world that electrification was the future of cars, and a few years ago, people were laughing at you. Now, not so much.”

With regard to “yeah”, as it belongs to one of the response tokens, it will be discussed in the following section (4.2.2.1).

Although not obvious from table 2, another noticeable point is that ‘er’ occurs in coursebooks with a significant lower frequency than that in authentic conversations (see table 8). The absence of the marker of ‘er’ is likely to reflect the nature that the dialogues are made up, in part, of scripted spoken language. From the video of the conversations, it is observed that speakers take their speaking lines without any hesitation. This might be due to them knowing what is coming next.

From table 3, it can be seen the total percentage of usage of top ten two-word chunks in coursebooks 1&2 is lower than that of CANCODE (1.74% versus 2.74%), though there is an evident increase in coursebooks 3&4 (2.19%). In contrast, these two-word chunks show higher frequency in conversation (IViE) (3.29%). Nevertheless, it is notable that TED Talks has the lowest percentage of usage of these two-word chunks.

To look at the data in table 3 closely, it is noted that ‘and I’ is absent in TED Talks and ‘I mean’ and ‘sort of’ are omitted in either in coursebooks 1&2 or coursebooks 3&4. The

absence of particular chunks will lead to missing the relevant functions because some chunks serve as discourse markers for different prominent functions in spoken language. However, it is interesting to find that the similar chunk 'and we' occur two times in TED Talks which might be due to the speakers (the interviewees are leaders of a big company) intending to use 'we' rather than 'I' to emphasize teamwork under many circumstances.

Other important discourse markers - 'I mean', 'sort of', 'I think' and 'you know', will be examined closely in the following section (4.2.1).

From table 4 and 5, it can be found that many of the three-, four-word chunks are absent in the sample discourses. This is possibly because the size of the sample texts (1550 words) was too small to show the level of frequency. Nevertheless, we can still find that three-word and four-word chunks are used with the most frequency (three-word chunks 0.65%, four-word chunks 0.19%) in conversations (IViE corpus) by looking at the percentage of different datasets in tables 4 and 5. These formulaic expressions are important in communication process of spoken language as they can be retrieved as a whole directly from our memory. This can help speakers on their utterances in an more efficient way, therefore more fluent (Wary, 2002, p. 36). However, Schmitt (2010, pp. 142) points out formulaic language is problematic for ESL learners which is caused by lack or misuse of it.

4.2 RQ2. Are the common features of spoken discourse, particular interaction features comparable between the coursebooks and authentic data?

#### 4.2.1 Findings

In tables 6, 7, 8 and 9, the words having function categories A and B (definition of categories A and B refer to section 3.3.3 on page 34) are counted and highlighted if they are in the list of the most frequent single words, two-word, three-word or four-word chunks in CANCODE (5m words) (the list of common words refer to chapter 4 on page 36).

Table 6. Frequency of common spoken discourse features (category A , B, C)

	Coursebooks book1&2	Coursebooks book3&4	TED Talks	Conversation (IViE corpus)
Category A Listenership and response				
Response token	8	6	3	22
Category B Relational language				
(1) Discourse marking	29	26	58	67
(2) Hedging	7	13	23	60
Category C Overlapping				
	0	0	4	15
Total:	44	45	88	164

Table 7. Common words serving for Category A - Listenership and response

CANCODE (5M)	Coursebooks book1&2	Coursebooks book3&4	TED Talks	Conversation (IViE corpus)
yeah (top 8)	0	0	2	7
mm (top 16)	1	0	0	0
oh (top 24)	0	0	0	0
right (top 30)	1	0	1	1
Others(out of top 30)	6	6	0	14
Total:	8	6	3	22

Table 8. Common words serving for Category B - Discourse marking

CANCODE (5m)	Coursebooks book1&2	Coursebooks book3&4	TED Talks	Conversation (IViE corpus)
Yeah(top 8)	0	0	12	9
er (top 17)	3	1	20	18
so (top 19)	4	4	13	5
oh (top 24)	4	2	0	6
well (top 27)	13	15	1	13
right (top 30)	2	3	2	0
you know (top 1)	3	0	5	10
I mean (top 2)	0	1	3	4
I mean I (top 3)	0	0	0	0
You know what I (top 1)	0	0	0	1
Know what I mean (top 2)	0	0	0	1
Total:	29	26	58	67

Table 9. Common words serving for Category B - hedging

CANCODE (5M) Top 30	Coursebook book1&2	Coursebook book3&4	Ted Talk	Conversation (IViE corpus)
like (top 26)	1	3	6	29
Just (top 31)	5	4	12	18
I think (top 3)	0	5	3	11
I don't know (top 1)	1	1	2	2
Total:	7	13	23	60

#### 4.2.2 Discussions

##### 4.2.2.1 Category A Listenership and response

Category A refers to some fixed lexico-grammatical items which serve a common feature of spoken interactions - these are “listener response tokens”. When listeners have no intentions of taking over speaking turn, they use the items to show they are responding to their speakers. By using the lexical units appropriately, listeners take a more active, responsive role in conversations. O’Keeffe et al (2007, p. 142) refer to this as ‘listenership’ and state ‘good listenership is natural and desirable for efficient spoken communication.’ Also, Carter and McCarthy (2006, pp. 190-191) suggest that response token can simultaneously signal a boundary and to signal agreement or simply to express friendly social support. The response tokens have different forms: ‘minimal response’ (e.g. mm, um hum), ‘non-minimal response’(e.g. really, definitely), clustering of response’(e.g.

yeah, mm), ‘negation of response’ (e.g. absolutely not). In this dissertation, the term “response token” (as described by O’Keeffe et al., 2007, p. 142) will be used as an umbrella to cover the different forms of the lexical items which are used by listeners to respond to speaker. The findings of frequency of response tokens and their particular functions in coursebooks and authentic conversations will be discussed below.

From table 6, it can be found that coursebooks 1&2, coursebooks 3&4 and the TED Talks have a lower frequency of response tokens than the frequency that occurs in conversations (IViE) (8; 6; 3 versus 22). It is noticeable that the TED talks show a lower frequency than others which is probably because the conversations in TED Talks share nature of interview in which both the guest speaker and the presenter are taking more roles as speakers rather than listeners. In the conversations, the guest speaker as interviewee is expected and encouraged to share more information - meanwhile, the presenter needs to take the lead of the conversation. As both speakers in TED Talks need to contribute their views in the talk rather than being a role of active listeners, the reason for low frequency of usage of response tokens may be explained. However, it might be normal in real-life conversation where one speaker takes more of the role of speaker and others act as listeners. Comparing the occurrence of response tokens in coursebooks 1&2 and 3&4 with the occurrence in conversations (IViE) (8 and 6 versus 22), it is clear that the response tokens are not used in coursebooks as adequately as in actual dialogue to demonstrate listenership actively. For example, in an excerpt of conversation (IViE) (example 2) below, it shows listener use response tokens to register that they are following



the speaker's drift. Examples 3 -7 are from coursebooks to be used for contrast.

### Example 2

S1: It is so dirty

S2: Really?

S1: Unbelievable!

S2: What's it's like? I've never been in

S1: It's like, uhm, you know, corrugated iron

S2: Yeah.

S1: It's like corrugated iron round it and - one bit's like spray painted.

S1: And you don't walk on the ground, you walk on the fag bach a it's like so horrible and you know we have to actually clean it. You're in litter duty.

S2: Yeah.

S1: There is a point when we have to clean it and I think it's horrible and I think any people who smoke should have to clean it.

S2: Definitely.

### Example 3

A: Well, it certainly used to be polluted. I remember it had a very distinctive smell. If you ...

J: That sounds revolting!

Example 4

A: I don't know if Janet's told you, but we're doing a series of reviews on the ethnic restaurants here ...

T: OK, I see.

Example 5

A: What I suggest is, if you've got time, we'll do some filming around here, and then we'll ...

J: OK.

A: Just remember to keep the needle out of the red zone. And don't forget to ...

J: OK, thanks.

Example 6

J: That's right. In fact, it's the first thatched building in London since ...

J: That's amazing.

Example 7

J: Yes, there are some seats, but most people stand while they watch the play. So they get ...

J: That's extraordinary!

Example 2 is from conversation (IViE) in which speaker 2 (S2) is taking the role of listener who wants to signal that she is listening, showing interest and using response tokens to allow the conversation to continue smoothly. For example, by using ‘really’, S2 shows her strong emotions of surprise and involves herself in the conversation. ‘Yeah’ is used to encourage speaker to keep talking and to continue the conversation. ‘Definitely’ demonstrates her agreement to speaker 1’s (S1) opinion. The response tokens make the conversation a successful communicative talk which show interlocuter’s mutual contributions.

Examples 3 - 7 are all the examples of response tokens used in coursebooks 3& 4. Though few response tokens are used, various functions are covered such as: engagement or assessment (e.g. ‘sounds revolting’, ‘fascinating’, ‘wow’, ‘happy ending’, ‘cool’, ‘that’s amazing’, ‘that’s extraordinary’), acknowledgement and show understanding (‘ok’), agreement (‘right’), strong emotion (‘you are joking!’, ‘oh, he didn’t!’ ). However, it is found that, in the coursebooks, the response tokens are not used continuously in the on-going interactive progress of the talk which might be because the speaker does not further develop his or her utterances after getting a response from the listener.

In Table 7, it is interesting to note that “mm”, the top 16 most frequently used single word (O’Keeffe et al., 2007, p. 35) is not used as expected from what was found in CANCODE (5m), and in fact, it is almost absent in all the four datasets (except “mm” occurs once in coursebooks 1&2). ‘Mm’ is used as an acknowledgement by listeners to show their

listenership. Likewise, it is perceived by speaker as a floor-yielding signal which marks the listener's desire for the talk to continue (e.g. O'Keeffe et al., 2007, p. 149; Carter et al., 2016, p. 173). The absence of 'mm' in the samples of the conversations may not mean it is not a common or frequent word with functions of response token, though the reason needs to be further explored.

Also, in Table 7, by looking at the similar function of 'yeah', it is noted that 'yeah' (the top 8 most frequent used words CANCODE 5 m) is missing in the coursebooks (see table 2), whilst it appears 24 times in the conversation (IViE). 'Yeah' is called 'response taken' and being used to show that we are listening to and interested in what is being said (Carter et al., 2016, p. 173). By looking at 'yes', the synonymous forms of 'yeah', we find that it is used six times in coursebooks 1 &2 and eight times in coursebooks 3 &4. According to Tao (2003, p. 201), 'yeah' has a lot more frequency than 'yes' and they are substantially different in terms of function. With closer inspection of the usage of 'yes' in coursebooks 1&2, it is found that 'yes' is used 4 times to answer questions, but only 2 times used as response token. However, in conversation (IViE), "yeah" is used very frequently as response token which shows a significant feature of interaction of spoken language. The marker 'yeah' provides a fairly non-committal response to what has just been said (Thornbury, 2005, p. 66).

Kendon (1967, p. 53) indicates that another function of response token is that they can serve as a guidance for speakers as to how the message is being received. For example,

to request information by using a response token. In dialogue, understanding the meaning is a mutual concern of both of the speakers. Interlocutors need to respond immediately to what the speaker has said. The following excerpt (example 8) reflects the relevant function:

Example 8 (from conversations (IViE)):

S2: They're changing it though aren't they. They're changing the smoking hut

S1: What? at Long Road?

S2: No, here

S1: What?

S2: They're building a new one

S1: Are they?

S2: They've got a new one with a building and they are taking the old one down I think

S1: Are they?

S2: I think that's what Mr Holmes's saying

The above example reflects a typical interaction feature in the conversation: speaker 1 (S1) uses the short question utterance to reflect her surprise and request for more information, while speaker 2 acts more as a speaker who adjusts her speaking message according S1's response. It is evident that how to signal that the listener understands what has been said is crucial to ensure the dialogue develops.

#### 4.2.2.2 Category B Relational language

##### (1) Discourse marking

Table 8 examines language items which are used commonly with discourse marking functions which are called discourse markers (DM). DMs are one of the important resources for listeners to indicate their involvement with what is being said and to manage their own responses (Carter and McCarthy, 2006, p. 221). O’Keeffe et al. (2007, p.172) define discourse markers as words and phrases, being structured outside of the clause, which link segments of the discourse to one another in ways for speakers to make choices of different functions in conversation. It works as important devices in: (1) managing and organizing conversation (e.g. markings of shifts and boundaries through talk) (2) maintaining relationship with listeners (e.g. marking of responses in conversation). It is common that speakers constantly signal their intentions to listeners all through conversations to show what they are going to say is connected to preceding utterances or what is expected to come up. They deliberately do this to smooth the cut-and-thrust nature of interactive talk.

The research has examined the data to see if common discourse markers are used with a comparable frequency in coursebooks, authentic data, and a general reference corpus. As mentioned in chapter three, the lexical items with discourse markers functions were chosen for examining. They occur with high frequency (single-word lists in the thirty

most frequent words, two-word, three-word, four-word chunks list in the top three most common chunks separately) in corpora of spoken English (CANCODE, (O’Keeffe et al., 2007, p. 35/65). The significant findings will be examined in detail and discussed in below.

**a. The coursebooks have the lowest frequency of the discourse makers.**

According to the frequency counts in Table 8, it is clear that the frequency of many of the discourse markers in coursebooks do not have similarities to the corpus data of CANCODE (5m), except for ‘well’. The frequency of total number of discourse markers and occurrence of each discourse marker differ within each dataset and across different datasets. The finding shows the number of discourse markers used in coursebooks is lower than authentic datasets, which is less than half of that in authentic conversation (IViE). However, a number of discourse markers are used as linguistic devices by speakers to manage and make their speech as smoothly as possible. A notably absence of discourse markers from the conversations in coursebooks, can result in the ‘failure of spontaneous and collaborative construction of talk’ (Thornbury, 2005, p. 78).

**b. ‘Yeah’ is missing in coursebooks.**

Previously, lexical item ‘yeah’ has been discussed as a response token which functions to show listenership without intention to start an utterance. However, it is noted that ‘yeah’ is used broadly in authentic conversation in TED Talks and the conversation (IViE). It is positioned at the beginning of the utterance by speaker to respond to what has been said and followed by his or her own turn and utterance. It appears not only to respond to previous speech but to start a new turn.

Example 9 (from TED Talks): Yeah

CA: Won't ever have to touch the wheel -- by the end of 2017.

EM: Yeah. Essentially, November or December of this year, we should be able to ...

**c. 'Well' occurs with the most frequent in coursebooks and across authentic data.**

McCarthy (2008, p. 34) states items such as 'well', 'actually', 'just', 'so', 'cos', 'like', and many others occur with extremely high frequency in any native speaker conversational corpus, but notably of much lower frequency in written corpora. They are typical lexico-grammatical items used in spoken language with significant interactive force.

According to Carter et al. (2016, p. 373), 'well' has the following functions:

1. It is positioned at the start of our utterance with the function to show what speakers are thinking about, what they are going to say, or how to organize their speeches.
2. It is used to show a slight change in topic, or when we are going to say something which is not quite as expected.
3. It is used to change what has been said slightly or express opinion in another way.
4. It can be used to admit or acknowledge that something is correct or true.
5. It can be used with a rising intonation to function as a type of question when we desire or wait for someone telling us something.



Well' is the twenty seventh most common single word in corpus of CANCOCD (O'Keeffe et al., 2007, p. 35). The frequency count in Table 8 shows that 'well' occurs 13 times in coursebooks 1&2 and 15 times in coursebooks 3&4 which has the equivalent occurrence in IViE corpus. However, it is rarely used in TED Talks. The following examples show how 'well' is being used in coursebooks. By examining the usage of 'well' in the coursebooks 3&4, some functions can be found among the 15 instances of using 'well': nine of them are used as the above function number 1, six of them are used as the above function number 2. The other functions mentioned above are missing in coursebooks. The finding suggests that the high frequency of using 'well' does not mean the coverage of various common functions as advised by corpus. "Well' in following excerpts 10.1 and 10.2 are examples for the above functions 1 and 2 respectively:

Example 10 (from coursebooks 3&4): Well

10.1

J: Does Tower Bridge still open?

A: Well, not so often.

10.2

J: Has anyone read any of the books?

A: Well, Joe, there are over 20 new books coming out next month, so ...

**d. 'You know', 'I mean' are used with lower frequency in comparison with authentic**

**data.**

Two two-word chunks ‘you know’ and ‘I mean’ are missing in coursebooks 1&2 and coursebooks 3&4 respectively, whilst in the CANCODE corpus (5 m), ‘you know’ is notably the most frequent two-word chunk. The total number of their occurrence in coursebooks and in the other two sets of authentic data, 4 versus 22, shows a big gap of the usage. Though they are the most common two-word chunks in CANCODE (O’Keeffe et al., 2007, p. 65), it appears they are not the most common chunks in the coursebooks.

In spoken English, ‘you know’ seems to be ubiquitous in spoken language which often marks what we think is old, shared or expected knowledge (Carter, McCarthy, Mark, & O’Keeffe, 2016, p. 173), or uncontroversial or logically linked knowledge (Carter, & McCarthy, 2006, p. 211). Speakers refer to shared knowledge constantly and appeal for agreement by using markers like ‘you know’ ... (Thornbury, 2005, p. 66). ‘You know’ is often used as an important signal of (assumed or projected) shared knowledge between collocutors or works as a pause marker (O’Keeffe et al., 2007, p.71). This is the way in which ‘you know’ is used in coursebooks to indicate shared knowledge to the listener when a new topic is launched.

Example 11 (from coursebooks 1&2): You know

11.1

M: I’m not usually shy, but – she’s so ... you know ...!

J: Oh, Mark!

11.2

M: The play which OUDS are producing. You know, the play I'm in ...

K: Oh, that play! Well, er ...

11.3

J: Well, it was nominated for an Oscar, you know.

K: That figures. It's a beautiful film.

The similar function of 'you know' is demonstrated in many instances of TED Talks and the authentic casual conversation (IViE). However, its missing in coursebook 3&4 might be due to the contrived nature of the interaction in the coursebooks, or it is being arranged purposely as it is already shown in lower level coursebooks (coursebooks 1&2). Either of the above reasons suggests that if they were authentic dialogue, 'you know' might appear naturally in the discourse because they are ubiquitous in spoken language.

'I mean' seems to function largely as a discourse marker in daily conversations, as shown in CANCODE, which is mainly used by speakers to reformulate or clarify something they have just said. In the following example 12, 'I mean' is used by speaker in the coursebooks to explain what he means by 'gender and racial stereotyping' to the listener, to make his utterance understandable. This is an example representing typical features of spoken language. However, it occurs only once in the sample conversations of the four

coursebooks. It might not be adequate for learners to be exposed to and learn.

Example 12 (from coursebooks 1&2): I mean

A: Gender and racial stereotyping. I mean, it depends on what job and sometimes where you work in London ...

## **(2) Hedging**

Table 9 examines language items which are used commonly with hedging functions for purposes of face, politeness, vagueness or approximation of softness. These language items are selected because they are common words with high frequency in spoken language. Similar to discourse marking functions, their frequency is compared among sets of data involving coursebooks and authentic data. The number of the usage of common linguistic items with hedging function in table 9 shows a gradual increase of usage from lower level coursebooks (coursebooks 1&2), to higher level coursebooks (coursebooks 3&4), TED Talks and authentic conversation (IViE). It is noticeable that the hedging function is used very rarely in coursebooks 1&2, though there is an increase in coursebooks 3&4. The frequency of the usage of the function in coursebooks 3&4 is only half of the number used in TED Talks and less than twenty percent of that used in conversation (IViE).

Hedging functions are important in interpersonal and interactive communication because

it has the important feature of interpersonal meaning. They indicate a range of expressions that are used in everyday spoken language to downtone the assertiveness of a segment of discourse (O’Keeffe et al., 2007, p. 175). For example, hedges play a crucial part of a polite conversation which is used by speakers to soften what they say or write, to make what we say less direct. A variety of markers serving the function of hedging are used by speakers to make their utterances not to sound too blunt and assertive (Carter and McCarthy, 2006, p. 223).

‘I think’ ‘just’, ‘like’ are used as hedging functions. Their usages are examined in the rest of this section.

‘I think’ is normally used by speakers to make utterances less assertive and less open to challenge for politeness and considering interlocuter’s face. From table 9 it can be seen that “I think’ appears five times in coursebook 3&4, but it is absent in coursebook 1&2. It’s not certain whether it is arranged purposely to introduce the chunk item in the higher level coursebooks. However, low level students may not expect to be introduced with the hedging function so late.

With regards to ‘just’, it can also serve as a softener in spoken English to downtone or soften utterances (Carter & McCarthy, 2006, p. 99). The word ‘just’ occurs frequently in a stretch of talk which is relaxed and secured by using the word to maintain a supportive and friendly tone and produce a softening effect. The omission of it makes the utterance

more direct and definite with less positive interaction (Carter & McCarthy, 1997, p. 109). The following example 13 gives the instances of using ‘just’ as a hedging function. There are no other different ways of usage detected in the authentic conversations (IViE) as far as hedging function is concerned. However, ‘just’ is obviously used more frequently than that in the coursebooks.

Example 13 (from coursebooks 1&2): Just

1. M: This is just so crazy!
2. K: ... How about just walking up to her and saying hi? Why don't you do that?
3. K: Yes, I'm just doing an essay. But it's great to see you.
4. K: Just tell him what you're up to.
5. M: Er ... Just a coffee.

In Table 9, the word ‘like’ (the top 26 most common words in CANCODE (O’Keeffe et al., 2007) is highlighted. ‘Like’ serves hedging function pervasively in spoken language and it can be heard very commonly in informal speaking. When ‘like’ serves as hedging function, it acts as vague language to soften expressions so that they do not appear too direct or unduly authoritative and assertive. There are times where it is necessary to give accurate, direct, and precise information in many informal contexts, but there are other times speakers prefer to make their choices to convey information which is softened in some way (O’Keeffe et al., 2007, p. 177). The deliberate using of vagueness is motivated which often marks the skill and sensitivity of a language speaker (Carter et al., 2006, p.

202). A number of functions of 'like' in spoken language related to hedging have been noted by Carter et al., (2016, p. 268). For example, to avoid making utterance too assertive, 'like' can be used at the end of our utterance to modify or soften what we have just said, especially if we are not certain whether it was the right thing to say. O'Keeffe et al., (2007, p. 189) state it is used to make statements approximate, and to add a note of deliberate vagueness to expressions. It is interesting to compare the different usage of 'like' in coursebooks and conversation (IViE). It can be clearly seen that their functions are significantly different in the two datasets (the number in brackets is the frequency of 'like' in each of functions).

'Like' occurs 11 time in coursebooks 3&4 with below functions distribution:

- with similar meaning 'I like something' (8)
- phrases 'feel like' (2)
- vague expressions to make groups or categories (1)

'Like' occurs 32 times in conversation (IViE) with the following functions:

- make statements approximate such as 'it is like' 'like ....' (18)
- stuff (things) like (6)
- filler (6)
- with similar meaning 'I like something' (1)
- attention to number (1)

In general, hedging function is not used adequately in coursebooks. For example, some of the conversations in coursebooks 3&4 are between friends and some of the conversations are between colleagues. Considering the function of hedging is to soften the utterance and to avoid making utterance too assertive, hedging should be used more frequently when the conversation is between colleagues. However, at present it is not the case based on the findings in the research.

#### **4.2.2.3 Category C Overlapping**

Overlapping in a conversation refers to the situation that other participants start speaking before the speaker has finished. It is noted that there is no overlapping in the coursebooks, although many cases of overlapping occur in the conversation (IViE) (0 versus 14). It is observed that participants of a conversation take turns to speak. In other words, they do not talk at the same time or leave an undue pause. However, overlapping talk may occur in any conversation for various reasons: speakers are competitive and each of them try to get the floor; or speakers show a high degree of collaboration and they follow each other's utterance very closely. It is normal in fast-paced talk, where overlap happens when other participants contribute to complete the message of the speaker as they can project how an utterance might continue, and the appropriate time for their contributions (Culpeper, Katamba, Kerswill, Woak, & MeEnergy, 2009, p. 503). If interlocutors have good interactions between each other in a dialogue, the conversation can continue and potentially have more chances to achieve the aim of the communication successfully.



In the example below, taken from the conversation (IViE), overlapping occurs when speaker 1 (S1) follows the speaker 2 (S2) very close and responds actively. In turn 1, S1 expresses that he never heard of Bernie Eccleston and asks what it is about. Then in turn 2 and turn 4 when S2 explains it to him, S1 realizes whom Bernie Eccleston is and responds to S2 “oh, I think I know it’. However, at the same time, S2 is trying to complete his message, so the overlap occurs:

(The sign “ [ “ indicates an overlap)

1. S1: well, I wouldn’t know about Bernie Eccleston, never heard of him. What is it anyway?

2. S2: + oh, it’s just - you know (?) donation to the labour party, [and got tobacco advertising

3. S1: [oh,

4. S2: and banned [for a while, but would it affect you? [Tobacco advertising

5. S1: [oh, I think I know it [affect me?

Based on the research findings, it is suggested that the coursebooks do not represent overlapping, the typical features of spoken English occurring in many cases of real-life conversations. It is also worth mentioning that TED Talks do not have overlap as much as that in the sample conversations (IViE). One possible reason might be that both speakers try to avoid talking at the same time for consideration of maintaining a polite

talking under formal setting.

Overlapping seems a common feature of conversations which occur in real life spoken language. Lack of it in the conversations of coursebooks suggests that the coursebooks do not provide a realistic model of spoken language to students who need to notice the language forms which are especially salient in the communicative uses of spoken English.

### **4.3 Other findings**

#### **(1) Turns**

Conversation participants interact with each other through the change of turns. In this dissertation, the definition of turn is adopted as described by Tao (2003, p. 189) - 'it is understood in the broadest sense, that is, any speaker change will be treated as a new turn'. Comparison of the four datasets shows that they have different numbers of turns: coursebooks 1&2, 146 turns (11 words per turn); coursebook 3&4, 108 turns (14 words per turn); TED Talk, 53 turns (30 words per turn); conversation (IViE), 81 turns (20 words per turn). It is interesting that it appears the turns in coursebooks are shorter and of equivalent length. As previously mentioned in the section of relational language, speakers usually use discourse markers to signal what they are going to say, and their intention to launch new topic or continue previous utterances, to clarify or rephrase, etc. If each turn is short without adequate development, it might not provide spaces for speakers to use the relevant lexical items to demonstrate interaction features of spoken language

sufficiently. This will result in the lack of relevant strategies to organize or construct the utterances when students need to participate in a discussion or express their opinion, as they do not have appropriate samples for learning. As students who learn from the coursebooks are not exposed to authentic use of language, it is not hard to imagine they will have difficulties in constructing their turns when they are involved in discussion in depth or need elaboration of their opinions.

## **(2 ) Speed of utterance**

It is observed that the speed of each streaming of the utterances (coursebooks 1&2 145.8/m, coursebooks 3&4 139.2/m, TED Talks 144.4/m and conversation (IViE) 187.8/m) show a great distinction between the speed of delivery in the coursebooks and that in conversation (IViE). We can find that the speed of TED Talks is equivalent to that in coursebooks, but if we listen to recording, we can find a significant difference between the two recordings. There are many places of laughter and pauses (especially due to laughter made by audience) during the process of some utterances in TED Talks, but there is rare laughter in the conversations of coursebooks. This may explain the reason for a lower speed of utterance of TED Talks than it is expected. If students get used to listening to dialogues with lower speed of utterance than the actual speech in real life, they will get familiar with that speed. The direct impact to the students would be the difficulty for them to follow the message of speakers in real life which is expected to be spoken with faster speed. On the other hand, they might speak with lower speed in real communication

which appears unnatural and odd and even difficult for native speakers to understand.

### **(3) Context and genre**

It is noted that most of the conversations in coursebooks occur among the same three characters who are friends and colleagues talking about something happening around their lives. Adolphs and Carter (2013, p. 38) mention four relationship categories: intimate, sociocultural, professional, transactional and pedagogic. The sociocultural category refers to interactions between friends and the professional category includes discourse that is related to professional interactions. In coursebooks, most of the participants' relationships are sociocultural and a few are professional.

However, the relationships of participants vary in actual life in which participants involve in both symmetrical and asymmetrical conversations. For example, speakers are equals when speaking to friends, but they also have encounters in classrooms, restaurants or shops, in which different social relations obtain. Therefore, students need to be exposed to spoken language where speakers are in various relationships.

In addition, the fixed relationship of the main participants in the conversations may lead to the lack of a variety of contexts and missing of particular lexis. O'Keeffe et al. (2007, p. 153) discuss some context-specific functions where some response tokens occur under particular contexts. Also, McCarthy (2003, p. 45) indicates items such as "fine" and

“excellent” are frequently associated with closing of deals, making of arrangements, achieving decisions or completing satisfactory transactions. Students need opportunities to watch and listen to real conversation in a range of genres, and in a variety of social contexts and interaction environments. There may need to be an emphasis on both informal speech genres and on more formal varieties.

Tomlinson (2017a, p. 5) states that many teachers indicate that the coursebooks they are using lack context authenticity. For university students living in EFL countries, they may have opportunities to be exposed to various contexts. For example, students may listen to lectures of foreign scholar visitors, they may wish to greet to them, ask questions or chat with them appropriately. Besides, service encounters include a variety of situations which students may encounter on and off campus when they live and study abroad. Unfortunately, the above-mentioned genres are either missing or not adequately provided. If students had learned the relevant examples of dialogues and been familiar with the context, they would be prepared for communication very well and would be more confident to deal with the communication in spoken English. It may not be appropriate to say the contexts included in coursebooks are not relevant to students’ lives and studies, but there are some contexts potentially encountered by students that are not covered.

#### **(4) Laughter**

Thornbury (2005, p. 66) states that casual conversation is often punctuated by laughter, or at least chuckles. However, by listening to the audio/video of sample conversations, it

is noticed that there is more laughter and chuckles in authentic dialogue. The frequencies are: coursebooks 1&2 (2 times), coursebooks 3&4 (1time), TED Talks (12 times) , authentic conversation (IViE) (8 times). It shows there are more active and various responses in authentic data which is not exclusively or simply verbal phenomenon.

## **Chapter 5 Conclusion**

### **5.1 Conclusion of the research**

#### **5.1.1 Summary of the findings**

From the small size survey of spoken discourse in the coursebooks, it appears that the percentage of the coverage of frequent words in coursebooks are equivalent to that of corpus CANCODE (O’Keeffe et al., 2007, Chapters two, three). However, it seems that coursebooks do not use many of the most frequent and common chunks, particularly two-word chunks of the CANCODE (5 m), and it shows a low frequency of the usage of the language items in comparison with authentic conversations (IViE).

Interaction features are represented frequently in authentic dialogues, but it was found that some typical interactive features were missing or lacking from the coursebooks - these included response and listenership (e.g. “yeah”), discourse marking (e.g. “you know”) and hedging (e.g. “just”) of spoken language.

Another interaction feature (slightly different from those in previous paragraph) is overlapping. This was also absent from the conversations of coursebooks which occur commonly in spoken English.

There is evidence that some of discourse features are reflected in the dialogues in coursebooks. For example, some discourse markers are used for response and turn management and constructions ('well' is used pervasively). Nevertheless, the gap exists between the typical features of spoken language and what is shown in the coursebooks.

Finally, there are some other findings including length of turn and context. Generally, the utterance in each single turn in coursebooks is shorter compared with that in authentic conversations. With respect to context, as mentioned in chapter 4, some contexts related to Chinese students living and study circumstances are not being covered.

The findings suggest it is unlikely that coursebooks are a realistic representation of typical conversational features since they do not match adequately with authentic data. This is reflected in four situations: (1) use of common words, (2) lack interaction features, (3) missing overlapping, (4) missing variety of turns and context.

### **5.1.2 What have researchers said**

As having said in the previous section, some chunks are not used adequately in coursebooks to serve particular discourse functions. However, it was pointed out by researchers (e.g. McCarten & McCarthy, 2010, p. 15) that some language patterns and chunks have particular discourse functions and play an important role to facilitate conversational strategies, particularly two-, three- four-word chunks etc. They indicated



that an important function of relational languages, in contrast to ‘transactional’ language which typically carries informational and factual content, is concerned with establishing and maintaining relationships with interlocutors.

As mentioned in chapter two, coursebooks are contrived in many cases. As Tomlinson & Masuhara (2017, p. 31) have stated, it overprotects the learners and contradicts what is known about how languages are acquired and ... does not prepare them for the reality of language use outside the classroom’. Therefore, it appears they do not help learners for their language acquisition as they are not showing language not in use.

The inadequacies of coursebooks leads to a failure to realize authenticity. As mentioned in chapter two by Tomlinson (2010, pp. 1-6), authenticity has various aspects: text authenticity, task authenticity, learner and teacher authenticity, context authenticity, etc. The consideration of representing target language in typical use is only one element to achieve authenticity. These aspects matter - e.g. how students interact with the materials; whether they engage with the texts when using them; whether they perceive the relevance or value of the materials. ‘The ideal is therefore for the designers to try to ensure that their materials achieve authenticity in design, in use and in reflection.’ (Tomlinson & Masuhara, 2017, p. 33). Although these aspects are out of scope of this research, they are important topics related to authenticity realization which are currently being researched. They are crucial elements to be considered for achieving text authenticity.

### 5.1.3 Suggestions

If learners want to communicate flexibly and effectively for collaborative ideas in a range of spoken contexts, they should be familiar with the multiple lexico-grammatical choices. Researchers advocate ‘noticing’ in English learning (e.g. Ellis, 2016, p. 212). However, if the texts have not exposed language features adequately for students to notice, ‘noticing’ will not help learners’ language acquisition.

However, due to various reasons, authentic spoken discourse cannot substitute the contrived coursebooks completely. It is suggested using at least some of authentic materials as supplementary materials in a spoken language class. By comparing the difference of language in the coursebooks and the authentic conversations, students may acquire awareness and notice how language can be used appropriately in speaking.

Because most of the recordings of real conversations need to be transcribed into spoken discourse for learning, the access to transcripts in social media is limited (as mentioned by Jones in chapter two). Choosing TED Talks as one source of authentic materials can be an alternative option for teaching and learning with the consideration that TED Talks is one of few convenient website resources with both video and transcripts provided and is accessible worldwide. With TED Talks, it is found that they have been used as a resource to study how to teach notetaking (Siegel, 2019). By using excerpts of TED Talks as authentic data for comparison in the research presented here, a brief and general

understanding of it has been gained, though further research in depth might be required to explore TED Talks' characteristics and potential value for using it in English teaching class. The small size of spoken discourse analysis shows that TED Talks represent some features of authentic conversations and it offers some insights into the benefits of using TED Talks as a source of authentic spoken language. These are:

- a. TED Talks can help students to notice language features as an authentic use of spoken language in formal setting (ask questions, express point of view, narrate a story, etc.)
- b. TED Talks can be used for students to notice different language features and lexicogrammatical options by comparing it with conversations in a more casual setting.

However, TED Talks does have limitations - for example, its formal setting context, which lack many contexts such as service encounters, and unchanged presenter (though guest speaker is different each time). Therefore, dialogues with different contexts should be accessed as supplementary materials of authentic teaching materials, though the resources are limited. Finally, the most important thing is to understand what Chinese students need and want, which is crucial for materials selections and adapting to realize the authentic interaction between students and teaching materials and achieve authenticity in teaching at the end.

## **5.2 Implications of the research**

### **5.2.1 To teachers**

Though this research may not be a comprehensive survey of the coursebooks, if teachers agree the coursebooks are not reflecting adequately the features of the spoken language, then they may have their own insights into how to adapt the books for their teaching through their own further study on these coursebooks. They may agree to use other supplementary appropriate authentic materials in the course of their teaching. Raised interest in referring to and use of extensive samples of naturalistic conversation data in English teaching materials will benefit teachers and English learners.

### **5.2.2 To teaching material designers**

This research is another small case of authenticity analysis of coursebooks and the result is not surprising that it appears that the coursebooks do not represent real languages. Timmis (2010, p. 63) argues that there is never a natural transfer process from description to pedagogy. However, it may raise materials writers' and teachers' attention when the findings accumulate to a certain degree. Timmis (2010, p. 64) states that designing materials which reflect the findings of recent research into spoken language presents a considerable challenge for teachers and materials writers. Many researches have been focusing on the proposals or procedures to help teachers and materials writers to experiment with teaching authentic spoken language in the class. The study of spoken discourse ought to hold some interests for programme designers, coursebook writers,

examiners and teachers themselves.

### **5.3 Limitations of the research**

#### **5.3.1 Limitation of the sample authentic materials for comparison**

As it is not easy to find one-one matching contexts between coursebooks and authentic conversations, the contexts of authentic dialogues do not correspond to that of spoken texts in coursebooks. For example, the conversation in TED Talks has a nature of interview in a formal setting which might not match with the conversations in coursebooks in terms of relationship of participants. In addition, in order to use the same amount of words to do the comparison, the authentic data is from a few long dialogues with fewer topics and contexts than coursebooks. However, the authentic conversations selected for the study are unscripted with a variation of formal and informal settings.

#### **5.3.2 Limitation of the corpus CANCODE (5m)**

As the language items chosen for the survey is according to the research of the five-million-word CANCOCD spoken corpus (O’Keeffe et al., 2007, chapter two, three, seven and eight), the size of data is limited and might not be up to date. Tomlinson (2010, p. 88) argues one of the limitations of using corpus for teaching is they are only up to date at the moments of production of their utterances. The available corpora may not cover all the

areas and contexts of language use and the language in use might not be captured because language is in the process of change. The authentic data are chosen from a range of time period - two TED Talks clips were produced in 2017 and 2018 respectively, but the speech data from the IViE corpus was released in 2001. The corpus (IViE) data is not very new, but it is the convenient resource which can be found, and the sample conversations are completely matching with the purposes of the study.

### **5.3.3 Limitation of the size of the research**

The research only looked at a small number of language items and features of spoken language which might show a few aspects of the features of the research materials. In addition, it might 'provide only partial facts when the result is derived from being as the observer to look at instead of introspective of the materials' (Widdowson, 2000, p. 6).

## **5.4 For the future**

### **5.4.1 About application of authentic materials**

A range of variety of resources can be considered for creating integrated learning materials inside and outside class. Timmis' study (2002, p. 247) shows 63% of students (400 responses) appear to want to use the kind of spoken English of native speakers and two thirds of them are living and studying in an EFL context. Though nowadays, it may

not be the same case since over ten years have passed. However, it is evident that native speakers' speaking, or their recordings are interesting and attractive to English language learners. Tomlinson (2010, p. 90) states media is one of the useful resources in our daily lives where excellent authentic materials can be found very often. He illustrates how teachers can use the language which they encounter in their everyday lives as resources in their teaching (Tomlinson, 2010, p. 87).

#### **5.4.2 About transcription of authentic spoken language**

It is important to recognize the fact that there is not sufficient spoken discourse available for pedagogic purposes as some of the audios and videos might be available, but most of them are not with transcripts. Carter & McCarthy (1997) suggest transcribing the recordings captured for the use of spoken corpora development and spoken texts creations. Though the transcription job might be difficult, taking time and with limited resources, the exploration of its resources and effort on its application may be worth it due to the potential value for learners. If applied linguists, language researchers, teachers and learners can contribute to projects to collect large quantities of spoken data, that will benefit language teaching and all the people involved in the area such as researchers, teachers and learners. The ideal situation is to expose learners to the spoken texts in a wide range of varieties of genres and text types. Different types of talk produce different types of language which is not limited to the difference between formal and informal or public and private conversational discourse (Carter & McCarthy, 1997, p. 8). Another

consideration is to ensure texts are relevant and engaging for the learners because that is crucial to realizing the potential value of authentic texts (Tomlinson, 2017, p. 2)

### **5.4.3 About corpus and language in use**

Tomlinson (2010, pp. 89-102) states that there has been progress towards representing the real language in teaching materials, but there is a long way to go to expose learners to authentic English in teaching and learning which allows them to imitate real life language. To close the gaps of usage of language between typical textbook and typical real-life, teachers and materials developers need to consult corpora but not to rely on them. On the other hand, it is necessary to observe and use the accessible authentic language resources around us and notice the language changes because what we know about English might not be equal to how language is actually used. Teachers and materials developers can help learners in various ways to discover and acquire the language in use and the language strategies in real life that they want and need.



## References

- Adolphs, S. and R. Carter. (2013). *Spoken Corpus Linguistics: From Monomodal to Multimodal*. Routledge.
- Allwright, R. (1979). Language learning through communication practice. In C. Brumfit, & K. Johnson, *The communicative approach to language teaching*. Oxford: Oxford University Press, 167-182.
- Angouri, J. (2010). Using textbook and real-life data to teach turn taking in business meetings. In N. Harwood (Ed.), *English language teaching materials - theory and practice*. Cambridge: Cambridge University Press, 373-394.
- Basturkmen, H. (2001). Descriptions of spoken language for higher level learners: the example of questioning. *ELT Journal*, 55(3), 4-13.
- Berk, R. A. (2009). Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom. *International Journal of Technology in Teaching and Learning*, 5(1), 1-21.
- Breen, M. (1985). Authenticity in the language classroom. *Applied Linguistics* 6(1), 60-70.
- Bygate, M. (1987). *Speaking*. Oxford: Oxford University Press.
- Carmen, H. (2016). The Film in Language Teaching Association (FILTA): a multilingual community of practice. *ELT Journal*, 70(2), 190-199.
- Carter, R., & McCarthy, M. (1997). *Exploring spoken English*. Cambridge: Cambridge

- University Press.
- Carter, R. & McCarthy, M. (2006). *Cambridge grammar of English: a comprehensive guide to spoken and written grammar and usage*. Cambridge: Cambridge University Press.
- Carter, R., McCarthy, M., Mark, G., & O’Keeffe, A. (2016). *English grammar today*. Cambridge: Cambridge University Press.
- Cullen, R., & Kuo, I. (2007). Spoken grammar and ELT materials: a missing link? *TESOL Quarterly* 41(2), 361-386.
- Culpeper, J., Katamba, F., Kerswill, P., Woak, R., & MeEnery, T., (2009). *English language description, variation and context*. London: Palgrave Macmillan.
- Ellis, R. (2016). Language teaching materials as work plans: An SLA perspective. In B. Tomlinson (Ed.), *SLA research and materials development for language learning*. Oxon: Routledge, 203 – 218.
- Gilmore, A. (2004). A comparison of textbook and authentic interactions. *ELT Journal*, 58(4), 363–374.
- Gilmore, A. (2007). Authentic materials and authenticity in foreign language learning. *Language Teaching*, 40(2), 97-118.
- Grabe, E., Post, B. and Nolan, F. (2001). Modelling intonational Variation in English. The IViE system. In Puppel, S. and Demenko, G. (eds). *Proceedings of Prosody 2000*. Adam Mickiewicz University, Poznan, Poland. (Downloadable from [http://www.phon.ox.ac.uk/files/apps/old\\_IViE/#find](http://www.phon.ox.ac.uk/files/apps/old_IViE/#find))
- Hedge, T. (1993). Key concepts in ELT. *ELT journal*, 47(3), 275-277.

- Jones, C., & Horak, T. (2014). Leave it out! The use of soap operas as models of spoken discourse in the ELT classroom. *The Journal of Language Teaching and Learning*, 4(1), 1-14.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, 26, 22-63.
- Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan horse. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed.). Cambridge: Cambridge University Press.
- McCarthy, M. J. (2008). Profiling spoken fluency. *The Language Teacher*, 32(7), 32-34.
- McCarthy, M. J., & Carter, R. (1995). Spoken grammar: What is it and how can we teach it? *ELT Journal*, 49(3), 207-218.
- McCarten, J., & McCarthy, M. J., (2010). Bridging the gap between corpus and coursebook: The case of conversation strategies. In F. Mishan, & A. Chambers (Eds.), *Perspectives on language learning materials development*. Bern: Peter Lang, 11-32.
- McCarthy, M. J., & McCarten, M. (2018). Now you're talking! Practising conversation in second language learning. In C. Jones, (Ed.), *Practice in second language learning*. Cambridge: Cambridge University Press, 7-29.
- Mishan, F. (2005). *Designing authenticity into language learning materials*. Bristol: Intellect.
- Morrow, K. (1977). Authentic texts and ESP. In S. Holden (Ed.), *English for Specific purposes*. London: Modern English Publications, 13-17.

- O’Keeffe, A., McCarthy, M. J., and Carter, R. A. (2007). *From Corpus to Classroom*. Cambridge: Cambridge University Press.
- Peacock, M. (1997). The effect of authentic materials on the motivation of EFL learners. *ELT Journal*, 51(2), 144-156.
- Schmitt, N. (2010). *Researching Vocabulary*. Hampshire: Palgrave Macmillan.
- Siegel, J. (2019). Teaching lecture notetaking with authentic materials. *ELT Journal*, 73(2), 124-133.
- Tao, H. (2003). Turn in initiators in spoken English: A corpus-based approach to interaction and grammar. In Leistyna, P. and Meyer, C. F. (Eds), *Corpus Analysis: Language Structure and language Use*. Amsterdam: Rodopi, 187-207.
- Thornbury, S. (2005). *Beyond the sentence introducing discourse analysis*. Oxford: Macmillan Education.
- Thornbury, S. (2008). *How to teach speaking*. Harlow: Pearson Education.
- Timmis, I. (2002). Native-speaker norms and international English: A classroom view. *ETL Journal*, 56(3), 240-249.
- Timmis, I., (2010). Teachers telling tales: Exploring materials for teaching spoken language. In F. Mishan, & A. Chambers (Eds.), *Perspectives on language learning materials development*. Bern: Peter Lang, 63-85.
- Tomlinson, B. (Ed.). (2007). *Language acquisition and development: studies of learners of first and other languages*. London: Continuum.
- Tomlinson, B. (Ed.). (2008). *English language learning materials - A critical review*. London: Continuum.

- Tomlinson, B. (2010). Helping learners to fill the gaps in their learning. In F. Mishan, & A. Chambers (Eds.), *Perspectives on language learning materials development*. Bern: Peter Lang, 87-108.
- Tomlinson, B. (Ed.). (2016). *SLA research and materials development for language learning*. New York: Routledge.
- Tomlinson, B. (2017). Introduction. In A. Maley, & B. Tomlinson (Eds.), *Authenticity in materials development for language learning*. Newcastle upon Tynne: Cambridge Scholars Publishing.
- Tomlinson, B., & Masuhara, H. (2018). *The complete guide to the theory and practice of materials development for language learning*. Hoboken, NJ: Wiley-Blackwell.
- Watkins, J., & Wilkins, M. (2011). Using YouTube in the EFL classroom. *Language Education in Asia*, 2(1), 113–119.
- Widdowson, H. (1998). Context, community and authentic language. *TESOL Quarterly*, 32(4), 705–716.
- Widdowson, H. (2000). On the limitations of linguistics applied. *Applied Linguistics*, 21(1), 3-25.
- Wong, V., Kwok, P., & Choi, N. (1995). The use of authentic materials at tertiary level. *ELT Journal*, 49(4), 318-322.
- Wray, A. (2002). *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

## **Appendix 1 Sources and sample texts for coursebook conversations:**

Jin, L., and Cortazzi, M. (2016). *The New Standard College English (2<sup>nd</sup> Ed.)*. Beijing:  
Beijing Foreign Language Teaching and Research Press.

### **Sample conversations:**

Book1-U2

Food, glorious food!

K: Oh, this looks nice.

M: Cool.

W: Good afternoon, table for three? Come this way.

M: Thank you

J: Thank you.

W: The specials are on the board.

Kate: So, what sort of food do you like, Janet?

J: Well, I like spicy food. And I'm not very fond of raw food! What would you  
recommend?

M: Why don't you try the chicken curry? That's nice and spicy.

J: What's in it?

M: Chicken cooked in tomatoes and onions with Indian spices.

J: I'll try it. Do we all choose a selection of dishes to share or only one dish per person?

M: Usually one dish per person

K: Or the moussaka looks good.

J: What's it made with?

K: It's made with lamb and eggplant. It's a Greek dish.

J: How is it cooked?

K: It's baked in the oven.

J: Mm, that sounds good too.

K: And as a starter?

J: What's minestrone soup?

M: It's an Italian soup with vegetables and pasta. It's delicious!

J: OK, I'll have that.

K: Waitress?

W: What can I get for you?

K: Well, for the starter, can we have two minestrone soups, and for the main course, one moussaka and one curry, please. What about you, Mark?

M: I'll have the prawns with garlic and the chilli con carne. And could you bring us some water, please?

W: OK.

M: Thank you.

W: Thanks.

J: Thanks.

J: Thanks ... What's chilli con carne?

M: It's a spicy Mexican dish with beef and beans. It's very hot!

Book1-U5

News 24/7

M: This is just so crazy!

J: What?

M: This story I'm reading.

K: So tell us.

M: A man within a wheelchair (was) crossing the road in front of a lorry at some traffic lights. Somehow, the back of the wheelchair got stuck on the front of the lorry.

When the lorry started moving, it took the wheelchair and the man with it!

K: You're joking!

M: The driver drove for several miles at 80 kilometres an hour before he stopped at a garage. The man was unhurt because his seat belt had stopped him falling out.

J: What a terrible story! Thank goodness the man was all right!

M: The police asked the driver if he'd realized he'd had a passenger. The driver said he had no idea at all.

M: Do you want to hear another one? A funny one this time.

K: Go on.

M: A woman reported that her car had been stolen and that she'd left her mobile phone in the car. The policeman suggested calling the mobile. When he did, the thief



answered. The policeman told the thief that he was answering an ad in the paper and that he wanted to buy the car. And the thief agreed to sell it!

J: He didn't!

M: So they arranged to meet and the thief was arrested and the woman got her car back.

J: A happy ending!

M: You get these great stories in the paper – I always read them. it

Book1-U7

All you need is love

K: Hi, Becky, how's it going?

B: Good!

M: Guys, look, can you help me with a problem?

J: Yes, of course.

M: The thing is, there's this girl I really like called Jenny Sparks. She's a fresher, really stunning, reads history. I know her name because someone pointed her out to me, but I've never actually spoken to her. Do either of you know her?

K: No.

J: No, I don't know her. Mark, how can you like her if you haven't met her?

K: It's because she's absolutely gorgeous, Janet.

M: That's right! I want to ask her out, but first I've got to meet her. Got any

suggestions?

B: Guys! You want to order?

M: Sorry.

K: Three cappuccinos?

B: Sure.

J: Do you know anyone who knows her? You could ask them to introduce you.

M: No, I don't, that's the problem.

K: Are you matchmaking, Janet?

J: What's matchmaking?

K: Making introductions between people who might like each other. We don't do that here. How about just walking up to her and saying hi? Why don't you do that?

M: No.

K: Why not?

M: I'm not usually shy, but – she's so ... you know ...!

J: Oh, Mark!

J: I understand Mark completely.

K: Well, it's the only way he's going to get to talk to her.

M: OK, I'll give it a try.

B: Solved the Jenny problem yet?

J: Thank you.

K: Thanks.

M: Thanks.

J: You'll be fine. Mark. She'll like him, won't she, Becky?

B: Of course she will!

Book2-U2

Mixed feeling

K: Come in. Hey, Janet.

J: Hi Kate, are you busy?

K: Yes, I'm just doing an essay. But it's great to see you. So what's new?

J: Well, nothing much.

K: You look a bit fed up. What's bugging you?

J: Well, I had a phone call from my parents and it made me feel homesick. It happens every time they call, and it gets me down.

K: I'm sorry to hear that. I know how you feel, I love speaking to my mum and dad, but I always feel miserable after the call.

J: My dad doesn't say much, and I want to speak to him, but I wish I knew what to say.

K: Don't let it get to you. My dad doesn't say much on the phone either. I call, he answers the phone, and says, "Hi, I'll pass you to your mother." It's really irritating.

J: But I miss him and my mother a lot, and I like to hear his voice.

K: Just tell him what you're up to.

J: Sometimes I feel as if I made a mistake leaving home and coming to Oxford.

Sometimes I feel like a moody teenager.

K: Try not to worry about it, Janet. It's normal to feel like that. I understand how you feel, but I bet everything will be fine next term. You'll get used to it. Hey, why don't you do what I do?

J: What's that?

K: When my dad calls, I ask him for more money! He usually says no, but at least I get to hear his voice!

J: Maybe. I'm sorry to take up your time, Kate, but I must go now. Bye!

Kate: Wait a minute ... !

Book2-U5

Time Off

M: Have you got your tickets for the play?

K: What play?

M: The play which OUDS are producing. You know, the play I'm in at the Oxford Playhouse.

K: Oh, that play! Well, er ...

M: What about you, Janet?

J: What's the play called?

M: Waiting for Godot, by Samuel Beckett. You are coming, aren't you?

K: Beckett?

J: Why not?

K: Well, um, I'm sure you'll be totally brilliant, Mark ... but I wish I could understand the play. It doesn't make sense.

M: If only you were more patient, Kate Beckett's a fascinating writer. You'll come through, won't you, Janet? You really ought to see something like this at least once during your stay in Oxford.

J: Well, I'm not sure.

M: Oh, come on! Please!

J: But if Kate doesn't understand the play, there's no way I'll be able to follow it.

K: Do you want to go?

J: Well, I love going to the theatre, and I'd really like to see Mark acting. And actually, yes, I think I should see a play by Samuel Beckett.

M: Good! So you're coming, Janet. I wish you'd come, too, Kate. It's a really good performance.

K: Well, OK, but I'm only doing it because you're in it. When is it on?

M: Next Tuesday to Saturday.

J: How about going Friday night?

M: That's great. But you'd better get your tickets soon, because we're expecting a full house.

Book2-U7

The world at war

M: Hi!

K: Hi, Janet! Have you been waiting long?

J: Not at all. What did you think of Hero?

K: It was brilliant, thanks for suggesting it.

J: Well, it was nominated for an Oscar, you know.

K: That figures. It's a beautiful film.

M: Yes. The costumes and scenery were amazing.

K: I'd love to know more about the emperor. He was cool. Who was he?

J: Qin Shi Huang. It's said he was the first emperor in the history of China-he unified  
China.

K: Did he? When?

J: Er ... 221 BC.

M: As long ago as that!

W: Hi guys! What can I get you?

K: Yes, I'll have a coke, thanks.

M: Er ... Just a coffee.

W: Sure.

M: Tell us more

J: Um ... Well, before that there were seven big states and they have been fighting each  
other for many years.

M: Right.

J: It's called the Warring States Period. Anyway – Qin was king of the largest state and he defeated the six other states, one after another. It took him ten years to conquer them, each with a different strategy.

M: What kind of man was he?

J: Well, he was brilliant, obviously. And also wise. He had this huge army – they were very powerful. After his army had attacked the first state, the next state surrendered without much fight. They were so terrified.

K: Wow!

J: What else? The army leaders were very clever – they used a river to flood a city.

M: That can't have been easy.

J: Yes, anyway, after conquering the last state, Qin made himself Emperor of the whole of China.

M: Was he the emperor who created the Terracotta Warriors?

J: That's right. He was so afraid of death that he wanted them to guard him in the afterlife.

K: Fascinating!

## Appendix 2 Sources and transcripts for conversations (IViE):

[http://www.phon.ox.ac.uk/files/apps/old\\_IViE/index.html](http://www.phon.ox.ac.uk/files/apps/old_IViE/index.html)

### Transcripts of conversations (IViE)

Cambridge\_Conversations\_clp

I Overlap utterances

+ where interruption occurred, or utterance resumed

= unfinished or truncated words

- Short pause

[ ] Laughter or other sounds

S1: Have you seen some of the latest billboard adverts for cigarettes?

S2: Uhm, I haven't seen any of them lately – |what about you?

S1: |We've+

S1: + Well, we've got lots up in the media studies room with| the +

S2: | all right

S1: + silk cut ones with the purple strange bits|with white waves and stuff like that –  
and - they

S2: |mhm

S1: look like really really nice adverts but I don't think there should be quite so many  
magazines and stuff.

S2: No, but - the fact they are so obscure it doesn't clearly really like clearly say - have



a cigarette does it?

S1: No, but - well those ones aren't as bad and they've got, uhm, a real strange one on Chesterton Road - which is like up on the big things you - drive past it - but - they are quite obscure - but - I think it should be - be a bit bigger where it says sort of like spanking homes or |and stuff like that.

S2: |Yeah,  
definitely

S1: Because there's so much crap in them then it does pretty +

S2: I don't think the adverts really influence like the ones the obscure ones I mean like the ones where they show like a man who's a model smoking |and like acting really cool, I think they're the

S1: |Yeah, yeah,

S2: ones that influence |people because I don't think it's really like the adverts that make people

S1: |yeah

S2: smoke.

S1: yeah, and so many pictures of people smoking is where I imagine stuff like that =

S2: |Yeah, exactly

S1: Is like with the top models | their embassy and| whatever +

S2: |Yeah, |exactly

S1: + It gives people the wrong idea. It's like you get thin by smoking and you end up

looking like me if you do.

S2: No

S1: I don't think that should be in there. I think that's really bad.

S2: Do you smoke?

S1: No, do you?

S2: No

S1: Do you like smoking hut?

S2: Never been there

S1: It is so dirty

S2: Really?

S1: Unbelievable!

S2: What's it's like? I've never been in

S1: It's like, uhm, you know, corrugated iron

S2: Yeah

S1: It's like corrugated iron round it and - one bit's like spray painted.

S1: And you don't walk on the ground, you walk on the fag bach a it's like so horrible and you know we have to actually clean it. You're in litter duty.

S2: Yeah

S1: There is a point when we have to clean it and I think it's horrible and I think any people who smoke should have to clean it.

S2: Definitely

S1: I think that would be really fair or a lot fairer, be fair to the one in Long road because their's is just like, have you been to Long Road?

S2: No

S1: Their smoking area is like - a concrete area - stuck in the middle of the field and that's it.

S2: They're changing it though aren't they. They're changing the smoking hut

S1: What? at Long Road?

S2: No, here

S1: What?

S2: They're building a new one

S1: Are they?

S2: They've got a new one with a building and they are taking the old one down I think

S1: Are they?

S2: I think that's what Mr Holmes's saying

S1: Who's he?

S2: A teacher

[Laughter]

S1: I don't know anything about it. I haven't heard anything about it, er I think it does really really damage your health though.

S2: definitely.

S1: Cos my mum used to smoke when I was younger and I've never had a problem with people smoking around me or anything even though they're dying but I used to get really really cross with my mum and do things like throw her cigarettes over the back fence and things like this. She used to get so cross with me. It was unbelievable.

S2: She stopped now

S1: Yeah yeah she stopped about 5 or 6 years ago but she used to get really really cross with me for doing it though - and they are so expensive as well

S2: Yeah, that's a big thing

S1: And I think I couldn't truly be able to afford to smoke

S2: Definitely not

S1: I just don't have that sort of money available so - and especially people who smoke like 10, 20 a day. Completely outrageous cos my boyfriend smokes and he doesn't spend a lot of money on it, but still it all adds up

S2: Yeh definitely

S1: That means

old\_IViE\_Cambridge\_Conversation\_cma

A: so, have you seen much of the news lately then?

B: uhm, well, which part?

A: oh, a bit much about tobacco advertising

B: oh, you mean Bernie Ecclestone and it's a +

A: well, I wouldn't know about Bernie Ecclestone, never heard of him. What is it anyway?

B: oh, it's just - you know (?) donation to the labour party, |and got tobacco advertising banned.

A: |oh,

B: and banned |for a while, but would it affect you? |Tobacco advertising

A: |oh, I think I know it |affect me?

A: Well it would erm, cut down the erm, well obviously it would cut down the uhm advertisements for tobacco's, in turn, uhm might cut down the amount of people buying it, so, uhm, from that, tobacco companies might think, well, you know, I will make tobacco a bit cheaper so that more people can get onto it.

A: so maybe it will be a good thing for me because I'm a smoker and so they won't reduce the price of tobacco to uhm increase the number of uhm smokers.

B: yeah, but would but would advertising encourage them to smoke because isn't it just a conscious decision that if they won't smoke, they +

A: well obviously it is but something has got to instigate the conscious decision, if doesn't just come from nowhere +

B: yeah

A: + but usually when you start smoking, its actually through a friend uhm rather than through an advertisement because you hardly ever look at adverts. So when you are watching TV, you just watch the program and when the adverts come on, you switch

off, don't you?

B: yeah

A: you just, uhm, I don't know

B: so it's, it's, it's more friends that pressure you but rather than just seeing roughmans on the side of a car that's gonna make you smoke

A: what exactly but if you are a big fan of uhm of uhm the grand prix or whatever then you are going to take more interest in it if you like aren't you? I mean erm erm |must think about that

B: |so+

A: one.

B: +so you are saying it's the fact that if someone has got roughmans on the side of the car and you follow them diligently you're gonna smoke roughmans

A: well if you're a big fan of the driver if you are a big fan of the thing in general , then you will want to gonna you get fanatical people that just want to be as much a part of the grand prix as possible , as you can. don't know why anyone would smoke roughmans though.

B: no, not your brand then

A: no, I smoke roll ups

B: but isn't it just the fact that if you have advertising, it's = it's just keeping the awareness of that brand between the people who already smoke them.

A: between the people who already smoke them. Well, then causing people to smoke them. Cos, you know, if you, if you see an ad for, if they stuck an ad for Barbie on the side of it. It's not going to make hundreds of thousands of people run out n buy Barbie just because it's on the side of a car.

B: Well, that's different from how I'm thinking about it.

A: But it, but it's still in a way the same thing. Even if they take it off. The awareness of cigarettes is going to be there. So, it +

B: well, no, it = it all goes to increase the image of smoking, really, doesn't it? If you get, if you get an advertisement on the side of big fast cars. Then everyone thinks, "Crist, that was just cool, wasn't it? You know? I want to try that out" It = it just, it will just raise the profile, like Channel 4 used to have really dull programs. Then decided to put things like Friends, and Frasier on there (?). It's an image, just went out. As to Tango ??? they may, an ad for it came on, with the funny man and things flying here and there, and whatever. It just goes to - raise the awareness and the profile of it, even if the thing is so much a bad thing +

A: Yeah, but on the +

B: + a dirty

A: + but is that really going to affect them. 'Cos something like smoking, a lot of people either choose to smoke, because it's, you know, the parents, you know, it's a rebel= it's a

rebellion against the parents, isn't it? Because they grow up to teach you, you know, don't start, like, smoking, it's bad for you to stop. So you do it, originally, just to rebel= rebel against your parents.

B: But if that one ???

A: I'm sorry, it's the advertising, it's not going to... You know what I mean? 'Cos if you are going to smoke, and you have decided you are going to smoke, you know, you are not going to see an ad and say, "Um, I'm going to smoke."

B: So, say you are a 10 year old kid, I mean, you are easily influenced. If you were to see something like, oh, these people, the age adults, see you look up to and respect and listen to, and they are all and - they've got pictures of fags, fag brands on their car, wherever. Then you are not going to think.



### **Appendix 3 Sources and transcripts for Ted Talks:**

1) The future we're building — and boring

[https://www.ted.com/talks/elon\\_musk\\_the\\_future\\_we\\_re\\_building\\_and\\_boring?language=en](https://www.ted.com/talks/elon_musk_the_future_we_re_building_and_boring?language=en)

2) How Netflix changed entertainment — and where it's headed

[https://www.ted.com/talks/reed\\_hastings\\_how\\_netflix\\_changed\\_entertainment\\_and\\_where\\_it\\_s\\_headed?language=en](https://www.ted.com/talks/reed_hastings_how_netflix_changed_entertainment_and_where_it_s_headed?language=en)

### **Transcripts of sample conversations of Ted Talks:**

Elon Musk · Chris Anderson

The future we're building — and boring

TED2017 | April 2017

10:20 CA: So you've started a new company to do this called The Boring Company.

Very nice. Very funny.

10:26 (Laughter)

10:28 EM: What's funny about that?

10:29 (Laughter)

10:32 CA: er, how much of your time is this?

10:35 EM: It's er, it's maybe ... two or three percent.

10:40 CA: You've called it a hobby. This is what an Elon Musk hobby looks like.

10:45 (Laughter)

10:46 EM: I mean, it really is, like – er, This is basically interns and people doing it part time. We bought some second-hand machinery. Er It's kind of puttering along, but it's making good progress, so --

11:03 CA: So an even bigger part of your time is being spent on electrifying cars and transport through Tesla. erm Is one of the motivations for the tunneling project the realization that actually, in a world where cars are electric and where they're self-driving, there may end up being more cars on the roads on any given hour than there are now?

11:25 EM: Yeah, exactly. A lot of people think that when you make cars autonomous, they'll be able to go faster and that will alleviate congestion. And to some degree that will be true, but once you have shared autonomy where it's much cheaper to go by car and you can go point to point, er the affordability of going in a car will be better than that of a bus. Like, it will cost less than a bus ticket. So er the amount of driving that will occur will be much greater with shared autonomy, and actually traffic will get far worse.

12:03 CA: You started Tesla with the goal of persuading the world that electrification was the future of cars, and a few years ago, people were laughing at you. Now, not so much.

yes

12:16 EM: OK.

12:17 (Laughter)

12:19 I don't know. I don't know.

12:21 CA: But isn't it true that pretty much every auto manufacturer has announced serious electrification plans for the short- to medium-term future?

12:31 EM: Yeah. Yeah. Er I think almost every automaker has some electric vehicle program. They vary in seriousness. Er Some are very serious about transitioning entirely to electric, and some are just dabbling in it. And some, amazingly, are still pursuing fuel cells, but I think that won't last much longer.

12:53 CA: But isn't there a sense, though, Elon, where you can now just declare victory and say, you know, "We did it." Let the world electrify, and you go on and focus on other stuff?

13:05 EM: Yeah. I intend to stay with Tesla as far into the future as I can imagine, and there are a lot of exciting things that we have coming. Obviously the Model 3 is coming soon. We'll be unveiling the Tesla Semi truck.

13:23 CA: OK, we're going to come to this. So Model 3, it's supposed to be coming in July-ish.

13:30 EM: Yeah, it's looking quite good for starting production in July.

13:35 CA: Wow. Er One of the things that people are so excited about is the fact that it's got autopilot. And you put out this video a while back showing what that technology would look like. Yeah

13:49 EM: Yeah.

13:51 CA: There's obviously autopilot in Model S right now. What are we seeing here?

13:55 EM: Yeah, so this is using only cameras and GPS. So there's no LIDAR or radar

being used here. This is just using passive optical, which is essentially what a person uses. Er The whole road system is meant to be navigated with passive optical, or cameras, and so once you solve cameras or vision, er then autonomy is solved. If you don't solve vision, it's not solved. So that's why our focus is so heavily on having a vision neural net that's very effective for road conditions.

14:34 CA: Right. Many other people are going the LIDAR route. You want cameras plus radar is most of it.

14:40 EM: You can absolutely be superhuman with just cameras. Like, you can probably do it ten times better than humans would, just cameras.

14:48 CA: So the new cars being sold right now have eight cameras in them. They can't yet do what that showed. When will they be able to?

14:59 EM: er I think we're still on track for being able to go cross-country from LA to New York by the end of the year, fully autonomous. er

15:09 CA: er OK, so by the end of the year, you're saying, someone's going to sit in a Tesla without touching the steering wheel, tap in "New York," off it goes.

15:20 EM: Yeah.

15:21 CA: Won't ever have to touch the wheel -- by the end of 2017.

15:25 EM: Yeah. Essentially, November or December of this year, we should be able to go all the way from a parking lot in California to a parking lot in New York, no controls touched at any point during the entire journey.

15:40 (Applause)

15:42 CA: Amazing. But part of that is possible because you've already got a fleet of

Teslas driving all these roads. You're accumulating a huge amount of data of that national road system.

22:49 EM: Yeah. Solar glass tiles where you can adjust the texture and the color to a very fine-grained level, and then there's sort of microlouvers in the glass, such that when you're looking at the roof from street level or close to street level, all the tiles look the same whether there is a solar cell...

Reed Hastings · TED2018

TED2018 | April 2018

How Netflix changed entertainment — and where it's headed

03:29 CA: And so you added all these other remarkable series, "Narcos," "Jessica Jones," "Orange is the New Black," "The Crown," "Black Mirror" -- personal favorite -- er "Stranger Things" and so on. And so, this coming year, the level of investment you're planning to make in new content is not 100 million. It's what?

03:51 RH: It's about eight billion dollars around the world. And it's not enough. er There are so many great shows on other networks. And so we have a long way to go.

04:03 CA: But eight billion -- that's pretty much higher than any other content commissioner at this point?

04:11 RH: No, Disney is in that realm, and if they're able to acquire Fox, they're even bigger. And er then, really, that's spread globally, so it's not as much as it sounds.

04:24 (Laughter)

04:26 CA: But clearly, from the Barry Dillers and others in the media business, it feels like from nowhere, this company has come and has really revolutionized the business. It's like, as if Blockbuster one day said, "We're going to make Blockbuster videos," and then, six years later, was as big as Disney. I mean, that story would never have happened, and yet it did.

04:47 RH: That's the bitch about the internet -- it moves fast, you know? Everything around us moves really quick.

04:53 CA: I mean, there must be something unusual about Netflix's culture that allowed you to take such bold -- I won't say "reckless" -- bold, well thought-through decisions.

...

07:26 CA: So you just wake up and read them on the internet.

07:29 RH: Sometimes.

07:30 CA: "Oh, we just entered China!"

07:32 RH: Yeah, well that would be a big one.

07:34 CA: er But you allow employees to set their own vacation time, and ... There's just --

07:42 RH: Sure, that's a big symbolic one, vacation, because most people, in practice, do that, anyway. er But yeah, there's a whole lot of that freedom.

07:53 CA: And courage, you ask for as a fundamental value.

07:58 RH: Yeah, we want people to speak the truth. And we say, "To disagree silently is disloyal." Er It's not OK to let some decision go through without saying your piece, and typically, writing it down. And so we're very focused on trying to get to good decisions through the debate that always happens. And we try not to make it intense, like yelling at each other -- nothing like that. You know, it's really curiosity drawing people out.

08:26 CA: You've got this other secret weapon at Netflix, it seems, which is this vast trove of data, a word we've heard a certain amount about this week. You've often taken really surprising stances towards building smart algorithms at Netflix. Back in the day,

you opened up your algorithm to the world and said, "Hey, can anyone do better than this recommendation we've got? If so, we'll pay you a million dollars." You paid someone a million dollars, because it was like 10 percent better than yours.

08:55 RH: That's right.

08:56 CA: Was that a good decision? Would you do that again?

08:58 RH: Yeah, it was super exciting at the time; this was about 2007. But you know, we haven't done it again. So clearly, it's a very specialized tool. And so think of that as a lucky break of good timing, rather than a general framework. Er So what we've done is invest a lot on the algorithms, so that we feature the right content to the right people and try to make it fun and easy to explore.

09:23 CA: And you made this, what seems like a really interesting shift, a few years ago. You used to ask people, "Here are 10 movies. What do you think? Which ones of these are your best movies?" And then tried to match those movies with recommendations for what was coming. And then you changed away from that. Talk about that.

...



11:29 CA: But -- yes, indeed. But isn't it the case that algorithms tend to point you away from the broccoli and towards the candy, if you're not careful? We just had a talk about how, on YouTube, somehow algorithms tend to, just by actually being smarter, tend to drive people towards more radical or specific content. It'd be easy to imagine that Netflix algorithms, just going on revealed values, would gradually --

11:56 RH: Right, get too base --

11:57 CA: We'd all be watching violent pornography or something. Or some people would, you know. But, how --

12:04 (Laughter)

12:06 Not me! I'm the child of a missionary...